1 Article type: Original Article

2

3 **Evidence for multiple founding lineages and genetic admixture in the evolution**
4 **of species within an oceanic island weevil (Coleoptera, Curculionidae) super-**
5 **radiation.**

6

7 Christiana M. A. Faria¹, Antonio Machado², Isabel R. Amorim³, Matthew J. G. Gage¹,
8 Paulo A.V. Borges³ and Brent C. Emerson^{1, 4}

9

10 ¹*Centre for Ecology, Evolution and Conservation, School of Biological Sciences,*
11 *University of East Anglia, Norwich NR4 7TJ, United Kingdom, ²c/Chopin 1, 38208*
12 *La Laguna, Tenerife, Canary Islands, Spain, ³Centre for Ecology, Evolution and*
13 *Environmental Changes - CE3C, Azorean Biodiversity Group and Universidade dos*
14 *Açores, Departamento de Ciências Agrárias, 9700-042 Angra do Heroísmo, Açores,*
15 *Portugal, ⁴Island Ecology and Evolution Research Group, IPNA-CSIC, 38206 La*
16 *Laguna, Tenerife, Canary Islands, Spain*

17

18 *Correspondence: Brent Emerson, Island Ecology and Evolution Research Group,
19 IPNA-CSIC, C/Astrofísico Francisco Sánchez 3, 38206 La Laguna, Tenerife, Canary
20 Islands, Spain

21 E-mail: bemerson@ipna.csic.es

22

23 Running head: Island speciation with multiple founding events and admixture

24

25 Word count: abstract, main text and references 6,980 / N pages for tables&figs.: 3-4

ABSTRACT

Aim To infer colonisation and speciation history for a closely related complex of nine species within the enigmatic Canary Island *Laparocerus* weevil radiation of 128 species. Using molecular dating and the spatial and temporal context that islands provide, we evaluate the possible explanations of incomplete lineage sorting and gene flow for the origin of shared genetic variation among species from different islands.

Location Canary Islands (Gran Canaria, Tenerife, La Palma and El Hierro).

Methods Phylogenetic analyses of mitochondrial (COII) and nuclear (ITS2) sequence data and molecular dating techniques were used to infer the origin of the group in the archipelago and their history of colonisation and differentiation.

Results Gran Canaria appeared to be the geographic origin of the complex. An unexpected result was that mtDNA revealed each of the single species on La Palma and El Hierro to be the product of more than one colonisation event from more than one source island. In both cases nuclear ITS2 data revealed these multiple colonisations to have been followed by admixture.

Main conclusions The two gene trees present very different topologies, with a rather simple colonisation history required to explain the pattern of nuclear gene relationships, while the mtDNA gene tree implicates a much more complex history of colonisation. Explanations of incomplete lineage sorting are ruled out and a history of colonisation and speciation for the *L. tessellatus* complex involving genetic admixture is inferred.

Keywords

- 50 **Canary Islands, colonisation, diversification, *Laparocerus*, lineage sorting,**
- 51 **Macaronesia, speciation**

52

53 **INTRODUCTION**

54

55 Charles Darwin's stop in the Galapagos Islands during his five-year voyage on the
56 HMS Beagle revealed to him that islands are an important source of evidence for
57 evolution. Since Darwin's time, islands have become not just a source of evidence,
58 but theatres for investigating and understanding mechanistic explanations for
59 evolution itself. This research effort has often focussed on particular islands and
60 archipelagos, reflecting the fact that the evolutionary proliferation of biodiversity has
61 progressed much farther on some islands than on others (Losos & Ricklefs, 2009).
62 Interestingly, this proliferation often appears to be disproportionately distributed
63 across co-occurring evolutionarily related taxa, with diversification having been much
64 more extensive within some lineages than others. For example, among the 88 genera
65 of weevil (Coleoptera, Curculionidae) occurring naturally within the Canary Islands,
66 there is an average of only three species per genus (Oromí *et al.*, 2010). However 128
67 species, more than 1/3 of all the native weevil species in the Canary Islands, belong to
68 a single genus, *Laparocerus* Shonherr, 1834 (A. Machado, in prep.). This raises an
69 important question, why do some lineages diversify so extensively on islands, while
70 others do not?

71

72 The addition of molecular phylogenetic techniques to the evolutionary biologist's
73 toolkit has seen renewed focus on the study of evolution within oceanic islands over
74 the last two decades, enabling researchers to address why a given lineage may have
75 speciated more extensively in some archipelagos compared to others. However, the
76 extent to which molecular phylogenetic analyses have gone beyond describing pattern

to inferring process has been limited, with most focusing on defining relationships among species, and inferring the timing of speciation events (e.g. Emerson *et al.*, 2000; Emerson & Oromí, 2005; Amorim *et al.*, 2012). Such analyses can help to elucidate the role of geography and the relative importance of within island and between island speciation. However, by typically sampling only a few individuals within a species, phylogenetic sampling does not capture diversification processes occurring below the species level.

Population-level studies that incorporate the sampling of individuals from across their range for the reconstruction of gene genealogies may help to reveal the extent to which natural selection or alternative mechanisms offer explanations of evolutionary change. Such studies have frequently revealed substantial genetic structuring within species within islands, and evolutionary processes underpinning diversification. The repeated convergent selection for, and evolutionary origins of, cave dwelling *Palmorchestia hypogaea* amphipod populations of the Canary Island of La Palma would have been all but missed with phylogenetic sampling (Villacorta *et al.*, 2008), as would have been conclusions of ancestral and derived ecological associations and niche shifts in the *Nesotes* beetles of Gran Canaria (Rees *et al.*, 2001). A revealing example of the combined power of geographic sampling with both mitochondrial and nuclear genomic sampling comes from Jordal *et al.*'s (2006) study of the sympatric and closely related *Aphanarthrum* weevil species *A. subglabrum* and *A. glabrum* on the island of La Palma in the Canary Islands. Representative geographic sampling of these two species and other taxa from the *A. glabrum* complex on other islands, combined with the analysis of mitochondrial and nuclear loci, revealed the combined

roles of geography (allopatric isolation and subsequent secondary contact) and species interactions (hybridisation and reinforcement) in driving diversification.

With 128 described species from the Canary Islands and 34 described species from the Madeira archipelago, the genus *Laparocerus* stands out as an evolutionary enigma. It is estimated that there may be as many as 200 species distributed in the Canary and Madeira archipelagos, with single species occurring in West Morocco, and the Salvages Islands (Machado, 2011). It is the most species rich of all animal and plant genera within the Canary Islands (Arechavaleta *et al.*, 2010). Species richness within the *Laparocerus* is eight times higher than that of the next most species rich weevil genus, *Acalles*, represented by 16 species (Oromí *et al.*, 2010), but there is little understanding of what evolutionary processes may underpin the speciation of this group. All *Laparocerus* species are flightless and most are oligophages that climb vegetation to feed upon leaves, while a few species dwell in the leaf litter or are adapted to the underground environment (Machado, 2003).

