MR. GUILLERMO FRIIS (Orcid ID : 0000-0002-0731-6468)

MR. BORJA MILÁ (Orcid ID : 0000-0002-6446-0079)

Article type : Original Article

Genome-wide signals of drift and local adaptation during rapid lineage divergence in a songbird

Running title: Drift and selection in a songbird radiation

Guillermo Friis,[1] Guillermo Fandos,[2] Amanda Zellmer,[3] John McCormack,[3,4] Brant C. Faircloth,[5] Borja Milá[1]

[1]National Museum of Natural Sciences, Spanish National Research Council (CSIC), Madrid 28006, Spain; [2]Department of Biodiversity, Ecology and Evolution, Complutense University of Madrid, Madrid 28040, Spain; [3]Department of Biology, Occidental College, Los Angeles, CA 90041, USA; [4]Moore Laboratory of Zoology and Department of Biology, Occidental College, Los Angeles, CA 90041, USA; [5]Department of Biological Sciences and Museum of Natural Science, Louisiana State University, Baton Rouge, LA 70803, USA

Corresponding author: Guillermo Friis, National Museum of Natural Sciences - CSIC, José Gutiérrez Abascal 2, Madrid 28006, Spain; Email: gfriis@mncn.csic.es; Tel: +34 914111328 x1266.

Abstract

The formation of independent evolutionary lineages involves neutral and selective factors, and understanding their relative roles in population divergence is a fundamental goal of speciation research. Correlations between allele frequencies and environmental variability can reveal the role of selection, yet the relative contribution of drift can be difficult to establish. Recently diversified taxa like the Oregon junco (Aves, Passerellidae, *Junco hyemalis oreganus*) of western North America provide ideal scenarios to apply genetic-environment association analyses (GEA) while controlling for population structure. Analysis of genome-wide SNP loci revealed marked genetic structure consisting of differentiated populations in isolated, dry southern mountain ranges, and less divergent, recently expanded populations in humid northern latitudes. We used correlations between genomic and environmental variance to test for three specific modes of evolutionary divergence: (i) drift in geographic isolation, (ii) differentiation along continuous selective gradients, and (iii) isolation by adaptation. We found evidence of strong drift in southern mountains, but also signals of local adaptation driven by temperature, precipitation, elevation and vegetation, especially when controlling for population history. We identified numerous variants under selection scattered across the genome, suggesting that local adaptation can promote rapid differentiation when acting over multiple independent loci.

Key words: local adaptation, isolation by adaptation, drift, selective gradients, redundancy analysis, postglacial expansion

Introduction

Lineage diversification involves both selective and neutral factors, and elucidating their relative strengths and interactions in the process of evolutionary divergence is essential to understand the mechanisms underlying the early stages of speciation (Coyne & Orr 2004;

Nosil 2012). Lineage differentiation can be driven by divergent natural selection (Coyne & Orr 2004; Darwin 1859), a mechanism that is at the core of 'ecological speciation' models, in which reproductive isolation arises as a by-product of cumulative, ecologically adaptive changes (Mayr 1947; Rundle & Nosil 2005; Schluter 2000). In turn, accumulation of genetic differences caused by drift in geographic isolation or in isolation-by-distance (IBD, Wright 1943; 1946) has been proposed as a mode of divergence driven by neutral factors (Mayr 1954, 1963), a mechanism that can be particularly strong in populations of small effective size (e.g. Carson 1975; Templeton 1981; Uyeda *et al.* 2009). Selection and drift can also act jointly and even interact during evolutionary divergence in a number of ways. Geographic distance usually implies environmental differences that may drive adaptation to local conditions and ecological differentiation, even if populations are connected by moderate gene flow (Egan *et al.* 2015; Rundle & Nosil 2005; Schluter 2000). The evolution of local adaptation can in turn result in isolation-by-adaptation (IBA), a mode of divergence where adaptive changes lead to intrinsic barriers to gene flow, enabling genome-wide differentiation at both neutral and selected loci (Funk *et al.* 2011; Nosil *et al.* 2008; Shafer & Wolf 2013; Wang & Bradburd 2014). Consequently, geographic distance and ecological divergence may promote similar patterns of genetic diversity among populations, so that teasing apart the roles of neutral evolution and ecological adaptation in evolutionary diversification requires approaches that account for both environmental heterogeneity and neutral population structure (Forester *et al.* 2016; Forester *et al.* 2018; Frichot *et al.* 2015; Rellstab *et al.* 2015; Wang & Bradburd 2014).

Our capacity to assess the relative roles of adaptation and neutral differentiation in driving population divergence has benefitted from our increased ability to survey genome-wide variation thanks to the development of next generation sequencing (NGS) techniques (Faria *et al.* 2014; McCormack *et al.* 2013; Seehausen *et al.* 2014). The increasingly large number

of loci afforded by NGS provides improved resolution to detect neutral population structure and patterns of gene flow among differentiated lineages. In addition, highly differentiated loci identified as outliers in an F_{ST} distribution can be interpreted as potential targets of divergent selection standing out in a background of balanced or neutrally maintained genomic variation (Cruickshank & Hahn 2014; Faria *et al.* 2014; Rellstab *et al.* 2015). However, methods of outlier detection relying solely on allele frequencies are sensitive to the confounding effects of historical factors, such as past sudden changes in population size or strong drift in small populations that may result in high rates of false positives (Billiard *et al.* 2005; Christmas *et al.* 2016; Edmonds *et al.* 2004; Kawecki & Ebert 2004). Moreover, changes in allele frequencies due to weak local selective pressures acting over large sets of independent loci can sometimes go undetected by outlier analyses (Bierne *et al.* 2011; Pritchard & Di Rienzo 2010; Rellstab *et al.* 2015). Alternative approaches that integrate environmental parameters by identifying allele frequencies that correlate with ecological variability have proven useful to detect signals of adaptation, especially when selective forces are weak (Forester *et al.* 2018; Frichot *et al.* 2015; Rellstab *et al.* 2015). These methods, known as genetic-environment association (GEA, Hedrick *et al.* 1976; Mitton *et al.* 1977) analyses, have the potential to reveal genetic patterns of differentiation due to local adaptation while testing for the role of multiple, specific environmental variables as drivers of selection. Importantly, GEA methods can correct for population history by controlling for general patterns of neutral genomic variation, allowing us to separate the respective effects of drift and selection in generating and maintaining variability (Forester *et al.* 2016; Forester *et al.* 2018; Rellstab *et al.* 2015). GEA analyses have greatly benefitted from the development of high-throughput sequencing techniques, resulting in a number of studies focusing on the genomic variability associated with environmental parameters in groups as diverse as plants (De Kort *et al.* 2014; Jones *et al.* 2013; Lasky *et al.* 2012; Nadeau *et al.* 2016; Sork *et al.* 2016), fungi (Ojeda

Alayon *et al.* 2017), wolves (Forester *et al.* 2018), lobsters (Benestan *et al.* 2016) and birds (Manthey & Moyle 2015; Safran *et al.* 2016; Szulkin *et al.* 2016; Termignoni-García *et al.* 2017).

Recently diversified systems provide an ideal scenario for studying the relative roles of selective and neutral factors in incipient divergence and speciation (Ahrens *et al.* 2018; Campagna *et al.* 2017). Specifically, GEA methods are particularly suitable when the system under study (i) is composed of closely related populations, among which the signals of selection are still recent and detectable; (ii) includes broad geographic distributions encompassing heterogeneous habitats across ecological clines (*i.e.* selective gradients) but also spatially discontinuous habitats so that adaptive and neutral divergence can be assessed in different spatial settings; (iii) shows large variability in the degree of geographical isolation among populations, from extensive gene flow to total isolation; and (iv) presents low variability in secondary sexual traits so that differential sexual selection can be ruled out as a major driver of population divergence.

