Helena R. Batalha[1], David J. Wright[1], Iain Barr[1], Nigel J. Collar[2,3], and David S. Richardson[1]

Genetic diversity and divergence in the endangered Cape Verde warbler *Acrocephalus brevipennis*

1. School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich, Norfolk, NR4 7TJ, UK
2. BirdLife international, Pembroke Street, Cambridge CB2 3QZ, UK
3. Bird Group, Department of Life Sciences, Natural History Museum, Tring, Herts HP23 6AP, UK

Corresponding author: David S. Richardson

Email: david.richardson@uea.ac.uk

Phone: (44) 01603 591496

Fax: (44) 01603 592250

Abstract

Genetic factors play an important role in the long-term persistence of populations and species, and conservation strategies should take such factors into account. We use neutral molecular markers to assess diversity and divergence between the three remaining island populations of a little-studied endemic passerine, the Cape Verde warbler, *Acrocephalus brevipennis*. Variation at both microsatellite loci and the cytochrome b gene reveal low diversity within the species overall, but considerable divergence among the populations on Santiago, Fogo and São Nicolau islands. The genetic markers show a gradient of genetic diversity with population size, with the smallest population of S. Nicolau being the least diverse, and the largest, Santiago, the most diverse. The more geographically isolated population on S. Nicolau is also more divergent from Santiago and Fogo than these two are from each other. The Cape Verde warbler diverged from its sister species, the greater swamp warbler *Acrocephalus rufescens*, within the last million years, and the three populations became isolated from one another 165,000–199,000 years ago. There is also evidence of population bottlenecks, especially in the smallest and most isolated population of S. Nicolau. This population seems to have decreased during the last century, with potential for further decreases and even extinction. As the three populations are genetically distinct, with no evidence of gene flow between them, we argue they should be treated as separate management units for the successful conservation of this species.

Keywords: microsatellites, cytochrome b, island passerine, endemic species, management units, Macaronesia

Introduction

Over the last few centuries many once common species have become threatened or extinct throughout the world, mainly due to anthropogenic impacts (Baillie et al. 2004; Butchart et al. 2010). Many of these declines and extinctions have occurred on islands (Steadman 1995; Brooks et al. 2002; Sax & Gaines 2008). Island species tend to be more vulnerable to extinction threats than mainland species for many reasons (Frankham 1998; Sax & Gaines 2008), including smaller population and range sizes, increased vulnerability to stochastic events, evolutionary/prey naivety, lower levels of genetic diversity and decreased gene flow (Frankham 1998; Frankham & Ralls 1998; Griffin et al. 2000; Sax & Gaines 2008; Frankham et al. 2009; Sih et al. 2010). Islands and archipelagos are often the focus of conservation actions not only because they contain threatened species but also because they can be biodiversity hotspots due to their high levels of endemism (Myers et al. 2000; Emerson 2002; Kier et al. 2009). Genetic factors can contribute significantly to extinction risk in the wild (Frankham & Ralls 1998; Saccheri et al. 1998; Spielman et al. 2004; Frankham 2005). Higher levels of genetic variation within a population provide more potential to adapt, for example to environmental changes and novel predators and parasites, whereas small, genetically depauperate populations are disadvantaged in this regard, and can also suffer from inbreeding depression (Crnokrak & Roff 1999; Hedrick & Kalinowski 2000; Brook et al. 2002). Such factors can drive small populations into an extinction vortex (Gilpin & Soulé 1986). Consequently, where possible, management plans for threatened species should take genetic factors into account (Brook et al. 2002; Frankham 2005).

Genetic data can be used to resolve taxonomic uncertainties, define management units and assess the extinction risk of a species, being therefore important in informing conservation and helping mitigate extinction (Frankham et al. 2009). This is especially important when populations have become isolated from one another, or fragmented (Moritz 1994b; Fouquet et al. 2010; Garcia-del-Rey et al. 2013). In an attempt to develop a framework for prioritizing conservation efforts, the terms 'evolutionarily significant units' (ESUs) and conservation 'management units' (MUs), mostly defined by genetic criteria, were coined (Ryder 1986; Moritz 1994a). Although there are many different definitions (see Fraser & Bernatchez 2001; Funk et al. 2012 for reviews), in a genetic framework ESUs should be reciprocally monophyletic for mitochondrial DNA and significantly divergent at nuclear loci (Moritz 1994b; Avise 2004), whereas the concept of MUs is applied to populations which are significantly divergent at nuclear or mitochondrial loci (Ryder 1986; Moritz 1994a). In other terms, ESUs identify historically isolated populations, and MUs can be defined as populations which show incomplete phylogenetic divergence but currently lack gene flow (Moritz 1994b). Although such definitions are not necessarily equivalent to subspecies (Moritz 1994b), they provide a first quantitative basis for defining conservation priorities below the species level (Funk et al. 2012).

Different types of genetic markers can have different mutation and recombination rates or transmission pathways, thus they can reveal different aspects of a species' evolutionary history (Avise 2004). Microsatellites have fast mutation rates (Selkoe & Toonen 2006), and can be used to infer current population structure, demography and gene flow (e.g., Illera et al. 2007; Dool et al. 2013; Ferchaud et al. 2015; Xenikoudakis et al. 2015). Furthermore, because of their high levels of polymorphism they are also useful for assessing demographic bottlenecks (Sunnucks 2000; Garza & Williamson 2001; Williamson-Natesan 2005). In contrast, mitochondrial DNA has slower mutation rates, and is commonly used to infer population history dating to, and before, the Last Glacial Maximum (Avise et al. 1987) circa 23 Kya ago (Clark et al. 2009). Combining data from microsatellites and mitochondrial DNA can provide a clearer picture of diversity within, and divergence among, populations, and can help reconstruct their relative demographic and phylogenetic histories (Reynolds et al. 2015). Additionally, patterns of variation at neutral markers can provide information about colonisation, gene flow, phylogeny and historical relations between populations (e.g., Hille et al. 2003; Dool et al. 2013; Dutta et al. 2013; Ferchaud et al. 2015; Padilla et al. 2015). Finally, genetic markers can be used to assess whether a lack of diversity may be contributing to current declines in the population (Allendorf & Luikart 2007; Frankham et al. 2009) and to determine if augmentation (Weeks et al. 2011) may be required to alleviate current problems, or ensure maximal adaptive potential is retained in the species for the future.

The Cape Verde warbler *Acrocephalus brevipennis* is an endangered passerine endemic to Cape Verde, a volcanic archipelago ca. 500 km off West Africa, between 14–18°N and 22–26°W in the Atlantic Ocean (Fig 1). It is most closely related to the greater swamp warbler *Acrocephalus rufescens* of sub-Saharan Africa (Leisler et al. 1997; Fregin et al. 2009). The Cape Verde warbler was discovered about 150 years ago on the island of São Nicolau (Keulemans 1866), and in the late 19th century it was known to exist at least on Santiago, S. Nicolau and Brava (Hazevoet 1993). During the 20th century the population decreased, possibly due to habitat loss caused by droughts (Hazevoet 1995) and by 1995 the bird was thought to be confined to Santiago (Hazevoet 1995; BirdLife International 2016). However, in 1998 the bird was rediscovered on S. Nicolau (Hazevoet et al. 1999), where the most recent published estimates suggested just 8–10 breeding pairs remain (Hazevoet et al. 1999; Donald et al. 2004). In 2004, another population, estimated at ca. 500 breeding pairs, was found on Fogo (Hering & Hering 2005; Hering & Fuchs 2009). Fogo and Santiago, in the southern Sotavento group of islands, are separated by ca. 60 km, whereas S. Nicolau, in the northwestern Barlavento group, is separated from the former two islands by more than 160 km (Fig 1). Such oceanic barriers can prevent gene flow in *Acrocephalus* warblers (Komdeur et al. 2004; Cibois et al. 2008), especially between the Sotavento and Barlavento groups (Hazevoet 1995; Hille et al. 2003), which could have important implications for this species' conservation. Currently, the Cape Verde warbler's global population is estimated at a maximum of 2,200–3,000 individuals, and is thought to be decreasing (BirdLife International 2016), but lack of census information means that such estimates may be somewhat unreliable (BirdLife International 2016). The species has legal protection status in Cape Verde but to date no countrywide conservation strategy has been developed. An understanding of the genetic variation within, and divergence between, the three warbler populations is clearly important to the success of any future conservation programme (Frankham et al. 2009; McCartney-Melstad & Shaffer 2015).

Here we use neutral molecular markers (mtDNA and microsatellites) to study the three known remaining populations of the Cape Verde warbler. Our aims were to clarify the degree of genetic divergence between the bird and its sister species, the greater swamp warbler, to assess genetic relationships between Cape Verde warbler populations, and to measure genetic diversity within each population so that this knowledge can be used to inform future conservation actions.

Materials and methods

Sample collection and DNA extraction

Samples were collected from November 2013 to January 2014, and September to November 2014, the main breeding season (Hazevoet 1995). We surveyed for warblers across all potential suitable habitat types and altitudes (the maximum altitude where a bird was sampled was 1,384 m above sea level) on each of the three islands to ensure that coverage was as comprehensive as possible (Fig 1, Table S1). We also conducted extensive transect and point counts in January 2014 on Brava (2 people x 4 days), where the bird was present until 1969 (Hazevoet 1995), and in September 2014 on Santo Antão (2 people x 7 days), where the vegetated valleys were thought to provide suitable habitat (Hazevoet 1995). When detected, birds were attracted with conspecific song playback and caught with mist nets. Each caught bird was fitted with a unique combination of a numbered Cape Verde metal ring, and three UV-resistant plastic colour rings. We determined if the birds were adults or juveniles based on previously described criteria for this and other warbler species (Komdeur 1991; Cramp & Perrins 1992; Baker 1997), and recorded each sampling location with a Garmin eTrex® H GPS. A small ca. 40 µl blood sample was collected from each bird by brachial venipuncture and stored at room temperature in 800 µl absolute ethanol. We aimed to sample a minimum of 25–30 unrelated individuals per island, as this is sufficient for accurate estimation of genetic parameters in microsatellite analyses (Hale et al. 2012). In addition, we used seven other Cape Verde warbler samples from S. Jorge dos Órgãos (central Santiago), seven greater swamp warbler samples (one from Senegal and six from Nigeria) and three Seychelles warbler *Acrocephalus sechellensis* samples (details on additional samples, sampling sites, dates, and bird age determination in Supplementary Material and Table S1). DNA was extracted using a salt extraction protocol following Richardson et al. (2001). Individuals were molecularly sexed following Griffiths et al. (1998).