The extensive diversity within *Laparocerus* has for many years both intrigued and confounded biologists, but recent efforts have partitioned this diversity into taxonomic units (e.g. Machado, 2006; 2009), and species complexes within this diversity have been delineated with both mtDNA and nuclear sequence data (Machado *et al.*, 2008; A. Machado, in prep.). As a first approach to understand why *Laparocerus* has diversified so dramatically within Macaronesia, we have sampled within the well-defined *Laparocerus tessellatus* species complex of nine species across its distribution on four of the Canary Islands (Gran Canaria, Tenerife, La Palma and El Hierro). Using sequence data from mitochondrial cytochrome c oxidase

subunit II gene (COII) and the nuclear DNA internal transcribed spacer 2 gene (ITS2), we investigated inter-island colonisation within the complex, and the distribution of genetic variation among species within the complex. Using the spatial and temporal context that islands provide, we evaluate the possible explanations of incomplete lineage sorting and gene flow for the origin of shared DNA sequences among species from different islands, and discuss the implications of this shared variation for the speciation process in *Laparocerus*.

METHODS

Sampling and laboratory procedures

The *Laparocerus tessellatus* complex consists of nine species that comprise a monophyletic group based on separate mitochondrial [cytochrome c oxidase subunit II gene (COII), 16S ribosomal RNA gene (16S), and the 12S ribosomal RNA gene (12S)], and nuclear [28SrDNA] sequence data. Bayesian Inference phylogenetic analysis of 243 *Laparocerus* operational taxonomic units (OTUs), including almost all described species of *Laparocerus* in the Canary Islands, reveals the complex to be monophyletic for both mtDNA and nuclear analyses with posterior probabilities (PP) of 1 in both cases (A. Machado, unpublished data). All species are single island endemics, with five on Gran Canaria (*L. microphthalmus* Lindberg, 1950, *L. obsitus* Wollaston, 1864, *L. osorio* Machado, 2012, *L. tirajana* Machado, 2012, and *L. sp. aff. tirajana*), two on Tenerife (*L. tessellatus* Brullé, 1839 and *L. freyi* Uyttenboogaart, 1940) and one on each of La Palma (*L. sp. 1*) and El Hierro (*L. bimbache* Machado, 2011). Samples were collected from 37 sites from 1999 to 2011 (Fig 1). A total of 173

specimens were collected from 10 sites in Gran Canaria, 18 in Tenerife, six in La Palma and three in El Hierro. Site locations and number of individuals per site are detailed in Table 1. *Laparocerus vicinus* Lindberg, 1953 (Gran Canaria) was used as an outgroup, based on its close phylogenetic relationship to the *L. tessellatus* complex. It is one of several closely related species that form a monophyletic group (PP = 1) with the complex in an analysis of mtDNA COII, 16S, 12S and nuclear 28S sequence data (A. Machado, unpublished data). Upon collection, samples were stored in absolute ethanol at 4°C prior to species identification and DNA extraction.

Total genomic DNA was extracted from the head and prothorax using the DNeasy 96 well Blood and Tissue Extraction Kit (QIAGEN, West Sussex, UK) following the manufacturer's instructions, with the single exception that the head and prothorax were not ground. After extraction, both head and prothorax were placed back in absolute ethanol with the remainder of the body and maintained at 4°C as vouchers within the collection of A. Machado. Part of the mitochondrial COII gene and the nuclear ITS2 gene were amplified and sequenced as detailed in Appendix S1 (see Supporting Information). Sequencing was performed with the forward primer only for COII and in both directions for the ITS2, due to heterozygosity and indel (insertion and deletion) variation.

Sequence alignment, haplotype reconstruction, and sequence properties

All sequences were processed and ambiguous base calls manually assessed with GENEIOUS PRO 5.4.5 (Drummond *et al.*, 2010). COII sequences were aligned using MAFFT 6.814 (Katoh *et al.*, 2002) and ITS2 consensus sequences were aligned using

FAST 1.15.7 (Bradley *et al.*, 2009), as it outperformed MAFFT when dealing with ITS2 indel variation. Both alignments were then checked by eye. Haplotypes for ITS2 sequences that were multi-site heterozygotes for single nucleotide polymorphisms (SNPs) were resolved either by direct comparison to homozygous sequences or with PHASE 2.1.1 (Stephens *et al.*, 2001; Stephens & Scheet, 2005). The web tool SEQPHASE (Flot, 2010) was used to create PHASE input files and to interpret PHASE output files. Haplotype determination for out of phase sequence traces, caused by alleles with indel variation, was inferred from related homozygous sequences, or with the program Indelligent (Dmitriev & Rakitov, 2008).

The total number of variable and parsimony informative sites, average and maximum pairwise genetic distances (both uncorrected and corrected) overall, within and between species were computed with MEGA 6.0 (Tamura *et al.*, 2013). The entropy-based index as implemented in DAMBE 5.2.78 (Xia *et al.*, 2003) was used to assess substitution saturation within the mtDNA and ITS2 genes and the RPD3 software (Martin *et al.*, 2010) was applied to detect potential historical recombination events within the ITS2 nuclear gene.

Evolutionary tree and haplotype network construction

Bayesian inference (BI) phylogenetic analyses were performed for the mtDNA and nuclear gene sequence alignments separately, using the parallel version of MRBAYES 3.2.1 (Ronquist, 2012). Eight analyses were run each for 10 million generations using 8 MCMC (Markov chain Monte Carlo) chains, discarding 25% of samples as burn-in. For both genes the general time reversible model of sequence evolution with a gamma

correction (GTR + G) was used without modelling invariant sites (I), as recommended by Ren et al., (2005), Siekmann et al., (2012) and Jia et al., (2014, and references therein), with priors set to the default values. Trees were rooted with *L. vicinus*. The parameter estimates were assessed for stationarity and convergence in TRACER v.1.5 (Rambaut & Drummond, 2009) with only estimated sample size (ESS) above 200 for all parameters being accepted. Trees were visualised in FIGTREE 1.3.1 (Rambaut, 2011). TCS v.1.21 (Clement, 2000) was employed to infer haplotype networks of the less divergent nuclear gene sequences using statistical parsimony (Templeton *et al.*, 1992) with 95% confidence limit, and the software HAPSTAR (Teacher & Griffiths, 2011) was used to draw the network.

Dating analysis

When trying to identify the evolutionary processes that shape species formation, distinguishing between introgression and incomplete lineage sorting is generally difficult (Toews & Brelsford, 2012). However oceanic islands offer an ideal framework to unravel these processes, mainly due to the dynamics of speciation by founder events. In an insular setting, incomplete lineage sorting is of minor consequence within species derived from a limited number of founding individuals from a different island. Thus, gene flow among species that are derived from independent colonisation events can be evaluated in a background of limited or no incomplete lineage sorting (e.g. Jordal *et al.*, 2006; Nietlisbach *et al.*, 2013). To further evaluate the possibility of incomplete lineage sorting, we use the spatial context provided by islands and the temporal information provided by DNA sequences to infer the timing and direction of colonisation events that are needed to

explain shared genetic variation between species on different islands. In the absence of a calibrated rate for *Laparocerus*, we applied a general coleopteran COII mutation rate of 0.015 substitution/site/myr (Cicconardi *et al.*, 2010) with a restricted uniform distribution interval [0.0149, 0.0151]. Our primary interest is not the absolute timing of lineage colonisation, but the relative timing. Thus absolute age estimates are intended to provide temporal context, under the untested assumption that the *Laparocerus* COII mutation rate is consistent with the general coleopteran rate estimate of Cicconardi *et al.* (2010). We use a strict molecular clock, as this is preferred over a relaxed clock for intraspecific genetic variation (Barnett *et al.*, 2014). The MRBAYES output trees for the mtDNA gene partition were used as starting trees for the dating analyses with BEAST 1.7.3 (Drummond *et al.*, 2012). Analyses used a GTR+G substitution model with 4 gamma categories, a Yule tree prior, and nodes with posterior probability (PP) of 0.90 or higher constrained to be monophyletic.