The Oregon junco complex (*Junco hyemalis oreganus*) of western North America provides a particularly well suited system to carry out genome-environment association analysis. The complex originated recently as part of the postglacial radiation of dark-eyed juncos across North America following a northward recolonization of the continent as ice sheets retreated after the last glacial maximum, c.a. 18,000 years ago (Friis *et al.* 2016; Milá *et al.* 2016; Milá *et al.* 2007). Among dark-eyed junco forms, the Oregon junco group presents the highest variability in terms of phenotype and ecological range, encompassing a broad latitudinal range from Baja California to Alaska (Fig. 1A). All forms of the Oregon junco share a characteristic dark hood, yet there is considerable population variation in plumage color,

mainly of the hood, dorsum and flanks, and the complex has been traditionally divided into at least 7 subspecific forms (Dwight 1918; Miller 1941; Nolan *et al.* 2002), which include, from south to north: *townsendi*, from the San Pedro Mártir mountains in northern Baja California, Mexico; *pontilis*, distributed just north of *townsendi* in the Sierra Juárez mountains, also in Baja California, Mexico; *thurberi*, from the mountains of southern California and Sierra Nevada; *pinosus*, a coastal form from central California, predominantly distributed in the Santa Cruz mountains; *montanus*, distributed across the interior of Oregon, Washington and British Columbia; *shufeldti*, a more coastal form from Oregon and Washington; and *oreganus* from coastal British Columbia and southern Alaska (Miller 1941; Nolan *et al.* 2002; Fig. 1A, Tables 1 and S1, Supporting Information).

The diverse spatial configuration of populations and environmental variability across the Oregon junco distribution are critical aspects that will affect our capacity to disentangle the roles of adaptive and neutral factors in explaining genomic variance. Here, we use a conceptual framework to classify into three main settings the distinct spatial scenarios observable in the Oregon junco with respect to gene flow and environmental variation. These include (i) geographically isolated populations in similar habitats, as in the case of the Baja California *townsendi* and *pontilis* forms, where limited divergent selection and low rates of gene flow should result in limited adaptive divergence and high neutral divergence by drift (Fig. 2A); (ii) parapatric populations under divergent ecological conditions, as exemplified by the *pinosus* and *thurberi* forms in California, where divergence is expected to increase due to local adaptation, while geographic proximity and moderate gene flow should lead to intermediate levels of neutral differentiation by drift (Fig. 2B); and (iii) populations found along a continuous environmental gradient, as in the northernmost forms of Oregon junco (*thurberi*, *shufeldti*, *montanus* and *oreganus*) where neutral divergence is expected to be low

due to high levels of gene flow, while local adaptation along the gradient may result in a pattern of high differentiation in adaptive variation (Fig. 2C).

We use the Oregon junco complex to study how geographic isolation, population history, and local adaptation to ecological variables have driven population differentiation across the range. We use extensive population sampling and genome-wide SNPs obtained from 'genotyping by sequencing' (GBS, Elshire *et al.* 2011), mapped against a new, high quality *Junco hyemalis* genome sequenced and assembled to be used as a reference. We first assess patterns of neutral genetic structure across the complex using selectively neutral SNPs, and we then look for correlations between environmental variables and allele frequencies across the Oregon junco distribution using two different GEA methods. These include redundancy analysis (RDA, Legendre & Legendre 1998; Van Den Wollenberg 1977), a canonical ordination method that allows computing the proportion of genomic variance explained by a number of select environmental variables while controlling for neutral genetic structure, and BayeScEnv (Villemereuil & Gaggiotti 2015), a method which examines correlation patterns among the same set of ecological variables and a number of potentially adaptive outlier loci identified by an F_{ST}-based method. We also use climatic variables to test for significant niche divergence while controlling for spatial autocorrelation.

Materials and methods

Population sampling

Oregon junco populations were sampled across the geographical range of the species using mist nets in order to obtain biological samples for DNA extraction (Fig. 1A). Each captured individual was aged, sexed, and marked with a numbered aluminum band. A blood sample was collected by venipuncture of the sub-brachial vein and stored in Queen's lysis buffer

(Seutin 1991) or absolute ethanol at -80ºC in the laboratory. Geographic coordinates were stored by means of a GPS device at each sampling site to be used as occurrence points in analyses involving georreferenced ecological data (see below). All sampling activities were conducted in compliance with Animal Care and Use Program regulations at the University of California Los Angeles, and with state and federal scientific collecting permits in the USA and Mexico. A high-quality tissue sample for whole-genome sequencing was obtained from a slate-colored junco (*Junco hyemalis carolinensis*) deposited at the Moore Laboratory of Zoology, Occidental College. Genomic DNA was extracted from blood and tissue samples using a Qiagen DNeasy kit (Qiagen™, Valencia, CA).

Whole-genome sequencing and assembly

We obtained a high quality genome of a *Junco hyemalis* specimen (Moore Laboratory of Zoology voucher number: MLZ:bird:69236) by combining shotgun and proximity ligation libraries at Dovetail Genomics, LLC. Briefly, a *de novo* draft assembly was first built using shotgun, paired-end libraries (mean insert size ~350 bp) and the Meraculous pipeline (Chapman *et al.* 2011). For the Chicago™ and the Dovetail HiC™ library preparation, chromatin was fixed with formaldehyde. Fixed chromatin was then digested with DpnII and free blunt ends were ligated. Crosslinks were reversed and the DNA purified from protein. Resulting nucleic material was then sheared to ~350 bp mean fragment size and sequencing libraries were generated using NEBNext Ultra enzymes and Illumina-compatible adapters. Sequencing of the libraries was carried out on an Illumina HiSeq X platform, producing 218 million 2x151 bp paired-end reads, which provided 233x physical coverage of the genome in the case of the Chicago™ library; and 121 million 2x151 bp paired-end reads in the case of the Dovetail HiC™ library, which provided 18,311x physical coverage of the genome. The *de-novo* draft assembly, shotgun reads, Chicago™ library reads, and Dovetail HiC™ library

reads were then used as input data for HiRise™, a software pipeline designed specifically for using proximity ligation data to scaffold genome assemblies (Putnam *et al.* 2016). The total length of the HiRise™ assembly was 958.32 Mb, with an N50 equal to 71.32 Mb and a total of 4,457 scaffolds. To recover the chromosomal coordinates of the scaffolds obtained with HiRise™, we mapped and oriented them against the zebra finch (*Taeniopygia guttata*) genome v87 available in Ensembl (Yates *et al.* 2016). We used the Chromosembler tool available in Satsuma (Grabherr *et al.* 2010) resulting in a final genome assembly of 955.9 Mb length and a N50 of 71.46 Mb.

Genotyping-by-sequencing

We used genotyping-by-sequencing (Elshire *et al.* 2011) to obtain individual genotypes from 126 Oregon juncos belonging to the following subspecific taxa: *townsendi* (n=16), *pontilis* (n=16), *thurberi* (n=35), *pinosus* (n=16), *montanus* (n=15), *shufeldti* (n=12), *oreganus* (n=16) (Table 1). GBS libraries were prepared and sequenced at Cornell University's Institute for Genomic Diversity, using the restriction enzyme PstI for digestion. Sequencing of the 126 individually-barcoded libraries was carried out in three different lanes (along with other 150 junco samples intended for other studies) of an Illumina HiSeq 2000, resulting in an average of 229.2 million good single-end reads 100 bp in length per lane. Samples of the same form were distributed among at least two of the lanes to avoid sequencing bias, except in the case of *shufeldti* individuals, which were all introduced in the last set of sequenced samples.

Alignment and variant calling

We evaluated GBS read quality using FASTQC after sorting reads by individual with AXE (Murray & Borevitz 2017). We performed the trimming and quality filtering treatment using Trim Galore (Krueger 2015), resulting in a set of reads ranging between 40 and 90 bp long.

Adapter removal stringency was set to 1 and the quality parameter 'q' to 20. The reads were then mapped against the assembled junco genome using the mem algorithm in the Burrows-Wheeler Aligner (BWA, Li & Durbin 2009). Read groups were assigned with Picard Tools version 1.126 (http://broadinstitute.github.io/picard) and BAM files generated for GATK and PCANGSD analysis (see 'Genetic structure analysis' section). We used the HaplotypeCaller tool from the Genome Analysis Toolkit (GATK, McKenna *et al.* 2010) version 3.6-0 to call the individual genotypes, and then the GenotypeGVCFs tool to gather all the per-sample GVCFs files generated in the previous step and produce a set of joint-called SNPs and indels in the variant call format (*vcf*) (GATK Best Practices, Auwera *et al.* 2013; DePristo *et al.* 2011). The subset of 126 samples used in this study included 1,056,359 biallelic SNPs before filtering. Because GBS data do not provide enough coverage for base quality score recalibration, we used VCFTOOLS (Danecek *et al.* 2011) to implement a 'hard filtering' process, customized for each of the downstream analyses (Table 2).