Molecular markers

Cytochrome b

We amplified and sequenced a 1,150 bp fragment of mtDNA comprising the last 13 nucleotides of the ND5 gene, 1,128 nucleotides of the cytochrome b, and 9 nucleotides in between the two genes (hereafter "cytochrome b", for simplicity). To avoid amplifying nuclear copies of the mitochondrial DNA (numts; Sorenson & Quinn 1998), we re-designed the primers ND5 and mt-F from Helbig et al. (2005). We used BLAST (Altschul et al. 1997) to map the primer sequence against the mitochondrial genome of the Eurasian reed warbler *Acrocephalus scirpaceus*, (accession number AM889139, Singh et al. 2008) in GenBank (Benson et al. 2013) and re-designed them to match that genome. We checked the redesigned primer pair (ND5-Acro 5'-GGCCTAATCAAAGCCTAC-3' and mt-F-Acro 5'-GGCTTACAAGACCAATGTTT-3') for compatible annealing temperatures, hairpins, complementarity, primer-dimer formation and % GC content using Primer3 (Koressaar & Remm 2007; Untergasser et al. 2012). Polymerase chain reactions (PCRs) were performed in 15 µl total volume, consisting of 7.5 µl TopTaq Master Mix (Qiagen, West Sussex, UK), 0.6 µl (10 µM) each primer, 4.8 µl ddH2O and 1.5 µl DNA. Thermal cycle conditions were 94°C for 3 min, followed by 40 cycles of 94°C for 30 s, 55°C for 45 s and 72°C for 1 min, with a final extension step at 72°C for 10 min and a cooling step at 20°C for 5 min. PCR products were quantified in a 1.5% agarose gel and inspected for signs of non-specific annealing, i.e. ghost bands (Sorenson & Quinn 1998; Bensasson et al. 2001). Successfully amplified products were cleaned with 5 µl of Exo and FastAP (Life Technologies, UK) to 10 µl of PCR product, and sequenced in both directions (Eurofins, Germany). PCRs were repeated twice to confirm unique haplotypes and to resolve any ambiguous sites.

Microsatellites

Individuals were genotyped at 17 polymorphic microsatellite loci, using multiplex combinations adapted from Spurgin et al. (2014; Table S2), using primers designed for the Seychelles warbler (Richardson et al. 2000). Each PCR included 1 µl Qiagen PCR multiplex master mix, 1 µl primer mix and 1 µl pre-dried DNA. The PCR cycling conditions were 15 min at 95°C, followed by 30 cycles of 30 s at 94°C, 1 min 30 s at 55 or 56°C and 1 min at 72°C, followed by a final extension step of 30 min at 60°C. PCR products were separated on an ABI 3730 DNA analyser (Applied Biosystems, UK), and allele sizes were assigned using GeneMapper 4.0 software (Applied Biosystems, UK). Genotypes were checked visually. PCRs were repeated twice to confirm unique alleles. In cases where two or more amplifications resulted in conflicting genotypes for the same individual we considered the individual as heterozygous. We calculated microsatellite error rate per locus according to Pompanon et al. (2005). One of the markers, Ase6, could not be reliably scored due to stutter, so it was excluded from further analyses.

Data analyses

Cytochrome b

Unless stated otherwise, statistical analyses were performed in R v. 2.14.1 (R Core Team 2016). Sequences were visually inspected in FinchTV v. 1.4.0 (Geospiza Inc., Seattle, WA, USA) and manually aligned in BioEdit v. 7.0.9.0 (Hall 1999), against partial CDS sequences of a Cape Verde and a greater swamp warbler (from Senegal) previously published (Fregin et al. 2009) – accession numbers FJ883026 and FJ883037, respectively. The sequences were translated in MEGA v. 6 (Tamura et al. 2013), and checked for stop codons. We calculated the number of haplotypes and segregating sites, haplotype and nucleotide diversity, Tajima's D (Tajima 1989) and Fu's FS (Fu 1997) in DnaSP v. 5 (Librado & Rozas 2009), and pairwise F_{ST} values (Wright 1951; Weir & Cockerham 1984) in Arlequin v. 3.5 (Excoffier & Lischer 2010). Significance of F_{ST} values was tested with 10,000 permutations. We calculated mean genetic distances within and among the three populations of Cape Verde warbler and the greater swamp warbler in DnaSP, using the pairwise distance model. A phylogenetic tree was constructed using the maximum likelihood method implemented in MEGA v. 6, based on the Tamura-Nei model of evolution (Tamura & Nei 1993) as selected by the Bayesian Information Criterion in the same software. We used 1,000 bootstrap replicates as a test of phylogeny, and a discrete Gamma distribution with four categories to model evolutionary rate differences among sites. However, haplotype networks are usually more appropriate to visualise relationships between haplotypes in intraspecific datasets than phylogenetic trees (Bandelt et al. 1999; Posada & Crandall 2001). Therefore, a median joining haplotype network was also built in Network v. 4.6.1.3 (Bandelt et al. 1999).

Time to most recent common ancestor

Time to most recent common ancestor (tMRCA) of the Cape Verde and greater swamp warblers was estimated using a Bayesian phylogenetic approach in BEAST v.2.3.0 (Drummond et al. 2012; Bouckaert et al. 2014). We estimated the best-fit model of nucleotide substitution in jModelTest v. 2.17 (Guindon & Gascuel 2003; Darriba et al. 2012), which tests for the best model of evolution using the Bayesian Information Criterion (BIC), Akaike Information Criterion (AIC) and decision theory (DT). AIC selected GTR+I whereas BIC and DT selected HKY+G. We therefore used the HKY model of substitution (Hasegawa et al. 1985) with a gamma site model comprising four rate categories (Yang 1994) and an estimated gamma shape parameter. Effective sample sizes (ESS) for all model parameters exceeded the recommended minimum 200 when using the HKY model, so this model was considered appropriate. The dataset was composed of unique haplotypes of Cape Verde and greater swamp warblers, using the Seychelles warbler as an outgroup. Codons were partitioned into 3 (1 + 2 + 3, frame 2) to allow each nucleotide position within the codon to be modelled separately. Empirical substitution model state frequencies were used along with an estimated HKY kappa parameter. We used a substitution rate of 0.0105 per site per million years, which corresponds to a 2.1% divergence rate, generally applicable to passerine cytochrome b sequences (Weir & Schluter 2008). Molecular clock rate variation is highly unlikely in our dataset, which comprises a single gene across closely related species with resulting low sequence information content. We therefore employed a strict molecular clock. As neither fossil nor relevant biogeographical calibration points were available for this clade of passerines, we utilised published molecular phylogenies to provide loose bounds for the priors. We specified a uniform distribution between 1 and 12 Mya for the tMRCA of the Cape Verde, greater swamp and Seychelles warblers group, as 12 Mya is the estimated divergence time of the small and large groups of *Acrocephalus* warblers; Cape Verde, greater swamp and Seychelles warblers are members of the latter group (Price 2008; Leisler & Schulze-Hagen 2011). We used a diffuse uniform prior bound between the present and 1 Mya to estimate tMRCA of Cape Verde and greater swamp warblers and the same for the tMRCA of the three Cape Verde warbler populations. This is because, even though a divergence time has never been proposed, previous phylogenetic analyses suggests that these two sister species are so similar, when compared to the dated splits in the *Acrocephalus* phylogeny, that they are not likely to have diverged more than 1 Mya (Leisler et al. 1997; Fregin et al. 2009; Leisler & Schulze-Hagen 2011). Finally, the birth rate and clock rate priors were both assigned diffuse gamma distributions (α = 0.001, β = 1000) (Drummond & Bouckaert 2014). The Markov chain Monte Carlo (MCMC) simulations were run with a standard 10%

burn-in. We did four runs, with a total of 100 million iterations, to check mixing and convergence, which was assessed with ESS. Runs were combined using LogCombiner (Rambaut & Drummond 2015) and the results were visualised using Tracer v.1.6 (Rambaut et al. 2015).

Microsatellite analysis

For each microsatellite locus and population we tested for deviations from Hardy–Weinberg equilibrium (HWE) and linkage disequilibrium (LD) using GENEPOP v. 4.0.10 (Raymond & Rousset 1995). We applied a Bonferroni sequential correction to control for type I errors (Rice 1989). Null allele estimates were calculated in MICROCHECKER (van Oosterhout et al. 2004). Allelic richness and number of private alleles in each population were calculated after controlling for differences in sample size, using a rarefaction approach implemented in HP-RARE (Kalinowski 2005). Inbreeding coefficients (F_{IS}) were calculated with GENEPOP. We tested for population structure by calculating global and pairwise F_{ST} values across the three Cape Verde warbler populations. Global and pairwise F_{ST} estimates, observed (Ho) and expected (He) heterozygosity were calculated using Arlequin v. 3.5 (Excoffier & Lischer 2010).

To determine the most likely number of genetic clusters (K), a Bayesian algorithm was implemented in STRUCTURE v. 2.3 (Pritchard et al. 2000). The goal was to detect overall population structure in the Cape Verde warbler, so we included information on population of origin in the analysis. We used a no-admixture ancestry model and independent allele frequencies. However, as a no-prior model may be better for revealing subtle similarities, or possible gene flow between populations (Pritchard et al. 2000; Falush et al. 2007; Hubisz et al. 2009), we also ran a model using admixture and correlated allele frequencies, disregarding prior information on sampling location. For both models, we undertook four runs of 500,000 MCMC repetitions with a burn-in of 20,000 at each clustering level for K = 1–6. To assess structure within the Santiago population, we ran both analyses using only the Santiago birds. The number of clusters best fitting the data was determined using both log probabilities and the ad hoc ΔK test (Evanno et al. 2005), implemented in STRUCTURE HARVESTER (Earl & vonHoldt 2012). Graphical results were visualised using DISTRUCT version 1.1 (Rosenberg 2004).

Bottleneck tests

We used three methods to assess evidence of population bottlenecks: the heterozygosity excess test (Cornuet & Luikart 1996) implemented in BOTTLENECK v. 1.2.02 (Piry et al. 1999); the mode shift test (Luikart et al. 1998), implemented in the same software; and the M-ratio test (Garza & Williamson 2001), as well as a modified version of this last test which accounts for monomorphic alleles (Excoffier et al. 2005), in Arlequin. Heterozygosity and mode-shift tests use excess heterozygosity and allele frequency data to detect recent bottlenecks, i.e. bottlenecks that have occurred within the past dozen generations (Luikart et al. 1998). M-ratio tests (Garza & Williamson 2001; Excoffier et al. 2005) are based on the reasoning that a bottlenecked population will lose some rare alleles, and this loss will not depend on allele size; therefore a bottlenecked population will show gaps in the allele size distribution. M-ratio tests can detect bottleneck signatures over longer periods than heterozygosity or shift mode tests, as M-ratios are likely to change more slowly than heterozygosity after a bottleneck (Garza & Williamson 2001; Williamson-Natesan 2005; Peery et al. 2012).