To account for the possibility of either the extinction or non-sampling of molecular lineages, nodes representing both the earliest and the most recent possible lineage colonisation times were estimated, and the time intervals among colonising lineages compared. For a more detailed discussion of this approach see Emerson (2002), Figure 6. Input files were generated in BEAUTI 1.7.3 (Drummond *et al.*, 2012), and 15 runs of 100 million generations each, sampled every 1000 generations, were performed and combined, checking sampling, mixing and convergence to a stationary distribution with TRACER 1.5 (Rambaut & Drummond, 2009).

RESULTS

Mitochondrial COII gene

All but one of the 173 specimens were successfully amplified and sequenced for the mitochondrial COII gene, producing an alignment of 172 sequences of 633bp. Across the ingroup, 122 polymorphic sites (of which 96 were parsimony informative) and 76 unique alleles were identified. The average pairwise p-distance was 4% across all species with a maximum of 6.9%. Within species, average pairwise p-distances ranged from 0.7% (*L. osorio*) to 3.9% (*Laparocerus* sp. 1), and between species, average pairwise p-distances ranged from 1.7% (*L. osorio* and *L. microphthalmus*) to 5.2% (*L. bimbache* and *L. tirajana*) (Table 2).

Xia's index for substitution saturation produced values of 0.015 (first and second codon positions) and 0.16 (third codon position) which were significantly lower than the critical value for symmetric topologies (0.69-0.79, $P < 0.001$; 0.64-0.77, $P < 0.001$, respectively), suggesting that sites have reached little saturation and sequences can be reliably used for phylogenetic reconstruction.

Nuclear ITS2 gene

All but two of the 173 sampled beetles were successfully amplified and sequenced for the ITS2 gene. Of the 170 beetles of the ingroup, 64 were homozygotes and 106 were heterozygotes. All but five heterozygotes were successfully resolved for indel and SNP variation (phase threshold $> 85\%$). The five unphased samples were removed from the dataset. The number of SNPs per heterozygote ranged from 1 to 6. Indels ranged in length from 1-19 bp and only one individual presented more than a single

indel differing between its two alleles. The ITS2 alignment consisted of 330 sequences resulting in a final alignment of 411 bp after the removal of two variable poly-A regions, and a hyper-variable region of 35 bp that could not be aligned. Across the ingroup there were 48 polymorphic sites (of which 24 were parsimony informative) and 52 unique alleles were identified. The average pairwise p-distance was 1%, with a maximum of 3.4%. Within species, average pairwise p-distance ranged from 0.3% (*L. tirajana*) to 1.4% (*L. sp. aff. tirajana*), and between species, average pairwise p-distance ranged from 0.4 (*Laparocerus* sp. 1 and *L. bimbache*) to 2.5 % (*L. microphthalmus* and *L. tessellatus*) (Table 3).

Xia's index for substitution saturation, performed on all sites for the nuclear ITS2 dataset, produced values of 0.10-0.15 which were significantly lower than the critical values for symmetric topologies (0.69-0.79, $P < 0.001$), suggesting that sites have reached little saturation and sequences can be reliably used for phylogenetic reconstruction. No signal of recombination was detected for the ITS2 gene for any of the methods used.

Phylogenetic analysis of mtDNA COII gene sequences

The bayesian tree of mtDNA sequence data reveals most of the morphologically described species within the *L. tessellatus* complex to be paraphyletic or polyphyletic for this gene (Fig 2). Four major groups are described by the mtDNA sequence data: two poorly supported but geographically distinct clades (1 and 2) and two well supported clades (3 and 4). Within the island of Gran Canaria, *L. osorio* is the only species clearly recovered as monophyletic (Posterior Probability, PP=0.98). The two

species from Tenerife, *L. freyi* and *L. tessellatus* are polyphyletic and share four mtDNA haplotypes (h15, h25, h29 and h76). *Laparocerus* sp. 1 from La Palma is polyphyletic, originating from three founding mtDNA lineages (uncorrected p-distance among lineages ranges from 3.2 to 6.3%). The first of these lineages (Lap-1) comprises clade 3 with sequences from Gran Canaria (PP=1). The second (Lap-2) and third (Lap-3) La Palma lineages comprise clade 4 with sequences from Tenerife and El Hierro (PP=1). Within clade 4, lineage Lap-2 forms a moderately supported clade with sequences from Tenerife (PP=0.86), while Lap-3 contains sequences from El Hierro and forms a highly supported clade with sequences from Tenerife (PP=0.95). Similar to La Palma, mtDNA sequences from the El Hierro species *L. bimbache* are polyphyletic with two lineages of independent origin (uncorrected p-distance between lineages 5.3%). The first of these (Bim-1) forms a strongly supported clade (PP=1) with sequences from La Palma while the second lineage (Bim-2) is more closely related to sequences from Gran Canaria.

Phylogenetic and network analyses of nuclear ITS2 sequences

Within the bayesian ITS2 tree the earliest branching events are comprised solely of DNA sequences from Gran Canaria, with the remaining sequences from Gran Canaria forming a well supported clade (clade C, PP=1) with sequences from the remaining islands (Fig. 3). Only *L. microphthalmus* and *L. tirajana* from Gran Canaria are recovered as monophyletic. The Tenerife species *L. freyi* and *L. tessellatus* are polyphyletic and share six haplotypes (h6, h20, h21, h24, h29, and h28). For a more resolved understanding of haplotype relationships within the clade comprising sequences from all islands (clade C), they were used for a network analysis, with

closely related sequences from *L. osorio* (h48, h49 and h50) included as an outgroup
 to provide temporal information regarding derived and ancestral haplotypes (e.g.
Zarza et al., 2008). A single network with no reticulations among haplotypes was
 recovered (Fig. 4) with haplotype h20 from Tenerife identified as the most recent
 common ancestor (mrca) haplotype within the clade. Haplotype h20 would appear to
 be either unsampled, or extinct on Gran Canaria, given that *L. sp. aff. tirajana*
 haplotypes (h42 and h43) are directly derived from it. Two haplotypes are shared
 between Tenerife and La Palma (h2 and h6), with one of these (h2) also being shared
 with El Hierro (Fig.4).

Dating analysis

The BEAST analysis of mtDNA COII sequences yielded an estimate of
 approximately 6.32 Ma [95% highest posterior density (HPD): 4.42-8.32 Ma] for the
 divergence of the *L. tessellatus* complex from *L. vicinus*, with initial diversification
 within the complex estimated at approximately 2.71 Ma [HPD 2.08-3.38]. These two
 absolute age estimates are tentative, as they are reliant on a general coleopteran
 mtDNA rate, and thus should be viewed as uncertain. Other node age estimates are
 used to compare earliest and most recent possible times for inter-island DNA among
 different sequence colonisation events (Fig. 2 and Table 4).

Lap-2 and Lap-3 are derived within lineage 4, a clade of mitochondrial diversity
 otherwise endemic to Tenerife (Fig. 2), indicative of La Palma having been colonised
 by two mitochondrial lineages from Tenerife. Lap-2 is estimated to have colonised
 between 0.21 Ma [0.06-0.4] (node D) and 0.07 Ma [0.02-0.16] (node E), and Lap-3

between 0.98 Ma [0.63-1.35] (node F) and 0.38 Ma [0.21-0.55] (node G). All remaining La Palma sequences occur together with Gran Canaria sequences in lineage 3. The absence of basal resolution within this lineage means that it is not possible to infer if La Palma sequences are derived from Gran Canaria, or visa versa. If La Palma sequences are derived from Gran Canaria, then this colonisation is estimated to have occurred approximately 0.96 Ma [0.66-1.28] (node C). The alternative scenario of a La Palma origin would mean that La Palma is estimated to have been colonised at least 0.96 Ma [0.66-1.28] from an undetermined source.