Genetic structure analyses

As a first approach to explore genome-wide population structure among all Oregon junco forms, we ran a principal components analysis (PCA) based on SNP data using the program PCANGSD (Meisner & Albrechtsen 2018). Because it takes genotype uncertainty into account, PCANGSD is well suited for low and medium depth data, similar to the depths of coverage obtained by GBS (Korneliussen *et al.* 2014). Using VCFTOOLS, we identified the 8 samples of each group (16 in the case of the geographically widespread *thurberi*) with the least amount of missing data. PCANGSD input files were produced with ANGSD (Korneliussen *et al.* 2014) by implementing genotype and mapping quality phred score thresholds of 20; a SNP p-value of $1e^{-6}$; a minor allele frequency (MAF) threshold of 0.02, corresponding to at least three copies of any given allele in the dataset; and a missing data

upper limit equal to 50%. In the PCANGSD analysis, we also implemented a threshold for SNPs showing highly significant deviations from Hardy-Weinberg equilibrium (HWE) with a p-value of 10^{-4} to filter out false variants resulting from the alignment of paralogous loci, resulting in a final matrix of 255,163 SNPs. We computed the eigenvectors from the resulting covariance matrix in R version 3.4.2 (R Core Team 2013).

We also examined population structure with the program STRUCTURE (Pritchard *et al.* 2000) using the same set of samples. To reduce computational time, we used a smaller SNP data matrix by applying more conservative filtering thresholds than in the PCA. We constructed a data matrix of biallelic SNPs excluding those with less than 4X and greater than 50X coverage, or those with a genotyping phred quality score below 40. Positions with less than 25% of individuals genotyped for each one of the forms of Oregon junco were removed from the data matrix, along with those presenting a MAF below 0.02. Once again, we implemented a threshold for SNPs showing highly significant deviations from HWE equilibrium with a p-value of 10^{-4}. For the STRUCTURE analysis, we also filtered out the SNPs under linkage disequilibrium (LD) using the function snpgdsLDpruning from the SNPrelate package (Zheng 2012). We applied the correlation coefficient method with a threshold of 0.2 (method ="corr", ld.threshold=0.2). Finally, we excluded the SNPs putatively under strong selection using BayeScan (Foll & Gaggiotti 2008). We applied BayeScan grouping by Oregon junco form (separating northern from southern *thurberi*) with default settings and a thinning interval size of 100 to ensure convergence. For each SNP we obtained the posterior probability for the selection model and the F_{ST} coefficient averaged over populations. For outlier detection and exclusion, we implemented a false discovery rate of 0.1, resulting in a final matrix of 16,858 SNPs. We converted the *vcf* file to STRUCTURE format using PGDspider (Lischer & Excoffier 2012) version 2.0.5.1. Bash scripts to perform

the analyses were created with STRAUTO (Chhatre & Emerson 2016), and we ran the program five times per K, with K ranging from 1 to 10 after running a preliminary analysis to infer the lambda value. The burn-in was set to 50K iterations and the analysis ran for an additional 100K iterations. Similarity scores among runs and graphics were computed with CLUMPAK (Kopelman *et al.* 2015).

Genotype-environment association analysis and variance partition

When applying GEA methods, there are two potentially confounding effects related to neutral factors: (i) structure among populations derived from strong drift in isolation may result in genetic patterns similar to those caused by adaptive divergence; and (ii) demographic expansions along latitudinal axes may create gradients of allele frequencies at neutral loci correlated with latitude, that in turn would correlate with any environmental variable that changes from south to north, mimicking a pattern of selective sweep and local adaptation (Excoffier *et al.* 2009; Excoffier & Ray 2008; Forester *et al.* 2016; Rellstab *et al.* 2015). Redundancy analysis (Legendre & Legendre 1998; Van Den Wollenberg 1977) is a canonical ordination method that allows computing the variance of a set of response variables explained by a number of constraining or explanatory variables. In addition, partial RDA enables computation of the shared variance between two sets of variables while conditioning or holding constant the effects of a third set of covariables. Here we used simple and partial RDA approaches as implemented in the R package vegan (Oksanen *et al.* 2016) to explore the associations between genetic variability and environmental data. RDA has been found to outperform other GEA methods, yielding a better tradeoff between false positive and true positive rates, especially under weak selective pressures acting over many loci (Forester *et al.* 2018). Nevertheless, and to further test the potential interactions between local adaptation and

the different levels of population structure, we combined RDA with an F_{ST}-based GEA method, BayeScEnv (Villemereuil & Gaggiotti 2015).

Ecological data were obtained from 7 of the 19 variables available in the BioClim database (Hijmans *et al.* 2005), specifically chosen in accordance to their relevance to junco ecology (Miller 1941; Nolan *et al.* 2002). They measured mean temperature and precipitation over the year (BIO1 and BIO12); mean temperature and precipitation over the warmest quarter (BIO10 and BIO18), which corresponds to the birds' breeding season; isothermality, referring to how the range of day-to-night temperature differs from the range of summer-to-winter, where a value of 100 indicates equality between them (BIO3); and seasonality of temperature and precipitation (BIO4 and BIO15). We also included three vegetation variables from the Moderate Resolution Imaging Spectroradiometer (MODIS) satellites as available in https://modis.gsfc.nasa.gov: percent tree cover (TREE), Normalized Difference Vegetation Index (NDVI, a measure of canopy greenness), and NDVI's annual standard deviation (std_NDVI). Finally, we included the high-quality elevation data provided by the NASA Shuttle Radar Topographic Mission (SRTM), downloadable from http://www2.jpl.nasa.gov/srtm for a total of eleven variables (Table 3). All ecological variables were centered and standardized. Following Blanchet et al. (2008), we implemented a forward selection method using the forward.sel function from the R-package packfor (Dray *et al.* 2009) to reduce the number of variables in the model. This procedure applies two stopping criteria: a significance level for each tested variable, which we set at 0.01; and a maximum limit for global adjusted R^2, equal to the adjusted R^2 of the RDA model including all initial variables. In doing so, we prevent inflation of the overall type I error and of the amount of explained variance. After this, we excluded those retained variables with a variance inflation factor (VIF) over 10 (Borcard *et al.* 2011) to avoid high collinearity. The final set of retained variables were BIO3, from now on referred to as isothermality; BIO10

and BIO18, from now on referred to as temperature and precipitation, respectively; TREE, or tree cover; NDVI, or greenness; and SRTM or elevation. Despite signs of low orthogonality observable among variables in the partial RDA (especially among precipitation, tree cover and greenness, see Results) we chose not to exclude more variables or to apply dimension-reduction treatments like PCA to the environmental space of variables so as to assess their specific and relative contributions to differentiation patterns (McCormack *et al.* 2010) and discuss signals of adaptation with higher confidence.

For building the SNP matrix to be used as response variables in the simple and partial RDA, we used VCFTOOLS to exclude indels, sex-linked and non-biallelic SNPs. Looking for a compromise between number of samples and loci, we retained up to 12 and no less than 9 samples per form (24 in the case of *thurberi*), conditioned on maximum threshold of missing data per sample of 0.6, resulting in a dataset of 91 individuals (Table 2). We filtered out all the sites with less than 4X and greater than 50X coverage, with a MAF below 0.02 or with a HWE equilibrium p-value lower than 10^{-4}. Because RDA does not allow any missing data, we also filtered out every position with any non-genotyped individual for a final subset of 15,252 SNPs. SNP data were coded as counts of the alternative allele for each position (*i.e.*, 0, 1 or 2 copies) with VCFTOOLS and transformed following Patterson et al. (2006).

Explanatory, ecological variables and the SNP matrix of response variables were used in a simple RDA with no conditional treatment; and in a partial RDA where the effects of neutral processes were subtracted. Neutral, population structure effects were approximated by the first two principal components of a new PCA performed with SNPrelate, computed from the same SNP matrix of response variables but after filtering for selection with BayeScan exactly as done for the STRUCTURE analysis. Statistical significance of RDA analyses was obtained

using a permutation-based procedure with 10,000 permutations, assuming $\alpha = 0.01$. We also used variance partitioning as implemented in the vegan R-package to estimate (i) the total proportion of genomic variation explained by ecological variables alone; (ii) by neutral structure alone; and (iii) the effects of both sets of variables. The analyses were conducted in R version 3.4.2 (code included in Appendix I, Supporting Information).