The distances between these islands are unequal (see Fig 1), and dispersal patterns can be differently influenced by the direction of the prevailing winds (Hazevoet 1995; Correia 1996; Hille et al. 2003). The islands have different geography, age, and habitat types available for the birds (Hazevoet 1995) and there are only three populations. It is therefore not possible to ascertain the effects of isolation by distance (Wright 1943), resistance (a metric which combines Euclidean distances and spatial heterogeneity) (McRae & Nürnberger 2006) or environment factors (Wang & Bradburd 2014) on the genetic structure of the Cape Verde warbler using these data.

Results

Cytochrome b

We observed 16 cytochrome b haplotypes in the Cape Verde warbler, five haplotypes in the greater swamp warbler and one haplotype in the Seychelles warbler (Table 1; Fig 2 and 3). The Cape Verde warbler populations do not share haplotypes except for the most common one, shared between Santiago and Fogo (Fig 2 and 3). The sequences have been deposited in the National Center for Biotechnology Information (NCBI) GenBank database (accession nos. KX518324–KX518347). Where the same haplotype was found in two different populations, we submitted two identical sequences. Basic diversity indexes are given in Table 1. Pairwise genetic distances between haplotype sequences within and among Cape Verde warbler populations were very low in all cases, ranging between 0.1% and 0.5% (Table 2). The Santiago population (the largest) showed the highest diversity within the Cape Verde warbler, in terms of number of haplotypes, number of segregating sites, haplotype and nucleotide diversity (Table 1, Fig 4). Nevertheless, haplotype and nucleotide diversity in the greater swamp warbler were much higher, despite the limited number of samples screened (Table 1). The population of S. Nicolau exhibits the lowest levels of diversity, while Fogo is intermediate (Table 1, Fig 4). Diversity levels for the Cape Verde warbler population overall are similar to those for the Santiago population (Table 1). The neutrality tests (Tajima's D and Fu’s FS), which reflect population expansion or contraction in the absence of selective forces, were not significant but showed tendencies for population expansion on Santiago, and contraction on Fogo and S. Nicolau (Table 1).

The maximum likelihood phylogenetic tree and the haplotype network show that the greater swamp warbler haplotypes cluster in one distinct group (Fig 2 and 3). Among the Cape Verde warbler haplotypes, some nodes have low bootstrap support values because of the low information content of the sequences, as seen in their very low genetic distances (Table 2). Importantly, the S.

Nicolau haplotypes seem to have derived from the most common haplotype (labelled CVW_ST_FG in Fig 2), which is found on Santiago and Fogo but not on S. Nicolau (Fig 2 and 3). Apart from this most common haplotype, there are no shared haplotypes between the three populations.

Time to the most recent common ancestor

The tMRCA estimate was sampled from an approximately log-normal distribution; hence geometric means are provided (Morrison 2008). The tMRCA for the Cape Verde and greater swamp warbler was estimated at about 292 Kya, with upper confidence limits well within the last million years (Table 3). The tMRCA of S. Nicolau and the two Sotavento populations was estimated at 199 Kya, and the tMRCA for Santiago and Fogo at 165 Kya, both with upper 95% confidence limits within the last 650 thousand years (Table 3). The lower 95% confidence limits for the tMRCA for all Cape Verde warbler populations, and for the Sotavento populations, was over 30 Kya ago, i.e., prior to the Last Glacial Maximum (Clark et al. 2009).

Microsatellites

Genetic diversity measurements are given in Table 4. The estimated error rate per locus was 0.25%. Evidence for null alleles was found in microsatellite loci Ase11 and Ase22 with null allele frequencies of 0.309 and 0.164, respectively, using the Oosterhout algorithm (van Oosterhout et al. 2004), so these were removed from subsequent analyses. Ase13 and Ase48 were in strong linkage disequilibrium (*P* < 0.001) across all populations. There was no difference between results of summary statistics using either only Ase13 or only Ase48 (Mantel test on F_{ST} matrices, *P* = 0.172; Wilcoxon tests on He and Ho, all *P* > 0.742), so Ase48 was removed from further analyses. After removing these markers, we found no deviations from HWE or significant LD between loci. Heterozygosity, allelic richness and number of private alleles show a clear gradient, with highest values in Santiago, intermediate in Fogo and lowest in S. Nicolau, concurring with the mtDNA results (Table 4 and Fig 4). The inbreeding coefficient (F_{is}) is very close to random expectation for all the populations of the Cape Verde warbler, i.e. between -0.0030 and 0.0047 (Table 4).

Global F_{ST} was 0.33 for cytochrome b and 0.16 for microsatellites (both $P < 0.001$), showing pronounced overall differentiation among the populations. Pairwise F_{ST} values are shown in Table 5. Differentiation between the Santiago and Fogo populations was moderate to high, and high between those populations and that of S. Nicolau. The pattern of differentiation identified among populations was the same for mitochondrial and microsatellite markers.

The Evanno method supports two main genetic clusters (k = 2; highest $\Delta K = 342.51$, Fig S1 and S2), one including the S. Nicolau population and one with the Sotavento populations (Santiago and Fogo, Fig S2). However, when k = 3 the plot shows clear separation between the three island populations (Fig 5). This indicates that even though the main differences are between S. Nicolau and the Sotavento populations, there are also marked differences between Santiago and Fogo. For the analyses using only the Santiago population, the best k = 1 (results not shown), showing no evidence of structure within the Santiago population.

Bottleneck tests

All tests show a clear signal of a past population bottleneck in the S. Nicolau population (Table 6). The M-ratio tests and modified M-ratio tests range between 0.15 and 0.30 (Table 6), well below the critical value of 0.68 (Garza & Williamson 2001), suggesting bottlenecks have occurred on all islands. However the shift mode and heterozygosity tests (using the strict stepwise mutation model) did not detect bottlenecks for the Santiago and Fogo populations (Table 6).

Discussion

Overall we found low genetic diversity within, but considerable divergence among, the three remaining populations of Cape Verde warbler. There is a gradient of genetic diversity decreasing from Santiago through Fogo to S. Nicolau (Fig 4). The gradient of genetic diversity from the larger to the smaller island is consistent with expectations derived from population and island size theories (MacArthur & Wilson 1967; Frankham 1996; Frankham et al. 2009). Santiago possibly holds the largest population, estimated at ca. 500 pairs in the 1980s, when the warbler was thought to be absent from the north part of the island (Hazevoet 1995). As birds can now be found all across this island, if the density and territory size are relatively constant, the actual population size is likely to be considerably higher (pers. obs.). However, to our knowledge there is no information on density or territory sizes for this island, and no censuses that could provide a reliable basis for an accurate estimate have been carried out on Santiago; thus estimates must be taken with extreme caution (BirdLife International 2016). Clearly, it would be advisable to undertake thorough population censuses on Santiago to better inform conservation. The population on Fogo is medium-sized, with ca. 500 pairs (Hering & Fuchs 2009) and birds limited to the well-vegetated north-eastern part of the island. On S. Nicolau there is a very small population (8-10 pairs; Hazevoet et al. 1999; Donald et al. 2004, 12-13 pairs; this study) limited to the reed patches and mango tree valleys within a small (< 20 km^2) central part of the island (pers. obs.).

Genetic divergence is greatest between S. Nicolau and the two Sotavento populations of Santiago and Fogo (Table 5). Our results are similar to those of previous studies addressing genetic divergence in other Cape Verdean taxa including geckos (Arnold et al. 2008; Vasconcelos et al. 2012), skinks (Brehm et al. 2001; Brown et al. 2001) and kestrels (Hille et al. 2003). There is usually a clear genetic differentiation between the populations on the Barlavento and Sotavento island groups (e.g., Brown et al. 2001; Hille et al. 2003), except where the focal species occupy micro-habitat niches and/or have very limited dispersal capacities. In such cases, they can be genetically divergent even among closely located islands (Brehm et al. 2001; Arnold et al. 2008; Vasconcelos et al. 2012); nevertheless the differentiation between Barlavento and Sotavento island groups is always present.

The warbler population on Santiago is widespread, inhabits a diverse range of habitats, from sugarcane plantations to mountain forests, and is genetically the most diverse of the three (Tables 1 and 5, Fig 3 and 4), but we did not find any population substructure. Considering that, based on area of occurrence of the warbler alone, it is potentially the largest population, and the fact that it is the most genetically diverse, we would argue that the Santiago population is currently of least conservation concern of the three populations. Although the Fogo population was only discovered in 2004 (Hering & Hering 2005), our results indicate that the island was colonised well before the Last Glacial Maximum (> 30 Kya) and has been isolated and diverging ever since. This population seems to be of medium conservation concern. Taken together, our results indicate that the split between the Santiago and Fogo populations was more recent than between those two and S. Nicolau. Current demographic trends in any of the three populations are unknown.

When the Cape Verde warbler was discovered on S. Nicolau in the 1860s, it was said to be “numerous” (Keulemans 1866). However, by the 1920s it was considered extinct on the island, and was only rediscovered in the 1990s (Hazevoet 1995; Hazevoet et al. 1999). One specimen, collected in the 1970s and housed in the Zoological Centre in Lisbon, is testimony that a remnant population persisted between the 1920s and the 1990s (Hazevoet et al. 1999). Our results support the occurrence of a severe bottleneck in this population, in agreement with the anecdotal evidence (Hazevoet et al. 1999; Donald et al. 2004). This population has lower diversity, different microsatellite allele frequencies and private mitochondrial alleles when compared to Santiago and Fogo. Phylogenetically the S. Nicolau birds group with the other Cape Verde warbler populations rather than with the greater swamp warbler (Fig 2). Additionally, the tMRCA of S. Nicolau and the two Sotavento populations is 199 kya. Consequently, the current S. Nicolau population cannot have been derived from birds that arrived from Santiago or Fogo in the late 20th century. It is also highly unlikely that S. Nicolau was recolonised by greater swamp warblers dispersing directly from Africa after the original population was extinct. There is a strong indication that a bottleneck has occurred in this population, and based on our own sampling effort we estimate a maximum of 20–25 breeding pairs currently on the island. S. Nicolau is known to have suffered from severe droughts during the last century (Sena-Martins & Moreno 1986; Hazevoet 1995; Correia 1996) and this and other factors, such as predation and/or pathogens, may have reduced the warbler population. However, we cannot accurately estimate by how much it has decreased given that there is no reported estimate of its original size, and assessing possible habitat loss is beyond the scope of this paper. Thus, any links between bottlenecks and constraints on population size are only speculative at this point. Clearly, the warbler population on S. Nicolau should be considered to be of high conservation concern.