Bim-1 and Bim-2 are derived within lineages of sequences otherwise endemic to La Palma and Gran Canaria respectively, indicating a colonisation of El Hierro by a La Palma lineage between 0.29 Ma [0.15-0.45] (node H) and 0.12 Ma [0.02-0.23] (node I) and by a lineage from Gran Canaria between 0.4 Ma [0.14-0.71] (node J) and 0.08 Ma [0.015-0.19] (node K).

DISCUSSION

We investigated the history of diversification within the *L. tessellatus* species complex in the Canary Islands using a combination of sequence data from one mitochondrial and one nuclear gene. Both gene trees are consistent with the progression rule hypothesis (Funk & Wagner, 1995), where younger islands are colonised from older islands. However, the two gene trees present very different topologies, with a rather simple colonisation history required to explain the pattern of nuclear gene relationships, while the mtDNA gene tree implicates a much more complex history of colonisation. Using the geographic context of the islands

themselves, and relative temporal information from the gene trees, we were able to identify the geographic origin of the complex, and dismiss explanations of incomplete lineage sorting to reveal a history of colonisation and speciation involving genetic admixture.

Geographic origin of the L. tessellatus complex

The nuclear sequence data indicates the geographic origin of the complex to be Gran Canaria. The earliest branching events within the ITS2 tree are uniquely composed of individuals from Gran Canaria, with sequences from all other islands restricted to a single well-defined clade that also includes DNA sequences from Gran Canaria (clade C, Fig. 3). The complex relationships within the clade composed of sequences from all islands are best understood when viewed as a network (Fig. 4). All nuclear ITS2 DNA sequence variation sampled on the islands of Tenerife, La Palma and El Hierro is derived from a single ancestral sequence that is identified as haplotype h20 from Tenerife. Given the inferred Gran Canaria origin for the complex, the most parsimonious explanation for the absence of h20 on Gran Canaria is that it is either unsampled or extinct. The alternative explanation that h20 was never present on Gran Canaria would require h42 and h43 to have both colonised Gran Canaria from Tenerife, with these two haplotypes being unsampled or now extinct on Tenerife. This alternative explanation is less parsimonious as it requires two colonisation events (c.f. 1), two unsampled or extinct haplotypes (c.f. 1), and would implicate admixture for *L. tirajana*. Relationships among mtDNA sequences within the complex provide little evidence that can be used to infer the geographic origin of the group. However, a Bayesian Inference phylogenetic ana

lysis of 243 *Laparocerus* OTUs, including almost all described species of *Laparocerus* in the Canary Islands, places the complex within a well-supported mtDNA clade (PP=1) comprised of an additional 19 species endemic to Gran Canaria, with only 3 non-Gran Canarian species (A. Machado, unpublished data), consistent with a Gran Canarian origin for the complex.

La Palma and El Hierro – single species with multiple origins

The ITS2 sequence variation on La Palma and El Hierro can be explained by a single source island in each case. ITS2 sequences from La Palma are consistent with Tenerife as a source, while those of El Hierro can be explained by an origin from either La Palma or Tenerife (Figs. 3 & 4). While more complex scenarios are also possible (e.g. El Hierro colonised from both Tenerife and La Palma), a simple single island origin is not rejected by the data. In contrast, mitochondrial relationships are in sharp disagreement with a simple colonisation scenario. *Laparocerus* sp. 1 (La Palma) is composed of multiple mitochondrial lineages, derived from either Tenerife (*L. freyi* and *L. tessellatus*) or Gran Canaria (*L. tirajana* and *L. sp. aff. tirajana*) (Fig. 2). *Laparocerus bimbache* (El Hierro) is composed of two mitochondrial lineages, one derived from *Laparocerus* sp. 1 (La Palma), while the other is derived from within *L. microphthalmus* (Gran Canaria) (Fig. 2). This pattern of mixed ancestry suggests either: (i) species origin on La Palma and El Hierro involving genetic admixture from multiple founding species (Fig. 5), or (ii) incomplete lineage sorting (Funk & Omland, 2003; Toews & Brelsford, 2012).

There is increasing evidence for genetic admixture following multiple founding events within oceanic island settings. Jordal *et al.* (2006) revealed historical gene flow among at least two founding populations of the *Aphanarthrum glabrum* species complex on the island of La Palma, followed by the evolution of reproductive isolation. Historical genomic admixture among species originating from different islands is invoked to explain discordant mito-nuclear tree topologies for a subspecies of the Galápagos mockingbird genus *Mimus* (Nietlisbach *et al.*, 2013) and species within the cricket genus *Lapaula* in the Hawaiian islands (Shaw, 2002). In contrast, Garrick *et al.* (2014) describe extensive and ongoing introgression within the Galápagos giant tortoise species *Chelonoidis becki* on the island of Isabela involving two founding populations of *Chelonoidis darwini* from the island of Santiago. It has recently been documented within Darwin's finches of the Galápagos archipelago that there has been extensive interspecific gene flow subsequent to independent colonisation events to the same island, with hybridisation giving rise to species of mixed ancestry (Lamichhaney *et al.*, 2015).

We argue that the genetic data for the *Laparocerus tessellatus* complex species of La Palma and El Hierro is also a consequence of genetic admixture involving multiple founding species, rather than incomplete lineage sorting. Distinguishing between these two processes is typically a challenge, but the dynamics of speciation by founder events between islands enables us to exclude lineage sorting by evaluating three expectations from incomplete lineage sorting. We deal with each of these three expectations below in turn.

Expectation 1: All sequence variation shared among islands was present in the ancestral gene pool. This would require colonisation events between islands involving large numbers of founding individuals, such that a substantial amount of the standing ancestral genetic variation would be transferred. This is not consistent with the colonisation dynamics of flightless beetles (Ikeda *et al.*, 2012; Vogler & Timmermans, 2012). Even if such a colonisation dynamic were to have occurred, it would then need to be followed by non-random extinction (or sampling) of genetic variation within both the source and founded islands. As an example, the following scenario would be necessary to explain patterns of genetic diversity on La Palma if *Laparocerus* sp. 1 was the result of a single colonisation event from either Gran Canaria or Tenerife. Extant sequence variation within lineage 4 (Fig. 2) prior to the colonisation event would have been present within the ancestral population of Gran Canaria. Additionally, sequence variation within lineage 3 (Fig. 2) would have been present within Tenerife. Reciprocal and extensive mtDNA lineage extinction on both islands would then be required to explain the absence of Tenerife sequences within lineage 3, and the absence of Gran Canaria sequences in lineage 4. We have formally evaluated incomplete lineage sorting as an explanation for non-shared genetic variation between Gran Canaria and Tenerife, which is clearly rejected (Appendix S2).