Finally, we repeated the simple RDA treatment for a subset of 49 SNPs identified as selectively divergent by a set of six BayeScEnv analyses using the same SNP matrix as in the RDA and the corresponding six ecological variables. Environmental data were computed as distances from the general mean of the means of each population, with the exception of elevation, which was measured from sea level, as recommended by BayeScEnv authors (Villemereuil & Gaggiotti 2015, https://github.com/devillemereuil/bayescenv/wiki). We used BayeScEnv with default settings and a thinning interval size of 100 to ensure convergence, and applied a FDR of 0.1 for outlier detection. Posterior probabilities of per-SNP selection-driven divergence for each one of the ecological variables were plotted with the package qqman (Turner 2016) in R 3.4.2.

Niche divergence tests

To further explore patterns of ecological divergence in the Oregon Junco, we tested for niche divergence applying the method developed by McCormack et al. (2010), which allows us to examine each environmental variable separately. To avoid a loss of statistical power due to multiple analyses, we conducted three specific comparisons of forms presenting different patterns of genetic divergence and geographical settings, including: (i) *townsendi* and southern *thurberi* forms, in order to estimate niche divergence between geographically isolated, genetically differentiated forms; (ii) the ecologically divergent *pinosus* with the

parapatric northern *thurberi* form, to further test a possible case of isolation-by-adaptation; and (iii) northern and southern populations of *thurberi*, as conspecific extremes of a potential adaptive gradient. We used occurrence points from our own georeferenced field sampling records, and this set of occurrence records was further revised to avoid spatial autocorrelation and to match the spatial resolution of environmental variables (1-km grid). Our final dataset comprised 80 localities: *pinosus* (n = 14), *thurberi* north (n = 26), *thurberi* south (n = 19), and *townsendi* (n = 21). We decided to improve quality (geographic accuracy) vs. quantity (number of occurrence records), by using fewer data with higher spatial accuracy (Engler *et al.* 2004). To generate a background dataset for each population, we drew 1000 random points from a background representing the geographic range of each junco population. In order to select an appropriately sized area for the niche divergence tests, we included accessible habitats according to the dispersal ability of each population (Soberon & Peterson 2005). We generated background samples from a 100-km "buffer zone" around known occurrences (Warren *et al.* 2008). For populations with small ranges or small dispersal ability (*thurberi* south, *pontilis* and *townsendi*), we restricted the buffer zone to 10 km to reduce spatial inaccuracies in the null distribution (Barve *et al.* 2011), after testing different buffer sizes to test the robustness in delimiting accessible areas for juncos. Next, we extracted the environmental data (identical to the data used for the RDA, see above) for both occurrence points and random background points from within the geographic range of each junco form. Niche divergence and conservatism were tested by comparing the observed environmental differences among forms against a null model of background divergence (generated by calculating the difference between background points using a bootstrapping approach and 1000 resamples) for each environmental variable using a two-tailed test. We conducted all analyses in R 3.2.2.

Results

Neutral genetic structure

The plot of the first two principal components from the PCA performed with PCANGSD revealed at least four distinct clusters in the Oregon junco group. The most differentiated groups were *townsendi* and *pontilis* from Baja California, which formed two highly divergent clusters with respect to each other and to other populations. A third differentiated group corresponded to southern *thurberi* individuals from Mount Laguna, showing less differentiation than the Baja California forms with respect to a fourth cluster, which included all the remaining forms in the PCA (Fig. 1B). Within this fourth cluster, *pinosus* from coastal California showed a certain degree of differentiation from the rest of the forms, a pattern that was clearly conspicuous when plotting the first and third components (Fig. 1C). The plot of third and fourth components also revealed a slight signal of divergence in the *oreganus* form (Fig. S1 in Supporting Information).

The STRUCTURE results were generally congruent with the PCA. The plot for K = 2 recovered *townsendi* as an independent population, which shared a considerable amount of variance with *pontilis* (Fig. 3). Although K = 2 was the best supported K value, several other values of K showed biologically relevant patterns. The analysis for K = 3 revealed *pinosus* as an independent genetic cluster, and K = 4 identified *pontilis* as differentiated population. The plot for K = 5 recovered the same main clusters seen in the PCA, capturing the differentiation of the southern *thurberi* form; and in K=6, *oreganus*, northern *thurberi*, *montanus* and *shufeldti* forms showed comparable, ambiguous probabilities of population assignment (Fig. 3).

Genotype-environment association and variance partition

Out of eleven potentially relevant ecological variables for juncos, six were retained after the forward selection method intended for excluding non-significant effects. Retained variables included isothermality, mean temperature of the warmest quarter, mean precipitation of the warmest quarter, vegetation cover, greenness, and elevation (Table 3). None of these variables were excluded due to excessive correlation as all VIF values were below 10 (maximum recovered VIF = 5.94).

All six ecological variables retained in the forward selection method were included as explanatory variables in the simple RDA and partial RDA models. RDA computes, in successive order, a series of axes that are linear combinations of the explanatory variables, and that best explain the variation in the matrix of response variables (Borcard *et al.* 2011). Six RDA axes (named RDA1 to RDA6 hereafter, ordered by the amount of variance explained by each one, Table 4 and Table S2, Supporting Information) explained 6.48% of the total genetic variance in the non-conditioned model, and 1.17% when removing the effects of neutral genetic structure (Table 4). The amount of explained variance increased to 31.07% when using only BayeScEnv outliers as response variables (Table 4, Table S2). The permutation tests for the full RDA models yielded a p-value below 0.001 in all three analyses, confirming the significance of the constraining variable effects.

Loadings of ecological explanatory variables on each of the axes varied across the three different RDA models (Fig. 4, Table 4, Table S2). In the non-conditioned RDA, the first four RDA axes presented significant linear relationship between the SNP data and the environmental variables (p-value < 0.001). The plot of the two first axes revealed that the RDA1 axis had a large negative contribution of tree cover and greenness, and loaded

positively on elevation (Fig. 4A). The RDA2 axis loaded mostly on isothermality and temperature. The plot of per-individual projections on these two axes revealed a pattern generally similar to the PCA. The forms *townsendi, pontilis* and to a lesser extent *pinosus* and southern *thurberi*, showed distinctive high values of correlation with both axes. The remaining forms showed similar correlation patterns with respect to RDA1 and 2, with *shufeldti* presenting a slightly differentiated association with RDA2 (Fig. 4A). The plot of significant RDA axes 3 and 4 revealed patterns of genetic correlation especially related to the first RDA axis for *pinosus*, which consequently presented the strongest association with isothermality. The individuals from the southern population of *thurberi* showed a more pronounced negative association with RDA2 than in the first two axes, revealing a positive correlation with mean temperature of the warmest quarter. Northern forms of the Oregon junco (northern *thurberi*, *shufeldti*, *montanus* and *oreganus*) distributed along the second RDA axis, with northernmost *montanus* and *oreganus* forms showed the strongest association with the particularly conspicuous precipitation gradient. *Townsendi* from Baja California occupied positions closer to the origin of coordinates, suggesting a weaker association between environmental and genetic variance (Fig. 4B).

The partial RDA yielded two first axes presenting significant genotype-environment associations (p-value < 0.001). The plot of RDA1 and RDA2 axes recovered association patterns highly similar to the third and fourth axes of the simple RDA, with RDA1 showing a large contribution of isothermality, while RDA2 loaded mostly on temperature and, to a lesser extent, precipitation and tree cover (Fig. 4C). Differential association patterns were more pronounced for boreal forms of Oregon junco, which differentiated along the precipitation-tree cover gradient more clearly, while *townsendi* and *pontilis* showed even lower association scores (Fig. 4C). Plots of the principal components of the PCA used as

proxies of neutral population structure were highly congruent with those based on PCANGSD (Fig. S2, Supporting Information).

The variance partition analysis showed that climate and neutral structure together explained 7.41% of the total genetic variability. Since variable sets are not orthogonal, a 5.31% of variation was explained jointly by the environmental data and the first two components of the PCA based on neutral loci. As recovered in the partial RDA, environmental variables alone explained 1.17% of the total variance, while the non-overlapping fraction of neutral genetic structure explained 0.92% of the variability in the SNP dataset. The p-value computed through the 10,000-step permutation test for each individual fraction of explained variance was below 0.001 in all cases, thus confirming the significant effects of both variable sets.