The estimated tMRCA between the three Cape Verde warbler populations (165–199 Kya) is not that much more recent than the tMRCA between the Cape Verde and greater swamp warblers (292 Kya; Table 3), suggesting that there was little gene flow between the Cape Verde warbler populations after the colonisation of the islands. However, divergence dates have an associated error (95% HPD; Table 3) and the use of different markers, substitution rates, calibration methods or molecular clocks can lead to variation in estimated tMRCA and divergence times (e.g. Ho 2007; Yoshikawa et al. 2008; Cibois et al. 2011). Consequently tMRCA must be treated with caution. For example, the molecular substitution rate we used was 2.1% sequence divergence per million years (Weir & Schluter 2008). However, there is evidence that no universal molecular clock rate exists for birds, and that substitution rates can be higher than 2.1% for recently diverged sister species or for intraspecific datasets (García-Moreno 2004; Lovette 2004; Subramanian et al. 2009). Nevertheless, it is clear that the Cape Verde warbler diverged from its sister species within the last million years (geometric mean estimate = 292 Kya), as previously suggested (Hazevoet 1995; Fregin et al. 2009), and that the three Cape Verde warbler populations have been isolated from each other since before the Last Glacial Maximum (> 30 Kya) and perhaps much longer (199 Kya). That Santiago and Fogo share the same most common cytochrome b haplotype (Fig 2 and 3), and have a tMRCA estimate of 165 Kya suggest that the split between these two populations was more recent than the split between them and S. Nicolau (199 Kya), with which they share no cytochrome b haplotypes.

Our results indicate that the populations have been through bottlenecks of variable intensity, with particularly strong evidence for the S. Nicolau population (Table 6). The M-ratio, which tests for occurrence of gaps in allele size range, can still show a bottleneck signal long after the signal is no longer detectable with heterozygosity or shift mode tests (Garza & Williamson 2001; Williamson-Natesan 2005; Peery et al. 2012). It is tempting to infer that, even though all the warbler populations appear to have undergone bottlenecks, the one affecting S. Nicolau occurred more recently, or is still ongoing, while in Fogo and Santiago the populations have stopped decreasing, or are even increasing in size. Contrasting results from M-ratio and heterozygosity tests are commonly used to infer differences in the timing of bottlenecks (Garza & Williamson 2001; Williamson-Natesan 2005; Peery et al. 2012), but these tests can be unequally affected if assumptions are violated (Peery et al. 2012), so results must be interpreted with caution. Therefore, even though our data show that the S. Nicolau population has been recently, or still is, bottlenecked, and that evidence for bottlenecks in the other two populations is not as strong, this does not exclude the possibility that the other populations are currently decreasing as well.

Despite extensive surveys, we did not find warblers on either Brava or Santo Antão, consistent with results from previous work (H. Dinis & E. Ramos, *pers. comm.*). This means that the Brava population, last reported in 1969 (Hazevoet 1995), is very likely to be extinct. A possible explanation is that, as the second smallest island in the archipelago, it would not have had sufficient suitable habitat to support a population through the drought periods that the archipelago suffered in the 1970s and 1980s (Hazevoet 1995; Correia 1996). In the case of Santo Antão, a lack of historical information means we are unable to make any inferences about the possible presence of a population there in the past.

The three populations of the Cape Verde warbler fall under the definition of MUs, but not under the definition of ESUs. There is no evidence of gene flow among them, they are genetically distinct, but there is sharing of one mitochondrial haplotype at least between Fogo and Santiago. The population of S. Nicolau is especially distinct, as it does not share mitochondrial haplotypes with either of the Sotavento populations, even though they are not reciprocally monophyletic. From a conservation genetics perspective, it may appear prudent to manage all three populations as separate units, such as suggested for similar situations in other taxa (Fouquet et al. 2010; Garcia-del-Rey et al. 2013; Reynolds et al. 2015). However, the decision to preserve each population's genetic uniqueness in the face of future risks of inbreeding depression or a limited adaptive potential is not straightforward (Amos & Balmford 2001; Miller et al. 2009; Weeks et al. 2011; Coleman et al. 2013; Weeks et al. 2016). Augmentation of the number of individuals on some islands, to help prevent inbreeding and maximise adaptive potential, may be beneficial to the populations' chances of survival (Weeks et al. 2011; Weeks et al. 2016). However, such actions could potentially lead to outbreeding depression (Edmands 2007) and the relative potential costs and benefits must therefore be assessed carefully. In this work, we did not investigate potential direct threats to any of the three populations, such as pathogens, predation or habitat destruction; we discuss our findings solely from a conservation genetics perspective.

Conservation conclusions and recommendations

Our genetic data suggest that each of the three remaining populations of the Cape Verde warbler should be treated as different MUs (Ryder 1986; Moritz 1994b), with different conservation priorities (high on S. Nicolau, medium on Fogo, low on Santiago). Special attention should be given to the more genetically divergent (and most genetically depauperate) population on S. Nicolau. Territory occupancy is currently being monitored at several locations on Santiago and Fogo (A. Rendall, J. Mascarenhas, A. Rodrigues, *pers. comm.*). Fewer conservation actions have been directed to S. Nicolau, where monitoring has been irregular over the last few years (L. Oliveira, *pers. comm.*). From previous surveys (Hazevoet et al. 1999; Donald et al. 2004) and our own sampling effort, we estimate that S. Nicolau currently holds a maximum of 20–25 breeding pairs (minimum 12; pers. obs.) and therefore could be considered to be on the verge of extinction. At the moment, the relative roles of genetic versus ecological factors in constraining population sizes on any of the islands are not clear and further information on this topic is required to help inform conservation actions. It is important to undertake baseline studies on this species (e.g. population sizes estimates, breeding behaviour, productivity, diet, habitat and predation), preferably on all islands as ecological differences might exist between populations. Until such studies are undertaken, managers should incorporate in their plans actions that take this bird's habitat into account. For example, in the Natural Park of Serra Malagueta, not all the invasive shrubs were removed in a recent restoration action because some warblers were nesting there (J. Mascarenhas, *pers. comm.*). Direct conservation actions and the use of augmentation in the population of S. Nicolau should be considered to help conserve this population. We hope that this genetic analysis will be of use to practitioners in refining action plans to maximise the conservation outcome for this endemic species.

Acknowledgements

We thank Andrew Power, Jaelsa Moreira, Josh Jenkins Shaw, Torbjörn Blixt, Naya Sena, Jeroen Arnoys, INIDA and Natural Parks of Serra Malagueta, Fogo and Monte Gordo for help with fieldwork; Juan Carlos Illera, Martim Melo, Martin Haase, the University of Oslo Natural History Museum and the Seychelles Warbler Research Group for extra samples; Urban Olsson, Juan Carlos Illera and Lewis Spurgin for input on molecular work; and Eduardo Garcia-del-Rey and an anonymous reviewer for valuable comments on the manuscript. Samples were collected under permits 36/2013 and 35/2014 issued by the General Direction for the Environment of the Republic of Cape Verde. This work was supported by a Fundação para a Ciência e Tecnologia (FCT Doctoral Research Grant SFRH/BD/84102/2012), through POPH funds from the European Social Fund (study design and data analysis), Rufford Foundation, African Bird Club, BirdLife International, Royal Society for the Protection of Birds (fieldwork), University of East Anglia and A. G. Leventis Foundation (molecular work).

Author contributions

DSR designed the study with input from IB and NJC. HRB, DSR, IB and NJC worked to raise the funding. HRB collected the samples, except where stated otherwise, did the molecular work and drafted the manuscript. HRB and DJW analysed the data with input from DSR. All authors had input into earlier drafts and agreed on the final manuscript.

Data accession statement

Haplotype sequences have been submitted to GenBank (accession numbers KX518324 -KX518347).

Conflict of interest

The authors declare that they have no conflict of interest.

Compliance with ethical standards

This work has been approved by the BIO Ethical Review Committee of the University of East Anglia.

References

Allendorf FW, Luikart GH (2007) Conservation and the genetics of populations. Blackwell Publishing, Malden.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research, 25, 3389-3402.

Amos W, Balmford A (2001) When does conservation genetics matter? Heredity, 87, 257-265.

Arnold EN, Vasconcelos R, Harris DJ, Mateo JA, Carranza S (2008) Systematics, biogeography and evolution of the endemic *Hemidactylus* geckos (Reptilia, Squamata, Gekkonidae) of the Cape Verde Islands: based on morphology and mitochondrial and nuclear DNA sequences. Zoologica Scripta, 37, 619-636.

Avise JC (2004) Molecular Markers: Natural History and Evolution, second edn. Chapman & Hall, New York.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC (1987) Intraspecific Phylogeography: The Mitochondrial DNA Bridge Between Population Genetics and Systematics. Annual Review of Ecology and Systematics, 18, 489-522.

Baillie JEM, Bennun LA, Brooks TM, Butchart SHM, Chanson JS, Cokeliss Z, Hilton-Taylor C, Hoffmann M, Mace GM, Mainka SA, Pollock CM, Rodrigues ASL, Stattersfield AJ, Stuart SN (2004) 2004 IUCN Red List of threatened species: a global species assessment. (eds. Hilton-Taylor C, Baillie J, Stuart SN). IUCN, Gland.

Baker K (1997) Warblers of Europe, Asia and North Africa. Princeton University Press, Princeton.

Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution, 16, 37-48.

Bensasson D, Zhang D-X, Hartl DL, Hewitt GM (2001) Mitochondrial pseudogenes: evolution's misplaced witnesses. Trends in Ecology & Evolution, 16, 314-321.

Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW (2013) GenBank. Nucleic Acids Research, 41, D36-D42.

BirdLife International (2016) Species factsheet: *Acrocephalus brevipennis*. http://www.birdlife.org. Accessed 11 September 2013.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A, Drummond AJ (2014) BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. PLoS Comput Biol, 10, e1003537.

Brehm A, Jesus J, Pinheiro M, Harris DJ (2001) Relationships of Scincid Lizards (*Mabuya* spp; Reptilia: Scincidae) from the Cape Verde Islands Based on Mitochondrial and Nuclear DNA Sequences. Molecular Phylogenetics and Evolution, 19, 311-316.