Expectation 2: If the species from La Palma and El Hierro are each the product of a single colonisation event, colonisation times of all molecular lineages shared between these and other islands should coincide. The relative time interval for the colonisation of a lineage to an island ranges from the lowest to the highest bound of the two time intervals (earliest and most recent possible) measured for each lineage, e.g. La Palma was colonised from one Tenerife lineage between 0.02-0.4 Ma (minimum and

maximum HPD values for nodes D and E, Fig. 2). There is substantial overlap in the time intervals estimated for different lineages colonising La Palma and El Hierro, and this is a consequence of the wide 95% posterior intervals estimated by the Bayesian MCMC method when dealing with shallow genetic divergence (Brown & Yang, 2010). Despite this difficulty, results clearly reject the hypothesis of a single colonisation for La Palma. MtDNA sequence diversity within Lap-2 is estimated to be derived from a colonisation event no older than 0.4 Ma (node D, Fig. 2). In contrast, mtDNA sequence diversity within Lap-1 is estimated to either be derived from a colonisation event from Gran Canaria no younger than 0.66 Ma (node C, Fig. 2), or to have originated much earlier than this date. Thus temporal information alone indicates that La Palma was colonised at least two times.

Expectation 3: A signature of incomplete lineage sorting for nuclear gene is expected to be more exaggerated than for the mitochondrial gene, due to its much bigger effective population size (Ballard & Whitlock, 2004). In contrast to this expectation, the signature of potential incomplete lineage sorting among islands presented by the nuclear (Fig. 4) is much less than expected if the patterns of mtDNA relatedness are due to incomplete lineage sorting. If we consider the shared mtDNA variation observed between Gran Canaria and both La Palma and El Hierro to be the result of incomplete lineage sorting, its absence within the more slowly evolving nuclear gene, with its much larger effective population size becomes difficult to explain. It should be noted that there are phenomena that can tend to increase the effective population size of mtDNA relative to nuclear DNA, such as female biased sex ratios, or mating systems that favour high variance in male reproductive success. While there is no information on mating system, sex ratios estimated from 142 individuals of *L.*

tessellatus, 67 *L. obsitus* and 88 *Laparocerus* sp. 1 are 0.9, 1.0 and 0.8 respectively (A. Machado, unpublished data), indicate a slight, or no bias, within species.

Species boundaries in Tenerife

Most DNA sequences sampled within *L. tessellatus* were also found in *L. freyi* suggesting that *L. tessellatus* may be of recent origin from an *L. freyi* like ancestor. Of the five mitochondrial and eight nuclear sequences sampled from *L. tessellatus*, four and six are shared with *L. freyi*. While the two species are not distinct at the genetic level, there is an interesting genetic observation at two sampling sites where two divergent mtDNA lineages were sampled in sympatry. Both sites are represented by only one of the two species. This observation provides an opportunity to evaluate whether mtDNA divergence within Tenerife may be associated with cryptic species (see details in Appendix 2). While at site TF48 (*L. freyi*), the two mtDNA lineages show no evidence that they might represent biological species, at site TF45 (*L. tessellatus*), there is a strong nuclear signature of limited gene flow among individuals from these two mtDNA lineages. This complex genetic pattern may be indicative of differences in environmental backgrounds facilitating contact and gene exchange between divergent gene pools at one site, while enhancing isolation and limiting gene flow at another. A similar phenomenon has been reported for Towhee birds *Pipilo maculatus* and *P. ocai* in southern Mexico, which hybridize extensively in several sites, but coexist in sympatry with no evidence of gene flow in other sites (Sibley & Sibley, 1964). However, FIS values at both sites, as well as three further sites sampled for only a single mtDNA lineage, were all found to be significantly positive (Appendix 2), indicating that rather than species boundaries, apparent nuclear genetic

differentiation between divergent sympatric mtDNA lineages could more plausibly be a consequence of non-random mating within sampling sites. Further sampling will be required to understand what drives the structuring of genetic variation within sampling sites.

Conclusions

Species that are the product of genetic admixture among multiple colonising lineages within island archipelagos, as demonstrated here, have rarely been reported for animal taxa. However, this phenomena may not be unexpected, because island archipelagos present a geographic matrix where both geographic isolation and secondary contact are a function of colonisation dynamics and frequency in space and time (Emerson & Faria, 2014). High colonisation rates among islands will push populations toward panmixia, while low colonisation rates will facilitate divergence. At intermediate rates, where genomic divergence is promoted, but where barriers to gene flow may not have reached completion, the potential for admixture may arise. Indeed, genomic admixture following multiple founding events has been demonstrated within Galápagos giant tortoises (Garrick *et al.*, 2014), and invoked to explain discordant mito-nuclear tree topologies for a subspecies of the Galápagos mockingbird genus *Mimus* (Nietlisbach *et al.*, 2013). Our findings, like those of Nietlisbach *et al.* (2013) and Garrick *et al.* (2014) raise an interesting question – to what extent might admixture be a driver of diversification itself? The potential importance of genetic admixture as a driver of speciation is well recognised (Shaw, 2002; Mallet, 2007; Schwenk *et al.*, 2008). In the context of island colonisation, founding events that involve only one or a few individuals will result in low genetic diversity within the

549 founding populations, which can only be recovered over a mutational evolutionary
550 time-scale mutation. Genetic admixture provides a potential escape from reduced
551 genetic variance via recombination among divergent genomes that may also facilitate
552 adaptation within novel adaptive landscapes (Mallet, 2007; Emerson & Faria, 2014).
553 Further work is needed to address this issue, and new techniques such as reduced
554 genome sequencing and genotyping by sequencing (e.g. Mastretta-Yanes *et al.*, 2014)
555 should prove very useful in this respect.

557 **Acknowledgements**

559 We are grateful to the constructive comments of three anonymous referees, which
560 helped to improve previous version of this manuscript. This study was supported by
561 FCT-PTDC/BIA-BEC/104571/2008 granted to I.R.A and P.A.V.B, by Spanish
562 MINECO grant CGL2013-42589-P awarded to B.C.E., and by the School of
563 Biological Sciences, University of East Anglia. C.M.A.F was funded by a Brazilian
564 Council for Scientific and Technological Development PhD studentship (CNPQ –
565 200439/2010-3).