The genome-wide survey of selection using BayeScEnv identified 49 potential outliers, seven of which were associated with more than one environmental variable: isothermality (2), temperature (5), precipitation (20), greenness (3), elevation (16), and tree cover (10) (Fig. S3, Supporting Information). The RDA based on these outliers produced 4 axes capturing significant associations (RDA1, 2 and 3, p-value < 0.001; RDA4 p-value < 0.002). The first axis showed moderate positive contributions from tree cover and greenness, and also a contribution of elevation (Fig. 4D). Variance in precipitation was almost entirely captured by RDA2, which also had a relatively high, negative contribution from isothermality. The plot of RDA1 versus RDA2 showed a pattern of correlation for *pontilis* and *townsendi* along RDA1, while genetic variance in *oreganus* appeared strongly associated with the gradient of precipitation along RDA2. The rest of boreal Oregon junco forms were distributed in an opposite fashion, with small differences in their patterns of correlation with environmental variability captured in the second axis of the RDA (Fig. 4D). A second plot differentiated

pinosus from *thurberi* and remaining boreal forms correlating positively with isothermality and negatively with elevation along the RDA3 axis, while *townsendi* and *pontilis* occupied opposite positions along the RDA4 axis correlating positively with greenness and temperature, respectively (Fig. S4, Supporting Information).

Niche divergence tests

We tested for niche divergence and conservatism on each of the environmental variables. We found significant niche divergence between *pinosus* and northern *thurberi* for three of the six environmental variables analyzed (isothermality, mean precipitation of the warmest quarter and elevation; Table 5). When considering northern *thurberi* vs. southern *thurberi*, we found significant divergence for mean temperature of the warmest quarter and conservatism for isothermality (Table 5). Tree cover was the only variable that exhibited significant divergence between *townsendi* and *thurberi* south (Table 5).

Discussion

Neutral population structure and local adaptation partly explain genomic variance among Oregon junco forms

Our results reveal that both neutral and selective factors have played a role in driving divergence among Oregon junco populations, and that the relative contributions of geographic isolation and environment-driven selection are not uniform across the distribution range of the complex. Environmental variables explained 1.17% of genomic variation when controlling for population structure, and environment and neutral structure together accounted for 7.41% of the variability in the 15,252 SNP matrix. The remaining 92.59% of the variance potentially corresponds to loci under balancing selection or selective pressures not represented in our ecological variables, and to neutral variation shared by all Oregon

junco forms because of their close relatedness and/or gene flow among them. The amount of variance explained solely by environmental variables in our study was comparable to the values reported in studies applying RDA to detect specific correlations between genomic variation and a given set of potentially correlated variables, as shown in plants (e.g. De Kort *et al.* 2014; Lasky *et al.* 2012; Vincent *et al.* 2013), wolves (Forester *et al.* 2018), or other avian species (Safran *et al.* 2016; Szulkin *et al.* 2016). Previous studies on birds have used simple spatial variables such as geographic distance to control for the effects of spatial autocorrelation (Safran *et al.* 2016; Szulkin *et al.* 2016). Here, we controlled for genome-wide patterns of neutral variation by subtracting the variance captured by the first two principal components of a PCA based on neutral genome-wide SNPs, a method which should better account for population history and structure, including changes in effective population size, geographic isolation and related effects (Forester *et al.* 2016; Forester *et al.* 2018). Since spatial autocorrelation is usually reflected by neutral genetic structure, we did not include spatial covariates to avoid over-conditioning the model (Rellstab *et al.* 2015). In addition, only a small fraction of the surveyed genome is expected to be related to genes coding for climatic adaptation or linked to them (Forester *et al.* 2018; Meirmans 2015). The reported 1.17% of association between genomic variation and environmental variability is within the expected values when identifying genotype-environment association signatures (Meirmans 2015), especially when analyzing only 15,252 genome-wide SNPs, obtained from a survey of just about 1.5% of the entire genome.

Genetic-environment association patterns in the diversification of the Oregon junco

The RDA revealed a number of strikingly different patterns of covariation between genetic variance and ecological variables that may have played a role in Oregon junco diversification. These patterns differed across the RDA treatments: the first two axes of the simple RDA

recovered a signal that is strikingly similar to that obtained in the PCA, suggesting that most of the variability captured corresponds to neutral, geographically structured variability. Interestingly, the third and fourth simple RDA axes yielded very similar association patterns to those recovered in the partial RDA when controlling for population structure. This suggests that the second set of two axes of the simple RDA may reveal variation more related to adaptive variance, once prominent neutral variability has been accounted for in the first two axes. In both simple RDA axes 3 and 4, and in the partial RDA, the forms *pontilis* and *townsendi* from Baja California, markedly isolated in terms of geography and neutral genetic variability, presented a low genetic-environment association. This suggests that the differentiation between *townsendi* and *pontilis* is due largely to geographic isolation, perhaps caused by unsuitable desert habitat in the lowlands surrounding their respective mountain ranges, a pattern also known as isolation by resistance (IBR, McRae & Beier 2007). In this scenario, our results suggest that drift under conditions of small population size and reduced gene flow may have been a significant driver of differentiation, fitting the classic allopatric speciation model (Coyne & Orr 2004; Mayr 1942, 1963). This hypothesis is consistent with the niche divergence test comparing *townsendi* with southern *thurberi*, for which all tested variables but greenness showed no signal of divergence beyond expectations based on background divergence.

The form *pinosus* showed considerable neutral genetic structure and a conspicuous pattern of genetic-environment association in both non-conditioned and partial RDA. When controlling for population structure, *pinosus* individuals showed high positive correlation values with isothermality, while correlating negatively with elevation. Indeed, *pinosus* presents the highest isothermality values, and the second lowest elevation after *oreganus*, reflecting a tolerance for low elevation conditions that are absent in neighboring *thurberi*, a pattern also

recovered in the niche divergence test. Unlike other differentiated forms like *pontilis* and *townsendi*, *pinosus* does not show high geographic isolation, and zones of intergradation with *thurberi* have been described (Miller 1941). Neutral population divergence despite the absence of geographic barriers to gene flow along with signs of local adaptation is a pattern consistent with isolation-by-adaptation, where barriers to gene flow may have arisen as individuals adapted to the distinct habitat of the coastal mountains of California. Niche distinctiveness and the genetic-environment association pattern of this form are thus congruent with a combination of warm latitude, low elevation and coastal influence that have seemingly resulted in the adaptive differentiation of *pinosus* from the rest of the Oregon junco taxa. As a result, differentiation by drift may have led to positive correlations between adaptive and neutral genetic divergence (Nosil *et al.* 2008).

The southern *thurberi* individuals from Mount Laguna showed moderate differentiation in terms of neutral genetic structure with northern *thurberi* and other boreal forms, and reduced differences in their genetic-environment association patterns when no controls for confounding factors were implemented that were only observable in the less informative third and fourth RDA axes. The Mount Laguna site represents the southernmost tip of the *thurberi* range in Southern California, which extends northward and reaches Oregon, forming a relatively continuous distribution (Miller 1941; Nolan *et al.* 2002), suggesting potentially high gene flow. However, the partial RDA revealed a distinctive pattern of high correlation with the mean temperature of the warmest quarter for the southern *thurberi* juncos, differentiating them from the rest of Oregon forms. They also correlated negatively with the mean precipitation of the warmest quarter. This pattern seems congruent with the habitat of Mount Laguna, and in general with the southern inland range of Oregon juncos, quite arid during summer but subject to snowfalls in winter due to the high elevations (Miller 1941),

contrasting sharply with the more climatically moderate coastal and northern populations. The moderate neutral genetic structure but considerable differentiation in the genetic-environment association patterns is consistent with a process of local adaptation that, similarly to *pinosus*, may be promoting selection-mediated divergence and ecological isolation between *thurberi* range extremes. Consistently, the niche divergence test comparing southern and northern *thurberi* populations was significant for temperature.

The boreal Oregon junco forms including *oreganus*, *montanus*, *shufeldti*, and northern *thurberi* showed very low neutral genetic structure or differences in ecological covariances in the chief axes of the simple, non-conditioned RDA, yet showed an increasing signal of association following their latitudinal distribution in the partial RDA. A strong association pattern emerged, especially for precipitation and for correlated environmental variables of tree cover and greenness, matching quite precisely their latitudinal distribution along a gradient of increasing humidity and vegetation cover. The ecology-related differences in genetic variance, consistent with the taxonomic classification of these forms, is especially relevant considering the relative phenotypic similarity of these taxa, and their apparent intergradation (Miller 1941; Nolan *et al.* 2002). These GEA and neutral population structure features are consistent with a potential process of diversification across a selective gradient (Forester *et al.* 2016), in which selection is the prominent evolutionary force driving differentiation (Felsenstein 1976; Haldane 1948; Nagylaki 1975; Slatkin 1973) while neutral alleles may move freely across space.