Brook BW, Tonkyn DW, Q'Grady JJ, Frankham R (2002) Contribution of inbreeding to extinction risk in threatened species. Conservation Ecology, 6, 16.

Brooks TM, Mittermeier RA, Mittermeier CG, Da Fonseca GAB, Rylands AB, Konstant WR, Flick P, Pilgrim J, Oldfield S, Magin G, Hilton-Taylor C (2002) Habitat Loss and Extinction in the Hotspots of Biodiversity. Conservation Biology, 16, 909-923.

Brown RP, Suarez NM, Smith A, Pestano J (2001) Phylogeography of Cape Verde Island skinks (*Mabuya*). Molecular Ecology, 10, 1593-1597.

Butchart SHM, Walpole M, Collen B, van Strien A, Scharlemann JPW, Almond REA, Baillie JEM, Bomhard B, Brown C, Bruno J, Carpenter KE, Carr GM, Chanson J, Chenery AM, Csirke J, Davidson NC, Dentener F, Foster M, Galli A, Galloway JN, Genovesi P, Gregory RD, Hockings M, Kapos V, Lamarque J-F, Leverington F, Loh J, McGeoch MA, McRae L, Minasyan A, Morcillo MH, Oldfield TEE, Pauly D, Quader S, Revenga C, Sauer JR, Skolnik B, Spear D, Stanwell-Smith D, Stuart SN, Symes A, Tierney M, Tyrrell TD, Vié J-C, Watson R (2010) Global Biodiversity: Indicators of Recent Declines. Science, 328, 1164-1168.

Cibois A, Beadell JS, Graves GR, Pasquet E, Slikas B, Sonsthagen SA, Thibault J-C, Fleischer RC (2011) Charting the course of reed-warblers across the Pacific islands. Journal of Biogeography, 38, 1963-1975.

Cibois A, Thibault J-C, Pasquet E (2008) Systematics of the extinct reed warblers *Acrocephalus* of the Society Islands of eastern Polynesia. Ibis, 150, 365-376.

Clark PU, Dyke AS, Shakun JD, Carlson AE, Clark J, Wohlfarth B, Mitrovica JX, Hostetler SW, McCabe AM (2009) The Last Glacial Maximum. Science, 325, 710-714.

Coleman RA, Weeks AR, Hoffmann AA (2013) Balancing genetic uniqueness and genetic variation in determining conservation and translocation strategies: a comprehensive case study of threatened dwarf galaxias, *Galaxiella pusilla* (Mack) (Pisces: Galaxiidae). Molecular Ecology, 22, 1820-1835.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics, 144, 2001-2014.

Correia E (1996) Contribuições para o conhecimento do clima de Cabo Verde. Garcia de Orta Série de Geografia, 81-107.

Cramp S, Perrins CM (1992) Handbook of the birds of Europe, the Middle East and Africa. Oxford University Press, Oxford.

Crnokrak P, Roff DA (1999) Inbreeding depression in the wild. Heredity, 83, 260-270.

Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. Nat Meth, 9, 772-772.

Donald PF, Taylor R, de Ponte Machado M, Pitta Groz MJ, Wells CE, Marlow T, Hille SM (2004) Status of the Cape Verde cane warbler *Acrocephalus brevipennis* on São Nicolau, with notes on song, breeding behaviour and threats. Malimbus, 26, 34-37.

Dool SE, Puechmaille SJ, Dietz C, Juste J, Ibáñez C, Hulva P, Roué SG, Petit EJ, Jones G, Russo D, Toffoli R, Viglino A, Martinoli A, Rossiter SJ, Teeling EC (2013) Phylogeography and postglacial recolonization of Europe by *Rhinolophus hipposideros*: evidence from multiple genetic markers. Molecular Ecology, 22, 4055-4070.

Drummond AJ, Bouckaert RR (2014) Bayesian evolutionary analysis with BEAST 2. Cambridge University Press, Cambridge.

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.

Dutta T, Sharma S, Maldonado JE, Wood TC, Panwar HS, Seidensticker J (2013) Gene flow and demographic history of leopards (*Panthera pardus*) in the central Indian highlands. Evolutionary Applications, 6, 949-959.

Earl D, vonHoldt B (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genet Resour, 4, 359-361.

Edmands S (2007) Between a rock and a hard place: evaluating the relative risks of inbreeding and outbreeding for conservation and management. Mol Ecol, 16.

Emerson BC (2002) Evolution on oceanic islands: molecular phylogenetic approaches to understanding pattern and process. Molecular ecology, 11, 951-966.

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

Excoffier L, Estoup A, Cornuet J-M (2005) Bayesian analysis of an admixture model with mutations and arbitrarily linked markers. Genetics, 169, 1727-1738.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources, 10, 564-567.

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

Ferchaud A-L, Eudeline R, Arnal V, Cheylan M, Pottier G, Leblois R, Crochet P-A (2015) Congruent signals of population history but radically different patterns of genetic diversity between mitochondrial and nuclear markers in a mountain lizard. Molecular Ecology, 24, 192-207.

Fouquet A, Green DM, Waldman B, Bowsher JH, McBride KP, Gemmell NJ (2010) Phylogeography of *Leiopelma hochstetteri* reveals strong genetic structure and suggests new conservation priorities. Conserv Genet, 11, 907-919.

Frankham R (1996) Relationship of Genetic Variation to Population Size in Wildlife. Conservation Biology, 10, 1500-1508.

Frankham R (1998) Inbreeding and Extinction: Island Populations. Conservation Biology, 12, 665-675.

Frankham R (2005) Genetics and extinction. Biological Conservation, 126, 131-140.

Frankham R, Ballou JD, Briscoe DA (2009) Introduction to Conservation Genetics, Second edn. Cambridge University Press, Cambridge.

Frankham R, Ralls K (1998) Conservation biology: Inbreeding leads to extinction. Nature, 392, 441-442.

Fraser DJ, Bernatchez L (2001) Adaptive evolutionary conservation: towards a unified concept for defining conservation units. Molecular Ecology, 10, 2741-2752.

Fregin S, Haase M, Olsson U, Alström P (2009) Multi-locus phylogeny of the family Acrocephalidae (Aves: Passeriformes) – The traditional taxonomy overthrown. Molecular Phylogenetics and Evolution, 52, 866-878.

Fu Y-X (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics, 147, 915-925.

Funk WC, McKay JK, Hohenlohe PA, Allendorf FW (2012) Harnessing genomics for delineating conservation units. Trends in Ecology & Evolution, 27, 489-496.

Garcia-del-Rey E, Marthinsen G, Calabuig P, Estévez L, Johannessen LE, Johnsen A, Laskemoen T, Lifjeld JT (2013) Reduced genetic diversity and sperm motility in the endangered Gran Canaria Blue Chaffinch *Fringilla teydea polatzeki*. J Ornithol, 154, 761-768.

García-Moreno J (2004) Is there a universal mtDNA clock for birds? Journal of Avian Biology, 35, 465-468.

Garza JC, Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. Molecular Ecology, 10, 305-318.

Gilpin M, Soulé M (1986) Minimum Viable Populations: processes of species extinction. In:Soulé M (ed) Conservation Biology: The Science of Scarcity and Diversity. Sinauer, Sunderland.

Griffin AS, Blumstein DT, Evans CS (2000) Training captive-bred or translocated animals to avoid predators / entrenamiento de animales para evitar depredadores. Conservation Biology, 14, 1317-1326.

Griffiths R, Double MC, Orr K, Dawson RJG (1998) A DNA test to sex most birds. Molecular Ecology, 7, 1071-1075.

Guindon S, Gascuel O (2003) A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. Systematic Biology, 52, 696-704.

Hale ML, Burg TM, Steeves TE (2012) Sampling for microsatellite-based population genetic studies: 25 to 30 individuals per population is enough to accurately estimate allele frequencies. PLoS ONE, 7, e45170.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. Nucl Acids Symp, 41, 95-98.

Hasegawa M, Kishino H, Yano T-a (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol, 22, 160-174.

Hazevoet C (1993) On the history and type specimens of the Cape Verde Cane Warbler *Acrocephalus brevipennis* (Keulemans, 1866) (Aves, Sylviidae). Bijdragen tot de Dierkunde, 62, 5.

Hazevoet CJ (1995) The birds of the Cape Verde islands: an annotated checklist. British Ornithologists' Union, Tring.

Hazevoet CJ, Monteiro LR, Ratcliffe N (1999) Rediscovery of the Cape Verde Cane Warbler *Acrocephalus brevipennis* on São Nicolau in February 1998. Bulletin of the British Ornithologists' Club, 4.

Hedrick PW, Kalinowski ST (2000) Inbreeding depression in conservation biology. Annual Review of Ecology and Systematics, 31, 139-162.

Helbig AJ, Kocum A, Seibold I, Braun MJ (2005) A multi-gene phylogeny of aquiline eagles (Aves: Accipitriformes) reveals extensive paraphyly at the genus level. Molecular Phylogenetics and Evolution, 35, 147-164.

Hering J, Fuchs E (2009) The Cape Verde Warbler: distribution, density, habitat and breeding biology on the island of Fogo. British Birds, 8.

Hering J, Hering (2005) Discovery of Cape Verde Warbler *Acrocephalus brevipennis* on Fogo, Cape Verde Islands. Bulletin of the African Bird Club, 12, 3.

Hille SM, Nesje M, Segelbacher G (2003) Genetic structure of kestrel populations and colonization of the Cape Verde archipelago. Molecular Ecology, 12, 2145-2151.

Ho SYM (2007) Calibrating molecular estimates of substitution rates and divergence times in birds. Journal of Avian Biology, 38, 409-414.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. Molecular ecology resources, 9, 1322-1332.

Illera JC, Emerson BC, Richardson DS (2007) Population history of Berthelot's pipit: colonization, gene flow and morphological divergence in Macaronesia. Molecular Ecology, 16, 4599-4612.

Kalinowski ST (2005) hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. Molecular Ecology Notes, 5, 187-189.

Keulemans JG (1866) Opmerkingen over de vogels van de Kaap-Verdische Eilanden en van Prins-Eiland (Ilha do Principe) in de bogt van Guinea gelegen. Nederlandsch Tijdschirft voor de Dierkunde, 3, 39.

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

Komdeur J (1991) Cooperative Breeding in the Seychelles Warbler. University of Cambridge

Komdeur J, Piersma T, Kraaijeveld K, Kraaijeveld-Smit F, Richardson DS (2004) Why Seychelles Warblers fail to recolonize nearby islands: unwilling or unable to fly there? Ibis, 146, 298-302.