References

- Amorim, I.R., Emerson, B.C., Borges, P.A.V. & Wayne, R.K. (2012) Phylogeography and molecular phylogeny of Macaronesian island *Tarphius* (Coleoptera: Zopheridae): why are there so few species in the Azores? *Journal of Biogeography*, **39**, 1583–1595.
- Arechavaleta, M., Rodríguez, S., Zurita, N. & García, A. (2010) *Lista de especies silvestres de Canarias. Hongos, plantas y animales terrestres. 2009*. Gobierno de Canarias.
- Ballard, J.W.O. & Whitlock, M.C. (2004) The incomplete natural history of mitochondria. *Molecular Ecology*, **13**, 729–744.
- Barnett, R., Yamaguchi, N., Shapiro, B., Ho, S.Y.W., Barnes, I., Sabin, R., Werdelin, L., Cuisin, J. & Larson, G. (2014) Revealing the maternal demographic history of *Panthera leo* using ancient DNA and a spatially explicit genealogical analysis. *BMC Evolutionary Biology*, **14**:70.
- Bradley, R.K., Roberts, A., Smoot, M., Juvekar, S., Do, J., Dewey, C., Holmes, I. & Pachter, L. (2009) Fast Statistical Alignment. *Plos Computational Biology*, **5**, e1000392.
- Brown, R.P. & Yang, Z. (2010) Bayesian dating of shallow phylogenies with a relaxed clock. *Systematic Biology*, **2**, 191–131.
- Cicconardi, F., Nardi, F., Emerson, B.C., Frati, F. & Fanciulli, P.P. (2010) Deep phylogeographic divisions and long-term persistence of forest invertebrates (Hexapoda: Collembola) in the North-Western Mediterranean basin. *Molecular Ecology*, **19**, 386–400.
- Clement, M., D. Posada, and K. A. Crandall (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.
- Dmitriev, D.A. & Rakitov, R.A. (2008) Decoding of Superimposed Traces Produced by Direct Sequencing of Heterozygous Indels. *Plos Computational Biology*, **4**, e1000113.
- Drummond, A.J., Suchard, M.A., Xie, D. & Rambaut, A. (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular biology and evolution*, **29**, 1969–73.
- Drummond, A.J., Ashton, B., Buxton, S., Cheung, M., Cooper, A., Heled, J., Kearse, M., Moir, R., Stones-Havas, S., Sturrock, S., Thierer, T. & Wilson, A. (2010) Geneious v5.3. In: Available from <http://www.geneious.com>
- Emerson, B.C. (2002) Evolution on oceanic islands: molecular phylogenetic approaches to understanding pattern and process. *Molecular Ecology*, **11**, 951–966.
- Emerson, B.C. & Oromí, P. (2005) Diversification of the forest beetle genus *Tarphius* on the Canary Islands, and the evolutionary origins of island endemics. *Evolution*, **59**, 586–598.
- Emerson, B.C. & Faria, C.M.A. (2014) Fission and fusion in island taxa - serendipity, or something to be expected? *Molecular Ecology*, **23**, 5132–5134.
- Emerson, B.C., Oromí, P. & Hewitt, G.M. (2000) Interpreting colonisation of the *Calathus* (Coleoptera: Carabidae) on the Canary Islands and Madeira through the application of the parametric bootstrap. *Evolution*, **267**, 2199–2205.
- Flot, J.F. (2010) SEQUENCE: a web tool for interconverting phase input/output files and fasta sequence alignments. *Molecular Ecology Resources*, **10**, 162–166.

- 616 Funk, D.J. & Omland, K.E. (2003) Species-level paraphyly and polyphyly: Frequency,
617 causes, and consequences, with insights from animal mitochondrial DNA.
618 *Annual Review of Ecology Evolution and Systematics*, **34**, 397-423.
- 619 Funk, V.A. & Wagner, W.L. (1995) Biogeographic patterns in the Hawaiian islands.
620 *Hawaiian biogeography: Evolution on a hot spot archipelago* (ed. by W.L.
621 Wagner and V.A. Funk), pp. 379-419. Smithsonian Inst Press, Washington
622 DC.
- 623 Garrick, R.C., Benavides, E., Russello, M.A., Hyseni, C., Edwards, D.L., Gibbs, J.P.,
624 Tapia, W., Ciofi, C. & Caccone, A. (2014) Lineage fusion in Galapagos giant
625 tortoises. *Molecular Ecology*, **23**, 5276-5290.
- 626 Ikeda, H., Nishikawa, M. & Sota, T. (2012) Loss of flight promotes beetle
627 diversification. *Nature Communications*, **3**, 648.
- 628 Jordal, B.H., Emerson, B.C. & Hewitt, G.M. (2006) Apparent 'sympatric' speciation in
629 ecologically similar herbivorous beetles facilitated by multiple colonizations
630 of an island. *Molecular Ecology*, **15**, 2935-2947.
- 631 Katoh, K., Misawa, K., Kuma, K. & Miyata, T. (2002) MAFFT: a novel method for
632 rapid multiple sequence alignment based on fast Fourier transform. *Nucleic
633 Acids Research*, **30**, 3059-3066.
- 634 Lamichhaney, S., Berglund, J., Almen, M.S., Maqbool, K., Grabherr, M., Martinez-
635 Barrio, A., Promerova, M., Rubin, C.-J., Wang, C., Zamani, N., Grant, B.R.,
636 Grant, P.R., Webster, M.T. & Andersson, L. (2015) Evolution of Darwin's
637 finches and their beaks revealed by genome sequencing. *Nature*, **518**, 371-375.
- 638 Losos, J.B. & Ricklefs, R.E. (2009) Adaptation and diversification on islands. *Nature*,
639 **457**, 830-836.
- 640 Machado, A. (2003) Sobre el método de coleccionar *Laparocerus* Schönherr, 1834 y el
641 reconocimiento de sus marcas en las hojas (Coleoptera, Curculionidae). *Vierae*,
642 **31**, 407-420.
- 643 Machado, A. (2006) The type material of the species of *Laparocerus* Schonherr, 1834
644 (Coleoptera, Curculionidae, Entiminae). *Journal of Natural History*, **40**, 2001-
645 2055.
- 646 Machado, A. (2009) New *Laparocerus* Schoenherr, 1834 from La Palma, Canary
647 Islands (Coleoptera, Curculionidae, Entiminae). *Graellsia*, **65**, 183-224.
- 648 Machado, A. (2011) Los *Laparocerus* Schoenherr, 1834 de las Canarias orientales,
649 islas Salvajes y Marruecos (Coleoptera, Curculionidae, Entiminae). *Graellsia*,
650 **67**, 205-241.
- 651 Machado, A., López, M., Almeida, T. & Hernández, M. (2008) Mitochondrial DNA
652 phylogenetic analysis of the genus *Laperocerus* (Coleoptera, Curculionidae,
653 Entiminae). I. The Madeiran clade. *Zoology Scripta*, **37**, 415-427.
- 654 Mallet, J. (2007) Hybrid speciation. *Nature* **446**, 279-283.
- 655 Martin, D.P., Lemey, P., Lott, M., Moulton, V., Posada, D. & Lefevre, P. (2010)
656 RDP3: a flexible and fast computer program for analyzing recombination.
657 *Bioinformatics*, **26**, 2462-2463.
- 658 Mastretta-Yanes, A., Arrigo, N., Alvarez, N., Jorgensen, T.H., Piñero, D. & Emerson,
659 B.C. (2014) RAD sequencing, genotyping error estimation and de novo
660 assembly optimization for population genetic inference. *Molecular Ecology
661 Resources*, DOI: **10.1111/1755-0998.1229**
- 662 Nietlisbach, P., Wandeler, P., Parker, P.G., Grant, P.R., Grant, B.R., Keller, L.F. &
663 Hoeck, P.E.A. (2013) Hybrid ancestry of an island subspecies of Galápagos
664 mockingbird explains discordant gene trees. *Molecular Phylogenetics and
665 Evolution*, **69**, 581-592.