GEA methods present a number of limitations, including potentially high rates of type I error (see Supporting Information for details). Here, rather than detecting specific loci under selection, we aimed to explore how selection and neutral processes shape the variability in Oregon juncos, but the risk of finding false significant associations between genetic variance and ecological parameters persists. Analyses on neutral genetic variability revealed different degrees of population structure that may have influenced the patterns obtained in the redundancy analyses, despite our best efforts to exclude demographic and historical effects. Correcting for population structure may have entailed a loss of selection signal especially for isolated forms, such as *pontilis* and *townsendi*, where adaptive variability is more likely to be also structured among populations and is thereby susceptible to being partialed out in the treatment (Forester *et al.* 2018). GEA based on constrained ordination methods that account for neutral divergence in scenarios of weak population structure may result in slight increases of false discovery rates when identifying specific loci under selection (Forester *et al.* 2018). This may be the case in the adaptive divergence signatures found in northern Oregon junco forms, yet evidence suggesting that the selection signal in our outlier SNP set is not spurious includes the overall congruence between our results and known ecological characteristics of juncos and their habitats. Nevertheless, and to further test the environmental associations revealed in this study, we included a complementary analysis combining GEA methods as suggested by some authors (Forester *et al.* 2018; Meirmans 2015; Rellstab *et al.* 2015). The simple RDA based on 49 SNP loci identified by BayeScEnv as potential targets of ecological divergent selection yielded comparatively less distinctive patterns of genotype-environment association among Oregon junco forms than the partial RDA. The RDA suggests that BayeScEnv correctly identified outliers related to low temperatures and high precipitation for *oreganus* samples, a pattern congruent with the habitat and with the outcomes of previous analyses for this form. It also detected highly differentiated positions in *pontilis* and

townsendi that correlate with RDA1, a pattern absent in the RDA based on the entire SNP dataset when correcting for population structure. This may suggest that BayeScEnv failed to exclude the effects of demographic history, or in turn, that controlling for the genetic variance captured by the PCA was indeed overly conservative in the RDA approach, as explained above. BayeScEnv was also less successful in detecting adaptive divergence in *pinosus* and in northern Oregon junco forms (with the exception of *oreganus*), where selection may be weaker or have acted during a shorter period. Arguably, the partial congruence between the full SNP dataset RDAs and the RDA based on BayeScEnv outliers entails a need for caution when interpreting the genotype-environment covariance patterns uncovered in this study. Nevertheless, the outlier SNP dataset explained 31.07% of the total climatic variability, a considerable amount compared with the full SNP data RDA models. This indicates a good fit of the retained outliers to the linear regression on the environmental parameters.

Interactions among environment, geography and demographic processes suggest three different modes of divergence within the Oregon junco lineage

The Oregon junco is one of the six phenotypically and genetically differentiated dark-eyed junco taxa that evolved during a northward expansion from Central America after the last glacial maximum, c.a. 18,000 years ago (Friis *et al.* 2016; Milá *et al.* 2007). Similar postglacial expansions have been reported for many other bird species (Alvarez *et al.* 2015; Hansson *et al.* 2008; Malpica & Ornelas 2014; Milá *et al.* 2000; 2006; Seutin *et al.* 1995). However, the population structure documented in this study reveals a variety of different spatial, selective and historical factors not previously documented in other avian taxa. In light of our genetic-environment association analyses, and of the patterns recovered by the PCA and STRUCTURE analyses, the results of this study suggest at least three different effects of geography and demographic history interacting to varying degrees with selection in the

process of Oregon junco diversification. First, the IBR pattern of differentiation presented by *pontilis* and *townsendi* may suggest that these forms are peripheral remnants of an original, broader distribution of the Oregon juncos, thereafter isolated in 'sky islands' of Baja California and diverging predominantly by drift. Indeed, in his thorough monograph on the geographic variation in juncos, Alden Miller (1941) had perceptively suggested early on that the habitat of Oregon juncos from Baja California did not seem to account for their phenotypic differentiation from Californian forms, and that their distinctive traits appeared to be predominantly "historical" (p. 306). The spatial configuration and recovered patterns of neutral and adaptive divergence for *pontilis* and *townsendi* fit the first model of neutral divergence of isolated population in approximately similar habitats (scenario A in Fig. 2). Alternatively, selection related to environmental variables such as tree cover, elevation and temperature might have played a role in the differentiation of a reduced set of outliers as depicted in the BayeScEnv-RDA combined approach, a signal that may have been lost when controlling for neutral population structure in the partial RDA.

Second, the potential IBA pattern found in *pinosus* suggests that the current area of intergradation may be the result of the formation of a secondary contact zone that formed after diverging in relative isolation, possibly linked to the ancient coastal closed-cone pine forest that has allegedly diminished since the Pleistocene (Miller 1941). Alternatively, *pinosus* could have differentiated in the absence of any geographic barrier to gene flow, and by the sole action of ecology-driven divergent selection. BayeScEnv's failure to identify differences in adaptive variability with respect to other boreal Oregon junco forms seems, in this case, a problem of analytical power, in light of the neutral structure and the rest of GEA analyses. In either case, the mode of divergence between parapatric populations in different habitats (scenario B in Fig. 2) has been only partially fulfilled, since local adaptation seems to

have resulted in reduced levels of gene flow, leading to increased neutral genome-wide differentiation (Flaxman *et al.* 2014; Funk *et al.* 2011; Nosil *et al.* 2008).

Third, the geographic continuum represented by northern *thurberi*, *shufeldti*, *montanus* and *oreganus* is also captured in the STRUCTURE analysis, suggesting that ongoing gene flow may occur among forms. Combined with the latitudinal pattern of increasing environmental association recovered in the partial RDA, these outcomes are consistent with a process of differentiation driven by local adaptation along a selective gradient in the direction of the northward expansion, fulfilling the third hypothesized mode of divergence for these forms of Oregon junco (scenario C in Fig. 2). BayeScEnv analysis uncovered a non-continuous pattern of differentiation among these northern forms, which may respond to a highest intensity of precipitation and vegetation related pressures at the extreme of the distribution, occupied by *oreganus*. Interestingly, ecological association approaches have been shown to perform better along clines of selection where demographic expansions align with the gradient of ecological variables, usually related to latitude (Frichot *et al.* 2015), as is the case in the *Junco* system.

A relevant aspect of the marked population structure found among Oregon junco forms is that it is based on a relatively small subset of genome-wide SNPs randomly sampled from across the genome, representing a genomic fraction not greater than 1.5% of the total of 1.1 Gb. The apparent signal of divergence mediated by environmental factors recovered in the RDA indicates that divergence may have taken place at the level of the entire genome, suggesting the role of multiple selective pressures during local adaptation along the latitudinally broad and heterogeneous distribution of the Oregon juncos. The presence of outliers potentially under positive selection scattered across the genome, as evidenced by BayeScEnv scans, seems to support this hypothesis of selection-driven genome-wide divergence, rather than

widespread drift among isolated populations. Other examples of such patterns of genomic differentiation due to divergent selection at early (e.g. Brawand *et al.* 2014; Egan *et al.* 2015; Parchman *et al.* 2013) and intermediate (e.g. Riesch *et al.* 2017) stages of speciation have been reported recently, contrasting with proposed models of speciation initiated by divergent selection in a few, localized genes involved in reproductive isolation (e.g. Nadeau *et al.* 2012; Poelstra *et al.* 2014).

Conclusion

Our analyses reveal the role of both local adaptation and demography in driving rapid diversification during the northward recolonization of western North America by the Oregon junco. The combined effects of a demographic expansion along a selective gradient with a heterogeneous landscape of environmental variability have resulted in a striking array of divergence modes within a single lineage, from isolated forms in Baja California that have differentiated largely by drift in isolated 'sky islands', to adaptive diversification along selective gradients with no obvious geographic barriers to gene flow. There is also a compelling example of isolation by adaptation in the case of *pinosus*, where ecological barriers to gene flow seem to maintain its divergence with respect to nearby forms. Genome-wide patterns of divergence indicate that Oregon junco diversification has been driven by multiple ecological factors acting on many loci across the genome, and suggests that selection may promote local adaptation in short periods of time, highlighting the role of adaptive divergence in the early stages of the speciation process. Future analyses of dense sequencing and functional gene characterization will be necessary to further identify adaptive changes promoting barriers to gene flow and reveal the genomic architecture of rapid diversification.