Koressaar T, Remm M (2007) Enhancements and modifications of primer design program Primer3. Bioinformatics, 23, 1289-1291.

Leisler B, Heidrich P, Schulze-Hagen K, Wink M (1997) Taxonomy and phylogeny of reed warblers (genus *Acrocephalus*) based on mtDNA sequences and morphology. Journal für Ornithologie, 138, 469-496.

Leisler B, Schulze-Hagen K (2011) The reed warblers: diversity in a uniform family. KNNV Publishing, Zeist.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics, 25, 1451-1452.

Lovette I (2004) Mitochondrial dating and mixed support for the "2% rule" in birds. The Auk, 121, 1-6.

Luikart G, Allendorf F, Cornuet J-M, Sherwin W (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. Journal of Heredity, 89, 238-247.

MacArthur RH, Wilson EO (1967) The Theory of Island Biogeography. Princeton University Press, Princeton.

McCartney-Melstad E, Shaffer HB (2015) Amphibian molecular ecology and how it has informed conservation. Molecular Ecology, 24, 5084-5109.

McRae B, Nürnberger B (2006) Isolation by resistance. Evolution, 60, 1551-1561.

Miller KA, Chapple DG, Towns DR, Ritchie PA, Nelson NJ (2009) Assessing genetic diversity for conservation management: a case study of a threatened reptile. Animal Conservation, 12, 163-171.

Moritz C (1994a) Applications of mitochondrial DNA analysis in conservation: a critical review. Molecular Ecology, 3, 401-411.

Moritz C (1994b) Defining 'Evolutionarily Significant Units' for conservation. Trends in Ecology and Evolution, 9, 373-375.

Morrison DA (2008) How to summarize estimates of ancestral divergence times. Evolutionary Bioinformatics Online, 4, 75-95.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. Nature, 403, 853-858.

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.

Peery MZ, Kirby R, Reid BN, Stoelting R, Doucet-Bëer E, Robinson S, Vásquez-Carrillo C, Pauli JN, Palsbøll PJ (2012) Reliability of genetic bottleneck tests for detecting recent population declines. Molecular Ecology, 21, 3403-3418.

Piry S, Luikart G, Cornuet J-M (1999) Computer note. BOTTLENECK: a computer program for detecting recent reductions in the effective size using allele frequency data. Journal of Heredity, 90, 502-503.

Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: causes, consequences and solutions. Nat Rev Genet, 6, 847-846.

Posada D, Crandall KA (2001) Intraspecific gene genealogies: trees grafting into networks. Trends in Ecology & Evolution, 16, 37-45.

Price T (2008) Speciation in birds. Roberts and Company Publishers, Greenwood Village.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics, 155, 945-959.

R Core Team (2016) R: A language and environment for statistical computing. R Foundation for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rambaut A, Drummond AJ (2015) LogCombiner v2.3.0 Available from http://beast.bio.ed.ac.uk/LogCombiner.

Rambaut A, Suchard MA, Xie D, Drummond AJ (2015) Tracer v1.6.0 Available from http://beast.bio.ed.ac.uk/Tracer.

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

Reynolds RG, Puente-Rolón AR, Platenberg R, Tyler RK, Tolson PJ, Revell LJ (2015) Large divergence and low diversity suggest genetically informed conservation strategies for the endangered Virgin Islands Boa (*Chilabothrus monensis*). Global Ecology and Conservation, 3, 487-502.

Rice WR (1989) Analyzing tables of statistical tests. Evolution, 43, 223-225.

Richardson DS, Jury FL, Blaakmeer K, Komdeur J, Burke T (2001) Parentage assignment and extra-group paternity in a cooperative breeder: the Seychelles warbler (*Acrocephalus sechellensis*). Molecular Ecology, 10, 2263-2273.

Richardson DS, Jury FL, Dawson DA, Salgueiro P, Komdeur J, Burke T (2000) Fifty Seychelles warbler (*Acrocephalus sechellensis*) microsatellite loci polymorphic in Sylviidae species and their cross-species amplification in other passerine birds. Molecular Ecology, 9, 2225-2230.

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. Molecular Ecology Notes, 4, 137-138.

Ryder OA (1986) Species conservation and systematics: the dilemma of subspecies. Trends in Ecology & Evolution, 1, 9-10.

Saccheri I, Kuussaari M, Kankare M, Vikman P, Fortelius W, Hanski I (1998) Inbreeding and extinction in a butterfly metapopulation. Nature, 392, 491-494.

Sax DF, Gaines SD (2008) Species invasions and extinction: The future of native biodiversity on islands. Proceedings of the National Academy of Sciences, 105, 11490-11497.

Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. Ecology Letters, 9, 615-629.

Sena-Martins DAd, Moreno JMG (1986) La desertification aux ilex du Cap-Vert. (ed. Fisheries MoRDa), Praia, Cape Verde.

Sih A, Bolnick DI, Luttbeg B, Orrock JL, Peacor SD, Pintor LM, Preisser E, Rehage JS, Vonesh JR (2010) Predator–prey naïveté, antipredator behavior, and the ecology of predator invasions. Oikos, 119, 610-621.

Singh TR, Shneor O, Huchon D (2008) Bird Mitochondrial Gene Order: Insight from 3 Warbler Mitochondrial Genomes. Molecular Biology and Evolution, 25, 475-477.

Sorenson MD, Quinn TW (1998) Numts: a challenge for avian systematics and population biology. The Auk, 214-221.

Spielman D, Brook BW, Frankham R (2004) Most species are not driven to extinction before genetic factors impact them. Proceedings of the National Academy of Sciences of the United States of America, 101, 15261-15264.

Spurgin LG, Wright DJ, van der Velde M, Collar NJ, Komdeur J, Burke T, Richardson DS (2014) Museum DNA reveals the demographic history of the endangered Seychelles warbler. Evolutionary Applications, 7, 1134-1143.

Steadman DW (1995) Prehistoric extinctions of pacific island birds: biodiversity meets zooarchaeology. Science, 267, 1123-1131.

Subramanian S, Denver DR, Millar CD, Heupink T, Aschrafi A, Emslie SD, Baroni C, Lambert DM (2009) High mitogenomic evolutionary rates and time dependency. Trends in Genetics, 25, 482-486.

Sunnucks P (2000) Efficient genetic markers for population biology. Trends in ecology & evolution (Personal edition), 15, 199-203.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics, 123, 585-595.

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Molecular Biology and Evolution, 10, 512-526.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Molecular Biology and Evolution, 30, 2725-2729.

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3—new capabilities and interfaces. Nucleic Acids Research, 40, e115.

van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) micro-checker: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes, 4, 535-538.

Vasconcelos R, Perera ANA, Geniez P, Harris DJ, Carranza S (2012) An integrative taxonomic revision of the Tarentola geckos (Squamata, Phyllodactylidae) of the Cape Verde Islands. Zoological Journal of the Linnean Society, 164, 328-360.

Wang IJ, Bradburd GS (2014) Isolation by environment. Molecular Ecology, 23, 5649-5662.

Weeks AR, Sgro CM, Young AG, Frankham R, Mitchell NJ, Miller KA, Byrne M, Coates DJ, Eldridge MDB, Sunnucks P, Breed MF, James EA, Hoffmann AA (2011) Assessing the benefits and risks of translocations in changing environments: a genetic perspective. Evolutionary Applications, 4, 709-725.

Weeks AR, Stoklosa J, Hoffmann AA (2016) Conservation of genetic uniqueness of populations may increase extinction likelihood of endangered species: the case of Australian mammals. Frontiers in Zoology, 13, 1-9.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution, 38, 1358-1370.

Weir JT, Schluter D (2008) Calibrating the avian molecular clock. Molecular Ecology, 17, 2321-2328.

Williamson-Natesan EG (2005) Comparison of methods for detecting bottlenecks from microsatellite loci. Conserv Genet, 6, 551-562.

Wright S (1943) Isolation by Distance. Genetics, 28, 114-138.

Wright S (1951) The genetical structure of populations. Annals of Eugenics, 15.

Xenikoudakis G, Ersmark E, Tison JL, Waits L, Kindberg J, Swenson JE, Dalén L (2015) Consequences of a demographic bottleneck on genetic structure and variation in the Scandinavian brown bear. Molecular Ecology, 24, 3441-3454.

Yang Z (1994) Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. J Mol Evol, 39, 306-314.

Yoshikawa N, Matsui M, Nishikawa K, Kim J-B, Kryukov A (2008) Phylogenetic relationships and biogeography of the Japanese clawed salamander, *Onychodactylus japonicus* (Amphibia: Caudata: Hynobiidae), and its congener inferred from the mitochondrial cytochrome b gene. Molecular Phylogenetics and Evolution, 49, 249-259.