- Oromí, P., Cruz, S. & Báez, M. (2010) Psocoptera, Mallophaga, Anoplura, Hemiptera, Thysanoptera, Coleoptera, Strepsiptera, Siphonaptera. *Lista de especies silvestres de Canarias. Hongos, plantas y animales terrestres. 2009* (ed. by M. Arechavaleta, S. Rodríguez, N. Zurita and A. García), pp. 230-302. Gobierno de Canarias, Santa Cruz de Tenerife.
- Rambaut, A. (2011) *FigTree v1.3.1*, Available from <http://tree.bio.ed.ac.uk/software/figtree/>.
- Rambaut, A. & Drummond, S.J. (2009) *Tracer v1.5*. Available from <http://beast.bio.ed.ac.uk/Tracer>.
- Rees, D.J., Emerson, B.C., Oromí, P. & Hewitt, G.M. (2001) Mitochondrial DNA, ecology and morphology: interpreting the phylogeography of the Nesotes (Coleoptera: Tenebrionidae) of Gran Canaria (Canary Islands). *Molecular Ecology*, **10**, 427-434.
- Ronquist, F., M. Teslenko, P. van der Mark, D. L. Ayres, A. Darling, S. Hohna, B. Larget, L. Liu, M. A. Suchard, J. P. Huelsenbeck (2012) MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. *Systematic Biology*, **61**, 539-542.
- Schwenk, K., Brede, N. & Streit, B. (2008) Introduction. Extent, processes and evolutionary impact of interspecific hybridization in animals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **363**, 2805-2811.
- Shaw, K.L. (2002) Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: what mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. *Proceedings of the National Academy of Science, USA*, **99**, 16122-16127.
- Sibley, C.G. & Sibley, F.C. (1964) Hybridization in the red-eyed towhees of Mexico: The populations of the southeastern plateau region. *Auk*, **81**, 479-504.
- Stephens, M. & Scheet, P. (2005) Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *American Journal of Human Genetics*, **76**, 449-462.
- Stephens, M., Smith, N.J. & Donnelly, P. (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, **68**, 978-989.
- 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.
- Teacher, A.G.F. & Griffiths, D.J. (2011) HapStar: automated haplotype network layout and visualization. *Molecular Ecology Resources*, **11**, 151-153.
- Templeton, A.R., Crandall, K.A. & Sing, C.F. (1992) A cladistic-analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA-sequence data. 3. Cladogram estimation. *Genetics*, **132**, 619-633.
- Toews, D.P.L. & Brelsford, A. (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, **21**, 3907-3930.
- Villacorta, C., Jaume, D., Oromí, P. & Juan, C. (2008) Under the volcano: phylogeography and evolution of the cave-dwelling *Palmarchestia hypogaea* (Amphipoda, Crustacea) at La Palma (Canary Islands). *BMC Biology*, **6**: 7.
- Vogler, A.P. & Timmermans, M.J. (2012) Speciation: don't fly and diversify? *Current Biology*, **22**, R284-R286.
- Xia, X.H., Xie, Z., Salemi, M., Chen, L. & Wang, Y. (2003) An index of substitution saturation and its application. *Molecular Phylogenetics and Evolution*, **26**, 1-7.

Zarza, E., Reynoso, V.H. & Emerson, B.C. (2008) Diversification in the northern neotropics: mitochondrial and nuclear DNA phylogeography of the iguana *Ctenosaura pectinata* and related species. *Molecular Ecology*, **17**, 3259-3275.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Table S1 and Details of PCR amplification and sequencing used in this study.

Appendix S2 Simulation of stochastic lineage sorting and Analysis of species boundaries.

Appendix S3 Full Bayesian phylogenetic trees of the mtDNA COII data set (Fig. S1) and ITS2 nuclear gene dataset (Fig. S2).

Data Accessibility

Mitochondrial DNA sequences and ITS2 nuclear sequences used in this study are available for download at GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>) under accession numbers #####

Biosketch

Christiana M. A. Faria is a PhD student at the Centre of Ecology, Evolution and Conservation, University of East Anglia. She is interested in applying molecular phylogenetic methods to studies of evolutionary ecology, biogeography, species adaptation and diversification processes.

742 Author contributions: C.M.A.F carried out molecular work and analysed the data;
743 C.M.A.F and B.C.E. wrote the manuscript, and all authors discussed the data and
744 commented on the manuscript; A.M. collected the samples; B.C.E. conceived and
745 supervised the project.

746

747 Editor: Aristeidis Parmakelis

748

749 Table 1 - Details of sampling within the Canary Islands for species within the *Laparocerus tessellatus* complex. Sites coded according to Fig 1.
 750 Geographic coordinates are provided in Table S1.

Species	Site code	Site	<i>n</i> individuals
<i>Laparocerus</i> sp. 1	LP1	La Palma: Montaña de Tagoja, 1250 m	3
	LP2	La Palma: Llanada de Barlovento, 650 m	5
	LP3	La Palma: s. El Paso, 870 m	7
	LP4	La Palma: Breña Alta, Pared Vieja, 1350 m	2
	LP6	La Palma: Mazo: Venijobre, 830 m	29
	LP7	La Palma: El Paso: Mña Don Mendo, 1075 m	3
<i>L. bimbache</i> Machado, 2011	EH9	El Hierro: Monte Ajares, 600 m	2
	EH10	El Hierro: San Andrés: Piedras Blancas, 912 m	2
	EH11	El Hierro: Cruz de Isora, Infra Masilva, 1247 m	2
<i>L. freyi</i> Uyttenboogaart, 1940	TF12	Tenerife: Cumbres de Arico, 1070 m	3
	TF13	Tenerife: Granadilla: Las Vegas, 630 m	6
	TF14	Tenerife: Arico: Contador, 1200 m	2
	TF15	Tenerife: Tanque Bajo, 500 m	1
	TF20	Tenerife: Cumbre de Bolicos, 1200 m	2
	TF21	Tenerife: s. Mña. Bermeja, 1600 m	2
	TF22	Tenerife: Güímar: Bco. del Agua, 700-800 m	9
	TF23	Tenerife: s. Icod El Alto, 1200 m	4

	TF24	Tenerife: Santa Úrsula: Bco. Bensa, 1463 m	3
	TF25	Tenerife: Tacoronte: FuenteFría, 1014 m	1
	TF27	Tenerife: Santa Úrsula: La Corujera, 600 m	3
	TF28	Tenerife: El Portillo, 2000 m	3
	TF48	Tenerife: Las Raíces, 930 m	11
	TF49	Tenerife: Ifonche, 990 m	2
<i>L. tessellatus</i> Brullé, 1839	TF43	Tenerife: Pista LasYedras, 740 m	1
	TF45	Tenerife: Anaga: El Pijaral Km 4.5, 700 m	12
	TF46	Tenerife: Anaga: Chinobre, 900 m	1
	TF47	Tenerife: Anaga: Cruz del Carmen, 900 m	2
<i>L. microphthalmus</i> Lindberg, 1950	GC32	Gran Canaria: Tamadaba NW, 1200 m	8
<i>L. obsitus</i> Wollaston, 1864	GC31	Gran Canaria: Valsendero: Bco. Oscuro, 900 m	2
<i>L. tirajana</i> Machado, 2012	GC35	Gran Canaria: San Bartolomé, Km 1, 940 m	12
	GC36	Gran Canaria: San Bartolomé: Bco. Tirajana, 900 m	1
<i>L. sp. aff. tirajana</i>	GC34	Gran Canaria: Bco. de los Cernícalos, 1400 m	4
	GC37	Gran Canaria: Degollada de Osorio, 875 m	5
	GC39	Gran Canaria: Cumbre: Roque Redondo, 1900 m	5
<i>L. osorio</i> Machado, 2012	GC31	Gran Canaria: Valsendero: Bco. Oscuro, 900 m	6
	GC40	Gran Canaria: Valsendero: Bco. Cazadores, 1080 m	3
	GC41	Gran Canaria: Las Huertecillas, 650 m	1

	GC42	Gran Canaria: Bco. de la Mina, 1200 m	2
<i>L. vicinus</i> Lindberg, 1953	GC31	Gran Canaria: Valsendero: Bco. Oscuro, 900 m	1
751			
752			
753			
754			
755			
756			
757			
758			
759			
760			
761			
762			
763			
764			
765			
766			

Table 2 - Mean pairwise genetic distances between mitochondrial species groups (n=76 alleles) within *Laparocerus tessellatus* complex in the Canary Islands. Fre = *L. freyi*, Lap = *Laparocerus* sp.1, Bim = *L. bimbache*, Tess = *L. tessellatus*, ATir = *L. sp. aff. tirajana*, Tir = *L. tirajana*, Mic = *L. microphthalmus*, Oso = *L. osorio*, Obs = *L. obsitus*.