Figure captions

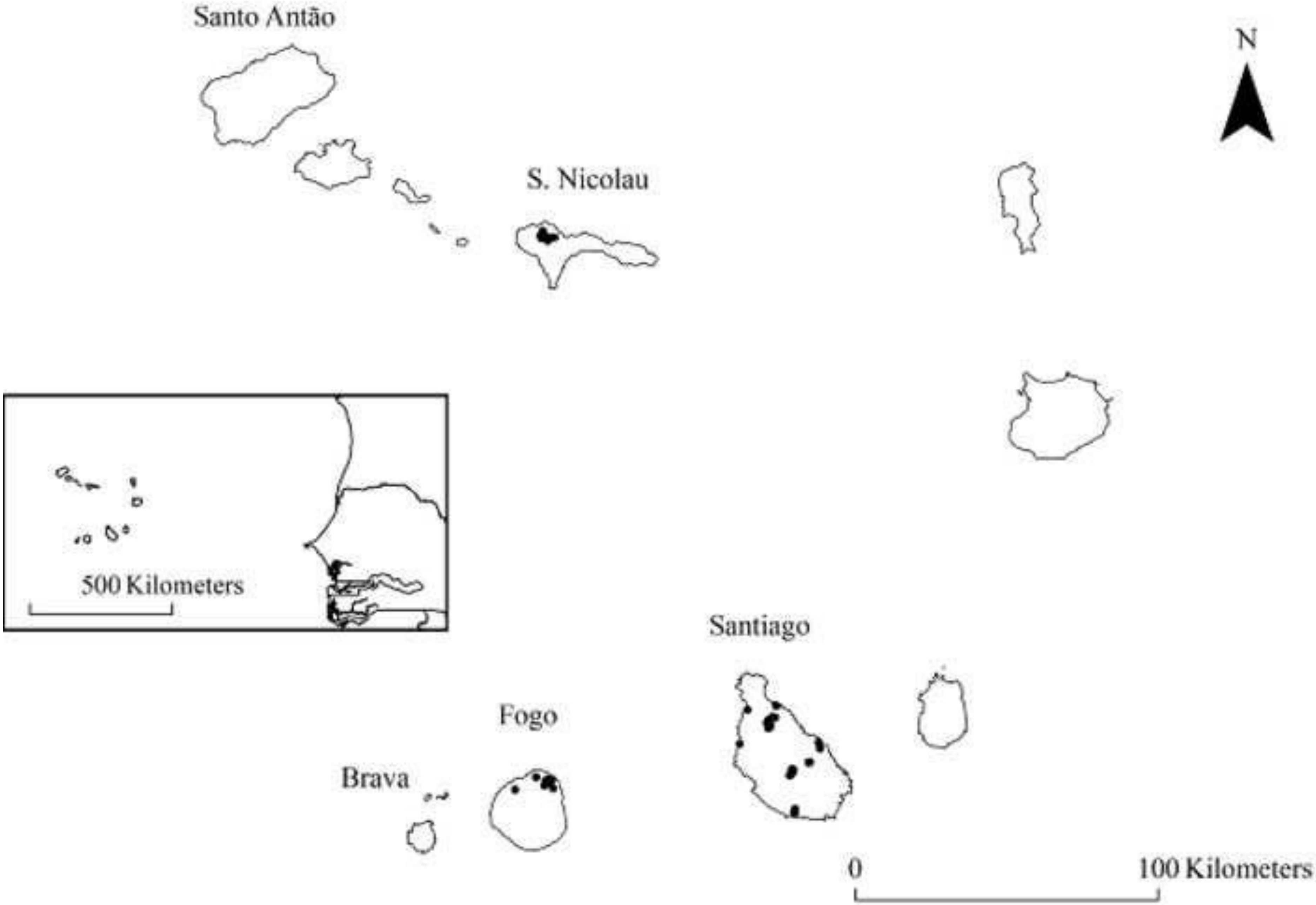
Fig. 1 The Cape Verde archipelago with the locations where Cape Verde warblers were sampled in 2013 and 2014. Inset shows the archipelago in relation to West Africa. For details on sampling sites and dates, see main text and supplementary material

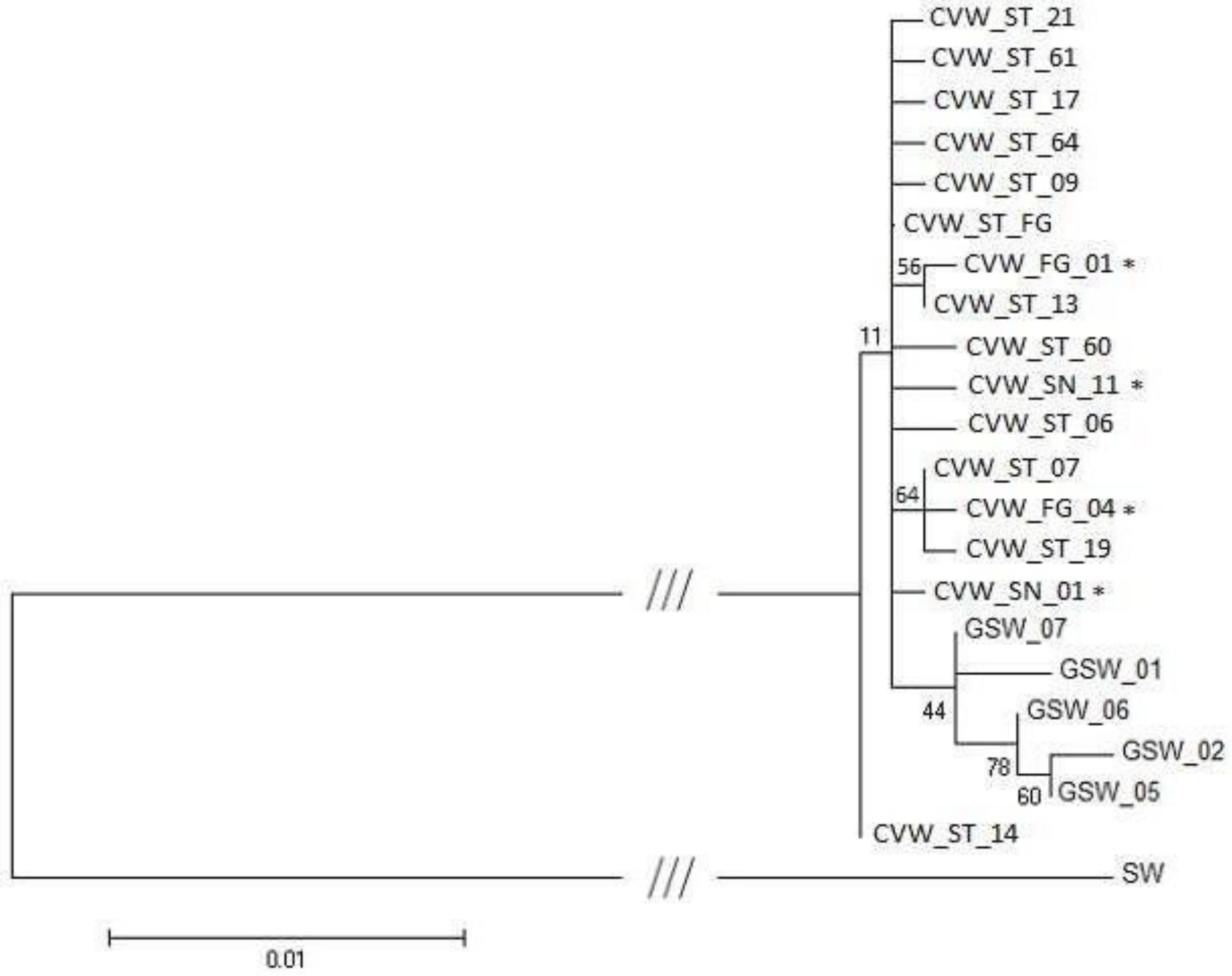
Fig. 2 Maximum likelihood phylogenetic analysis of the Cape Verde warbler (CVW, n = 145) and greater swamp warbler (GSW, n = 7) species, based on cytochrome b; haplotypes names are coded by island (ST = Santiago, FG = Fogo, SN = S. Nicolau); asterisks mark haplotypes only found on Fogo or S. Nicolau; the Seychelles warbler haplotype (SW, n = 3) is used as an outgroup

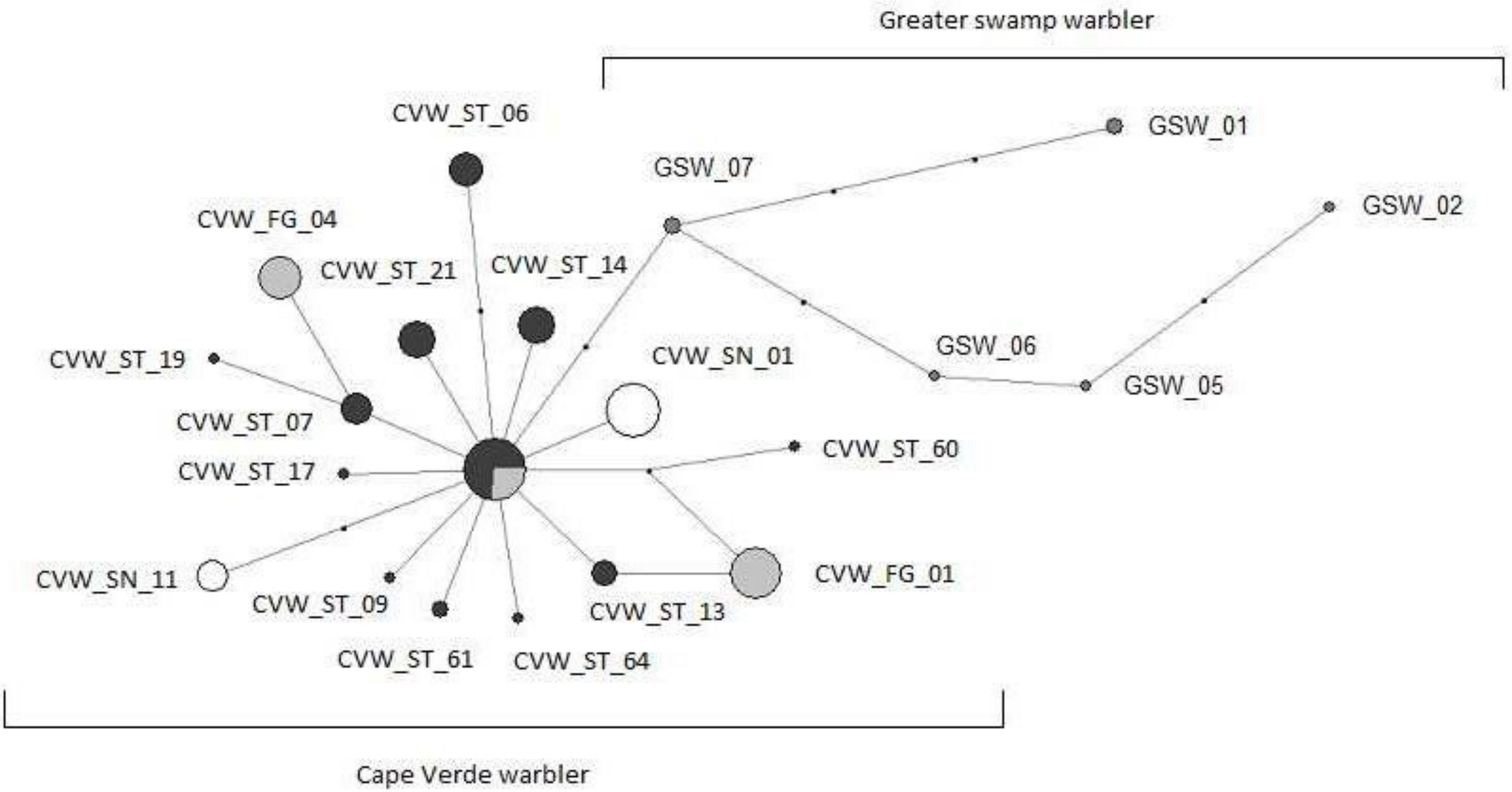
Fig. 3 Median-joining haplotype network based on cytochrome b of the Cape Verde warbler (CVW, n = 145) and the greater swamp warbler (GSW, n = 7). Haplotypes found in the Cape Verde warbler populations are represented in dark grey for Santiago, light grey for Fogo and white for S. Nicolau; the central haplotype is the most common one (labelled CVW_ST_FG in Fig. 2); haplotypes found in the greater swamp warbler, in medium grey, cluster separately. The size of the circles is proportional to the number of individuals that share that haplotype. Black dots represent intermediate (unsampled) haplotypes, with 1 bp between consecutive haplotypes

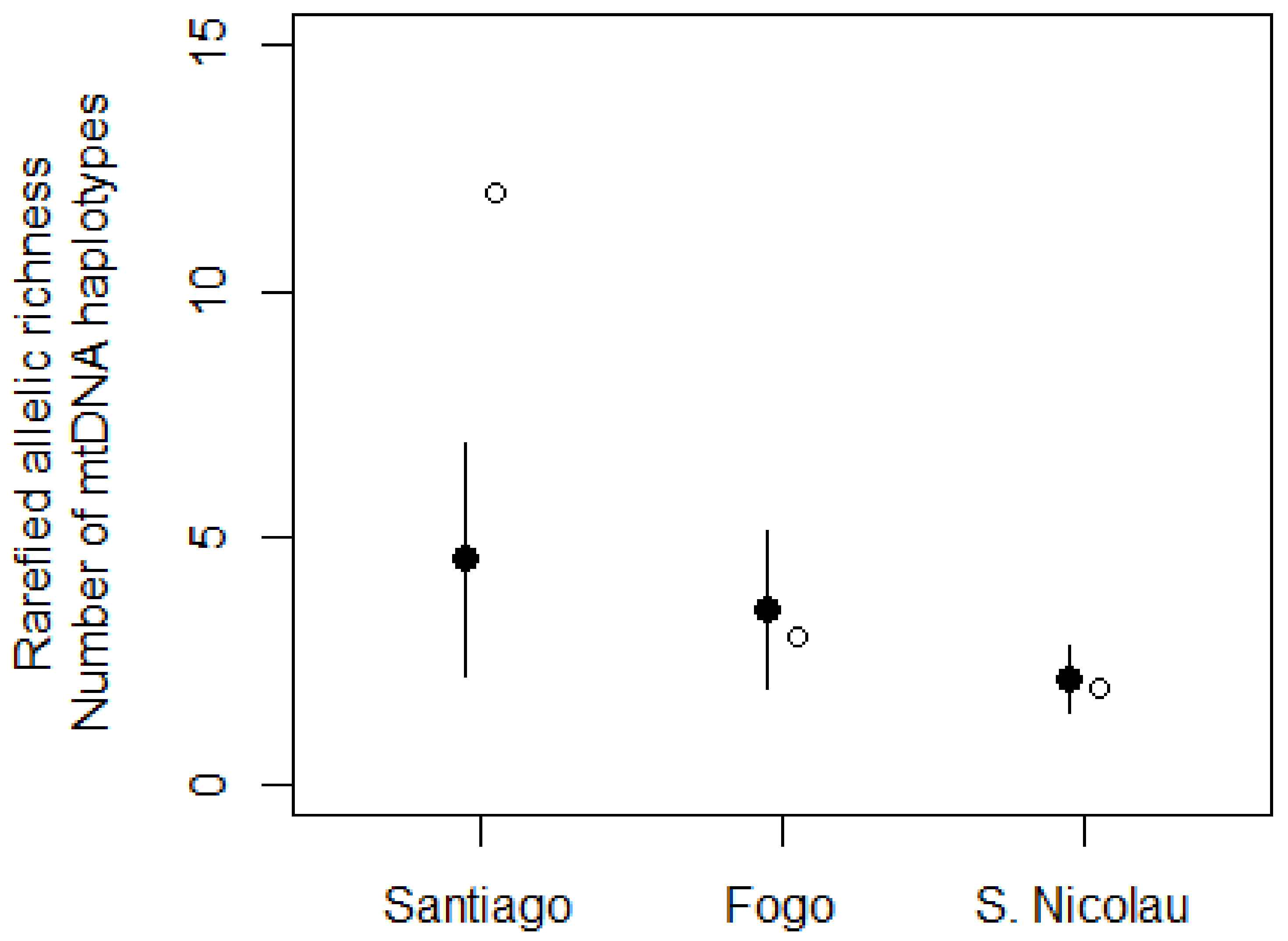
Fig. 4 Gradient of decreasing genetic diversity with estimated population size: rarefied allelic richness (full circles, n = 131; error bars = SD) and number of haplotypes (empty circles, n = 145) for each of the Cape Verde warbler (CVW) populations, ordered by decreasing population size from Santiago to S. Nicolau

Fig. 5 STRUCTURE plot showing k = 3 distinct genetic clusters in the Cape Verde warbler samples where Santiago, n = 66; Fogo, n = 38; and S. Nicolau, n = 27. Each bar represents an individual's proportional probability of membership to each cluster, given in dark grey, light grey and white, respectively


Santo Antão
S. Nicolau
N
500 Kilometers
Santiago
Fogo
Brava
0
100 Kilometers


CVW_ST_21
CVW_ST_61
CVW_ST_17
CVW_ST_64
CVW_ST_09
CVW_ST_FG
56
CVW_FG_01 *
CVW_ST_13
11
CVW_ST_60
CVW_SN_11 *
CVW_ST_06
CVW_ST_07
64
CVW_FG_04 *
CVW_ST_19
CVW_SN_01 *
GSW_07
GSW_01
44
GSW_06
GSW_02
78
60
GSW_05
CVW_ST_14
SW
0.01


Greater swamp warbler
CVW_ST_06
GSW_01
GSW_07
GSW_02
CVW_FG_04
CVW_ST_21
CVW_ST_14
GSW_06
CVW_ST_19
CVW_SN_01
GSW_05
CVW_ST_07
CVW_ST_60
CVW_ST_17
CVW_SN_11
CVW_ST_09
CVW_FG_01
CVW_ST_13
CVW_ST_61
CVW_ST_64
Cape Verde warbler


Rarefied allelic richness
Number of mtDNA haplotypes
0
5
10
15
Santiago
Fogo
S. Nicolau


S. Nicolau
Fogo
Santiago

Tables

Table 1 Genetic diversity and demography in the Cape Verde warbler (CVW) and greater swamp warbler (GSW) based on cytochrome b; N = number of individuals sequenced; S = segregating sites; h = number of haplotypes; Hd = haplotype diversity; Pi = nucleotide diversity; for Tajima's D, all *P* > 0.05 and for Fu's FS, all *P* > 0.02 (non-significant)

Population	**N**	**S**	**h**	**Hd ± SD**	**Pi ± SD**	**Tajima's D**	**Fu's FS**
All CVW	145	17	16	0.887 ± 0.011	0.00182 ± 0.00009	-0.887	-4.335
Santiago	72	13	12	0.841 ± 0.024	0.00134 ± 0.00011	-1.222	-4.356
Fogo	43	4	3	0.642 ± 0.035	0.00172 ± 0.00008	2.601	4.450
S. Nicolau	30	3	2	0.370 ± 0.084	0.00097 ± 0.00022	1.097	3.704
GSW	7	8	5	0.905 ± 0.103	0.00315 ± 0.00063	0.569	-0.333

Table 2 Estimates of evolutionary divergence among haplotype pairs (pairwise genetic distances), between the Cape Verde warbler populations and greater swamp warbler (GSW), based on cytochrome b sequences: within (diagonal in underlined italic) and among population (below diagonal)

Population	**Santiago**	**Fogo**	**S. Nicolau**	**GSW**
Santiago	<u>*0.001*</u>			
Fogo	0.002	<u>*0.002*</u>		
S. Nicolau	0.002	0.002	<u>*0.001*</u>	
GSW	0.004	0.005	0.005	<u>*0.003*</u>

Table 3 Estimates of the time to the most recent common ancestor (tMRCA; geometric mean shown) for the Seychelles warbler (SW), Cape Verde warbler (CVW) and greater swamp warbler (GSW), based on the cytochrome b; HPD = highest posterior density; Mya = million years; Kya = thousand years

Split	tMRCA	95% HPD
Seychelles warbler / Cape Verde warbler and greater swamp warbler	3.23 Mya	1.00–9.63 Mya
Cape Verde warbler / greater swamp warbler	292 Kya	67–872 Kya
S. Nicolau / Santiago and Fogo	199 Kya	40–615 Kya
Santiago / Fogo	165 Kya	32–511 Kya

Table 4 Genetic diversity for the Cape Verde warbler (CVW) populations, based on 13 microsatellite loci and using only unrelated birds (N); observed (Ho) and expected (He) heterozygosity, allelic richness, private alleles and inbreeding coefficient (F_{IS})

Population	N	Ho ± SD	He ± SD	Allelic richness ± SD	Private alleles ± SD	F_{IS}
All CVW	131	0.483 ±0.181	0.543 ±0.207	5.536 ±3327	na	0.1100
Santiago	66	0.567 ±0.196	0.565 ±0.195	4.598 ±2.446	1.290 ±1.271	-0.0030
Fogo	38	0.507 ±0.241	0.509 ±0.228	3.574 ±1.651	0.362 ±0.596	0.0047
S. Nicolau	27	0.355 ±0.120	0.355 ±0.126	2.151 ±0.685	0.183 ±0.376	0.0018

Table 5 Pairwise F_{ST} values for the Cape Verde warbler populations of Santiago, Fogo and S. Nicolau showing microsatellite pairwise F_{ST} above the solid line and Cytochrome b F_{ST} values below; all values are significant at $P < 0.001$

Population	Santiago	Fogo	S. Nicolau
Santiago		0.062	0.232
Fogo	0.256		0.275
S. Nicolau	0.349	0.450	

Table 6 Wilcoxon test for heterozygosity excess *P*-values (IAM = infinite allele model; TPM = two phase model with either 70% or 90% of stepwise mutations and SMM = stepwise

mutation model), mode shift test and Garza-Williamson’s M and modified M ratio indexes for bottleneck detection in the three Cape Verde warbler populations

Population	**IAM**	**TPM**		**SMM**	**Shift mode test**	**M index**	**Modified M index**
		(70%)	**(90%)**				
Santiago	0.0006	0.0210	0.0327	0.1465	no	0.307 ±0.117	0.289 ±0.120
Fogo	0.0002	0.0057	0.0061	0.0640	no	0.272 ±0.135	0.221 ±0.098
S. Nicolau	0.0002	0.0005	0.0005	0.0024	yes	0.201 ±0.106	0.151 ±0.101