	Fre	Lap	Bim	Tess	ATir	Tir	Mic	Oso	Obs
Uncorrected p-distances (%)									
Fre	4.5 (2.7)	4.35	4.17	2.77	4.60	4.80	3.83	4.46	4.34
Lap	8.1	8 (3.9)	4.3	4.6	3.6	3.6	4.6	4.8	4.6
Bim	7.6	8.9	7 (3.3)	4.3	5.2	5.2	3.8	4.4	4.1
Tess	4.6	8.7	7.9	5.6 (3.3)	4.7	4.8	4.0	4.8	4.5
ATir	8.7	6.9	10.6	9.0	2 (1.6)	1.7	4.0	4.3	3.7
Tir	9.1	6.9	10.6	9.2	2.3	1.4 (1.1)	4.0	4.6	3.9
Mic	6.5	9.1	7.9	7.1	6.9	7.1	0.9 (0.8)	1.7	1.7
Oso	8.6	9.7	9.0	9.9	7.9	8.7	2.1	0.8 (0.7)	2.5
Obs	7.9	8.8	7.7	8.6	6.2	6.9	2.2	3.6	2.7 (2.1)
Tamura Nei + G corrected (%)									

Average pairwise genetic distances within species are represented in diagonal. Values in diagonal inside brackets and above diagonal represent uncorrected genetic distances (p-distances). Values in diagonal outside brackets and under diagonal represent genetic corrected distances using the Tamura Nei model for nucleotide substitution.

Table 3 - Mean pairwise genetic distances between nuclear species groups (n=52 alleles) within *Laparocerus tessellatus* complex in the Canary Islands. Lap = *Laparocerus* sp.1, Fre = *L. freyi*, Bim = *L. bimbache*, Tess = *L. tessellatus*, ATir = *L. sp. aff. tirajana*, Oso = *L. osorio*, Tir = *L. tirajana*, Mic = *L. microphthalmus*, Obs = *L. obsitus*.

	Fre	Lap	Bim	Tess	ATir	Tir	Mic	Oso	Obs
Uncorrected p-distance (%)									
Fre	0.9 (0.9)	1	0.7	0.9	1.5	1.8	2.5	1.3	1.9
Lap	0.9	0.7(0.7)	0.4	1	1.5	1.8	2.5	1.3	1.9
Bim	0.8	0.4	n/a	0.8	1.3	1.5	2.3	1.1	1.7
Tess	0.9	1.1	0.9	0.9 (0.9)	1.6	1.9	2.5	1.4	1.9
ATir	1.6	1.7	1.4	1.7	1.5 (1.4)	1.4	2	1.1	1
Tir	1.9	1.9	1.6	2	1.5	0.3 (0.3)	2.3	1	1.2
Mic	2.7	2.8	2.5	2.8	2.2	2.5	0.3 (0.3)	1.8	1.6
Oso	1.4	1.4	1.2	1.5	1.1	1	1.7	0.7 (0.7)	0.9
Obs	2.1	2.1	1.9	2.1	1.1	1.3	1.7	1	0.8 (0.8)
TamuraNei + G corrected (%)									

Average pairwise genetic distances within species are represented in diagonal. Values in diagonal inside brackets and above diagonal represent uncorrected genetic distances (p- distances). Values in diagonal outside brackets and under diagonal represent genetic corrected distances using the Tamura Nei model for nucleotide substitution. n/a – no variation within *L. bimbache* (unique haplotype)

791 Table 4 - Estimated relative times in million years ago (Ma) of *Laparocerus tessellatus*
 792 complex mitochondrial lineages expressed as mean values with 95% highest posterior
 793 density (HPD) intervals. ET = earliest possible time; MRT = most recent possible time.
 794 Island codes: EH- El Hierro, GC-Gran Canaria, LP- La Palma, TF-Tenerife. La Palma
 795 lineages= Lap-2 and Lap-3; El Hierro lineages = Bim-1 and Bim-2.

Node	Description	Mean value (Ma)	95% HPD
A	root of the tree	6.32	4.45-8.36
B	Ingroup	2.71	2.08-3.38
C	Divergence of LP sequences from GC sequences	0.96	0.66-1.28
D	ET for colonisation of LP from TF sequences (Lap-2)	0.21	0.06-0.4
E	MRT for colonisation of LP from TF sequences (Lap-2)	0.07	0.02-0.16
F	ET for colonisation of LP from TF sequences (Lap-3)	0.98	0.63-1.35
G	MRT for colonisation of LP from TF sequences (Lap-3)	0.38	0.21-0.55
H	ET for colonisation of EH from LP sequences (Bim-1)	0.29	0.15-0.45
I	MRT for colonisation of EH from LP sequences (Bim-1)	0.12	0.02-0.23
J	ET for colonisation of EH from GC sequences (Bim-2)	0.4	0.14-0.71
K	MRT for colonisation of EH from GC sequences (Bim-2)	0.08	0.015-0.19

796

797

798

799

800

801

Figure Legends

Figure 1. Distribution of sampling sites in the Canary Islands; a complete list of the site names, species collected and number of individuals can be found in Table 1. Non-consecutive numbering is for consistency with an unpublished database. Geographic coordinates (XY) for each site can be found in Table S1. Numbers in parentheses refer to the proposed maximum estimated geological ages of the islands in millions of years ago (Ma).

Figure 2. Bayesian phylogenetic tree inferred from mtDNA COII sequences (633 bp) for the Canary Islands *Laparocerus tessellatus* complex using MrBayes with the GTR+G model of sequence evolution. The tree is rooted with *L. vicinus*. Bayesian posterior probabilities are shown above nodes. Italic numbers at selected nodes indicate BEAST estimated ages of divergence with 95% highest posterior density intervals (see Materials and Methods). Letters and numbers immediately to the right of species names correspond to island codes and sampling sites, see Table 1. Shared haplotypes are highlighted in bold. A full version of the tree with all sampled individuals is provided in Fig. S1 (Supporting Information).

Figure 3. Bayesian phylogenetic tree inferred from ITS2 nuclear gene sequences (411 bp) for the Canary Islands *Laparocerus tessellatus* complex using MrBayes with the GTR+G model of sequence evolution. The tree is rooted with *L. vicinus*. Bayesian posterior probabilities are shown above nodes. Letters and numbers immediately to the right of species names correspond to site codes, see Table 1. Shared haplotypes are

highlighted in bold. A full version of the tree with all sampled individuals is provided in Fig.S2 (Supporting Information).

Figure 4. Haplotype network corresponding to the ITS2 sequence variation in the less divergent group of the *Laparocerus tessellatus* complex (clade C in Fig 3, see text for more details) in the Canary Islands. Colours correspond to islands, and differently coloured segments within circles represent haplotype sharing across islands. The Gran Canarian outgroup *L. osorio* is represented by green haplotypes 48, 49 and 50. Lines represent a mutational step; black circles represent missing or unsampled haplotypes.

Figure 5. Interisland colonisation history within the *Laparocerus tessellatus* complex inferred from DNA sequence data from the mitochondrial COII gene and the nuclear ITS2 gene. The islands of La Palma and El Hierro are inferred to have been colonised from more than one island, suggesting a complex history involving genetic admixture between different founding species for the origin of the single species occurring on each island (see text for details).

Figure 1

Figure 2

Figure 3

892

893

894 Figure 4

Figure 5