Acknowledgements

We thank Pau Aleixandre, Jonathan Atwell, Elena Berg, Setefilla Buenavista, Steve Burns, Jatziri Calderón, Adrián Gutiérrez, Alfonsina Hernández, Fritz Hertel, Ellen Ketterson, Adán Oliveras de Ita, César Ríos, Sahid Robles, Vicente Rodríguez, Whitney Tsai, Rich Van Buskirk and Alvar Veiga for invaluable help in the field and lab. We are also grateful to Yoann Anselmetti, Clémentine Francois, Rachel Johnston, Etienne Kornobis, Benoit Nabholz, Sergio Nigenda, Jacqueline Robinson, Marjolaine Rouselle and Robert K. Wayne for their assistance with bioinformatic analyses. We also thank Laura Barrios and Luis M. Carrascal for reviewing the statistical analyses and providing orientation, and to Marta Szulkin for her kind assistance with GEA analyses. We thank the subject editor and three anonymous reviewers for constructive comments on a previous version of the manuscript. Funding was provided by grant CGL-2011-25866 from the Spanish Ministerio de Ciencia e Innovación to BM.

Data Accessibility

SNP and georreferenced datasets deposited in Dryad (doi:10.5061/dryad.0nb2307); *Junco hyemalis* reference genome deposited in NCBI (Accession number: QZWM00000000).

Author Contributions

G. Friis and BM designed the study and carried out field sampling; G. Friis, JM, and BF generated and analyzed genomic data; G. Friis, G. Fandos and AZ generated and analyzed environmental data; G. Friis and BM wrote the manuscript with input from all co-authors.

References

Ahrens CW, Rymer PD, Stow A, *et al.* (2018) The search for loci under selection: trends, biases and progress. *Molecular Ecology* **27**, 1342-1356.

Alvarez S, Salter JF, McCormack JE, Milá B (2015) Speciation in mountain refugia: phylogeography and demographic history of the pine siskin and black-capped siskin complex. *Journal of Avian Biology* **47**, 335-345.

Auwera GA, Carneiro MO, Hartl C, *et al.* (2013) From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. *Current protocols in bioinformatics*, 11.10. 11-11.10. 33.

Barve N, Barve V, Jiménez-Valverde A, *et al.* (2011) The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling* **222**, 1810-1819.

Benestan L, Quinn BK, Maaroufi H, *et al.* (2016) Seascape genomics provides evidence for thermal adaptation and current-mediated population structure in American lobster (Homarus americanus). *Molecular Ecology* **25**, 5073-5092.

Bierne N, Welch J, Loire E, Bonhomme F, David P (2011) The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Molecular Ecology* **20**, 2044-2072.

Billiard S, Lenormand T, Gavrilets S (2005) Evolution of migration under kin selection and local adaptation. *Evolution* **59**, 13-23.

Blanchet FG, Legendre P, Borcard D (2008) Forward selection of explanatory variables. *Ecology* **89**, 2623-2632.

Borcard D, Gillet F, Legendre P (2011) Chapter 6. Canonical Ordination. In: *Numerical Ecology with R*, pp. 153-226. Springer.

Brawand D, Wagner CE, Li YI, *et al.* (2014) The genomic substrate for adaptive radiation in African cichlid fish. *Nature* **513**, 375.

Campagna L, Repenning M, Silveira LF, *et al.* (2017) Repeated divergent selection on pigmentation genes in a rapid finch radiation. *Science advances* **3**, e1602404.

Carson HL (1975) The genetics of speciation at the diploid level. *The American Naturalist* **109**, 83-92.

Chapman JA, Ho I, Sunkara S, *et al.* (2011) Meraculous: de novo genome assembly with short paired-end reads. *PLOS One* **6**, e23501.

Chhatre V, Emerson K (2016) StrAuto: Automation and parallelization of STRUCTURE analysis. *See http://strauto. popgen. org*.

Christmas MJ, Biffin E, Breed MF, Lowe AJ (2016) Finding needles in a genomic haystack: targeted capture identifies clear signatures of selection in a nonmodel plant species. *Molecular Ecology* **25**, 4216-4233.

Coyne JA, Orr HA (2004) *Speciation* Sinauer Associates, Inc., Sunderland, Massachusetts.

Cruickshank TE, Hahn MW (2014) Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology* **23**, 3133-3157.

Danecek P, Auton A, Abecasis G, *et al.* (2011) The variant call format and VCFtools. *Bioinformatics* **27**, 2156-2158.

Darwin C (1859) *The origin of species by means of natural selection, or the preservation of favoured races in the struggle for life.* John Murray, London.

De Kort H, Vandepitte K, Bruun HH, *et al.* (2014) Landscape genomics and a common garden trial reveal adaptive differentiation to temperature across Europe in the tree species Alnus glutinosa. *Molecular Ecology* **23**, 4709-4721.

DePristo MA, Banks E, Poplin R, *et al.* (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics* **43**, 491-498.

Dray S, Legendre P, Blanchet G (2009) packfor: Forward Selection with permutation. *R package version 0.0-7/r58*.

Dwight J (1918) The geographical distribution of color and of other variable characters in the genus *Junco*: a new aspect of specific and subspecific values. *Bull. Am Mus. Nat. Hist.* **38**, 269-309.

Edmonds CA, Lillie AS, Cavalli-Sforza LL (2004) Mutations arising in the wave front of an expanding population. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 975-979.

Egan SP, Ragland GJ, Assour L, *et al.* (2015) Experimental evidence of genome-wide impact of ecological selection during early stages of speciation-with-gene-flow. *Ecology Letters* **18**, 817-825.

Elshire RJ, Glaubitz JC, Sun Q, *et al.* (2011) A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLOS One* **6**, e19379.

Engler R, Guisan A, Rechsteiner L (2004) An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal of applied ecology* **41**, 263-274.

Excoffier L, Hofer T, Foll M (2009) Detecting loci under selection in a hierarchically structured population. *Heredity* **103**, 285-298.

Excoffier L, Ray N (2008) Surfing during population expansions promotes genetic revolutions and structuration. *Trends in Ecology & Evolution* **23**, 347-351.

Faria R, Renaut S, Galindo J, *et al.* (2014) Advances in Ecological Speciation: an integrative approach. *Molecular Ecology* **23**, 513-521.

Felsenstein J (1976) The theoretical population genetics of variable selection and migration. *Annual Review of Genetics* **10**, 253-280.

Flaxman SM, Wacholder AC, Feder JL, Nosil P (2014) Theoretical models of the influence of genomic architecture on the dynamics of speciation. *Molecular Ecology* **23**, 4074-4088.

Foll M, Gaggiotti OE (2008) A genome scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics* **180**, 977-993.

Forester BR, Jones MR, Joost S, Landguth EL, Lasky JR (2016) Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes. *Molecular Ecology* **25**, 104-120.

Forester BR, Lasky JR, Wagner HH, Urban DL (2018) Comparing methods for detecting multilocus adaptation with multivariate genotype-environment associations. *Molecular Ecology*.

Frichot E, Schoville SD, de Villemereuil P, Gaggiotti OE, François O (2015) Detecting adaptive evolution based on association with ecological gradients: orientation matters! *Heredity* **115**, 22-28.

Friis G, Aleixandre P, Rodríguez-Estrella R, Navarro-Sigüenza AG, Milá B (2016) Rapid postglacial diversification and long-term stasis within the songbird genus Junco: phylogeographic and phylogenomic evidence. *Molecular Ecology* **25**, 6175-6195.

Funk DJ, Egan SP, Nosil P (2011) Isolation by adaptation in Neochlamisus leaf beetles: host-related selection promotes neutral genomic divergence. *Molecular Ecology* **20**, 4671-4682.

Grabherr MG, Russell P, Meyer M, *et al.* (2010) Genome-wide synteny through highly sensitive sequence alignment: Satsuma. *Bioinformatics* **26**, 1145-1151.

Haldane JB (1948) The theory of a cline. *Journal of genetics* **48**, 277-284.

Hansson B, Hasselquist D, Tarka M, Zehtindjiev P, Bensch S (2008) Postglacial colonisation patterns and the role of isolation and expansion in driving diversification in a passerine bird. *PLOS One* **3**, e2794.

Hedrick PW, Ginevan ME, Ewing EP (1976) Genetic polymorphism in heterogeneous environments. *Annual Review of Ecology and Systematics* **7**, 1-32.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**, 1965-1978.

Jones MR, Forester BR, Teufel AI, *et al.* (2013) Integrating landscape genomics and spatially explicit approaches to detect loci under selection in clinal populations. *Evolution* **67**, 3455-3468.

Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. *Ecology Letters* **7**, 1225-1241.

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I (2015) Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources* **15**, 1179-1191.

Korneliussen TS, Albrechtsen A, Nielsen R (2014) ANGSD: analysis of next generation sequencing data. *BMC bioinformatics* **15**, 356.

Krueger F (2015) Trim Galore!: A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files.

Lasky JR, Des Marais DL, McKAY J, *et al.* (2012) Characterizing genomic variation of Arabidopsis thaliana: the roles of geography and climate. *Molecular Ecology* **21**, 5512-5529.

Legendre P, Legendre L (1998) Numerical ecology: second English edition. *Developments in environmental modelling* **20**.

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754-1760.

Lischer HEL, Excoffier L (2012) PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* **28**, 298-299.

Malpica A, Ornelas JF (2014) Postglacial northward expansion and genetic differentiation between migratory and sedentary populations of the broad-tailed hummingbird (*Selasphorus platycercus*). *Molecular Ecology* **23**, 435-452.

Manthey JD, Moyle RG (2015) Isolation by environment in white-breasted nuthatches (*Sitta carolinensis*) of the Madrean Archipelago sky islands: a landscape genomics approach. *Molecular Ecology* **24**, 3628-3638.

Mayr E (1942) *Systematics and the origin of species* Columbia Univ. Press, New York.

Mayr E (1947) Ecological factors in speciation. *Evolution*, 263-288.

Mayr E (1954) Change of genetic environment and evolution. **157**.

Mayr E (1963) *Animal species and evolution* Belknap Press, Cambridge, MA.

McCormack JE, Hird SM, Zellmer AJ, Carstens BC, Brumfield RT (2013) Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular Phylogenetics and Evolution* **66**, 526-538.

McCormack JE, Zellmer AJ, Knowles LL (2010) Does Niche Divergence Accompany Allopatric Divergence in Aphelocoma Jays as Predicted Under Ecological Speciation?: Insights from Tests with Niche Models. *Evolution* **64**, 1231-1244.

McKenna A, Hanna M, Banks E, *et al.* (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* **20**, 1297-1303.

McRae BH, Beier P (2007) Circuit theory predicts gene flow in plant and animal populations. *Proceedings of the National Academy of Sciences* **104**, 19885-19890.

Meirmans PG (2015) Seven common mistakes in population genetics and how to avoid them. *Molecular Ecology* **24**, 3223-3231.

Meisner J, Albrechtsen A (2018) Inferring Population Structure and Admixture Proportions in Low Depth NGS Data. *bioRxiv*, 302463.

Milá B, Aleixandre P, Alvarez-Nordström S, *et al.* (2016) More than meets the eye: Lineage diversity and evolutionary history of Dark-eyed and Yellow-eyed juncos. *Snowbird: Integrative Biology and Evolutionary Diversity in the Junco (ED Ketterson and JW Atwell, Editors). University of Chicago Press, Chicago, Illinois, USA*, 179-198.

Milá B, Girman DJ, Kimura M, Smith TB (2000) Genetic evidence for the effect of a postglacial population expansion on the phylogeography of a North American songbird. *Proceedings of the Royal Society Biological Sciences Series B* **267**, 1033-1040.

Milá B, McCormack JE, Castaneda G, Wayne RK, Smith TB (2007) Recent postglacial range expansion drives the rapid diversification of a songbird lineage in the genus Junco. *Proceedings of the Royal Society B-Biological Sciences* **274**, 2653-2660.

Milá B, Smith TB, Wayne RK (2006) Postglacial population expansion drives the evolution of long-distance migration in a songbird. *Evolution* **60**, 2403-2409.

Miller A (1941) Speciation in the avian genus *Junco*. *University of California Publications in Zoology* **44**, 173-434.

Mitton J, Linhart Y, Hamrick J, Beckman J (1977) Observations on the genetic structure and mating system of ponderosa pine in the Colorado Front Range. *Theoretical and Applied Genetics* **51**, 5-13.

Murray KD, Borevitz JO (2017) Axe: rapid, competitive sequence read demultiplexing using a trie. *bioRxiv*, 160606.

Nadeau NJ, Whibley A, Jones RT, *et al.* (2012) Genomic islands of divergence in hybridizing Heliconius butterflies identified by large-scale targeted sequencing. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**, 343-353.

Nadeau S, Meirmans PG, Aitken SN, Ritland K, Isabel N (2016) The challenge of separating signatures of local adaptation from those of isolation by distance and colonization history: The case of two white pines. *Ecology and Evolution* **6**, 8649-8664.

Nagylaki T (1975) Conditions for the existence of clines. *Genetics* **80**, 595-615.

Nolan VJ, Ketterson ED, Cristol DA, *et al.* (2002) Dark-eyed Junco (*Junco hyemalis*). In: *The Birds of North America* (eds. Poole A, Gill F). The Birds of North America, Inc., Philadelphia, Pennsylvania.

Nosil P (2012) *Ecological Speciation* Oxford University Press.

Nosil P, Egan SP, Funk DJ (2008) Heterogeneous genomic differentiation between walking-stick ecotypes:"isolation by adaptation" and multiple roles for divergent selection. *Evolution* **62**, 316-336.

Ojeda Alayon DI, Tsui CK, Feau N, *et al.* (2017) Genetic and genomic evidence of niche partitioning and adaptive radiation in mountain pine beetle fungal symbionts. *Molecular Ecology* **26**, 2077-2091.

Oksanen J, Blanchet F, Kindt R, Legendre P, O'Hara R (2016) Vegan: community ecology package. R Packag. 2.3-5.

Parchman T, Gompert Z, Braun M, *et al.* (2013) The genomic consequences of adaptive divergence and reproductive isolation between species of manakins. *Molecular Ecology* **22**, 3304-3317.

Patterson N, Price AL, Reich D (2006) Population structure and eigenanalysis. *PLoS genet* **2**, e190.

Poelstra JW, Vijay N, Bossu CM, *et al.* (2014) The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science* **344**, 1410-1414.

Pritchard JK, Di Rienzo A (2010) Adaptation–not by sweeps alone. *Nature Reviews Genetics* **11**, 665-667.

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

Putnam NH, O'Connell BL, Stites JC, *et al.* (2016) Chromosome-scale shotgun assembly using an in vitro method for long-range linkage. *Genome Research* **26**, 342-350.

R_Core_Team (2013) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rellstab C, Gugerli F, Eckert AJ, Hancock AM, Holderegger R (2015) A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology* **24**, 4348-4370.

Riesch R, Muschick M, Lindtke D, *et al.* (2017) Transitions between phases of genomic differentiation during stick-insect speciation. *Nature Ecology & Evolution* **1**, 0082.

Rundle H, Nosil P (2005) Ecological speciation. *Ecology Letters* **8**, 336-352.

Safran R, Scordato E, Wilkins M, *et al.* (2016) Genome-wide differentiation in closely related populations: the roles of selection and geographic isolation. *Molecular Ecology* **25**, 3865-3883.

Schluter D (2000) *The ecology of adaptive radiation* Oxford University Press, Oxford.

Seehausen O, Butlin RK, Keller I, *et al.* (2014) Genomics and the origin of species. *Nature Reviews Genetics* **15**, 176-192.

Seutin G (1991) Preservation of avian blood and tissue samples for DNA analyses. *Canadian Journal of Zoology* **69**, 82-90.

Seutin G, Ratcliffe LM, Boag PT (1995) Mitochondrial DNA homogeneity in the phenotypically diverse redpoll finch complex (Aves: Carduelinae: *Carduelis flammea-hornemanni*). *Evolution* **49**, 962-973.

Shafer ABA, Wolf JBW (2013) Widespread evidence for incipient ecological speciation: a meta-analysis of isolation-by-ecology. *Ecology Letters* **16**, 940-950.

Slatkin M (1973) Gene flow and selection in a cline. *Genetics* **75**, 733-756.

Soberon J, Peterson AT (2005) Interpretation of models of fundamental ecological niches and species' distributional areas.

Sork VL, Squire K, Gugger PF, *et al.* (2016) Landscape genomic analysis of candidate genes for climate adaptation in a California endemic oak, Quercus lobata. *American Journal of Botany* **103**, 33-46.

Szulkin M, Gagnaire PA, Bierne N, Charmantier A (2016) Population genomic footprints of fine-scale differentiation between habitats in Mediterranean blue tits. *Molecular Ecology* **25**, 542-558.

Templeton AR (1981) Mechanisms of speciation-a population genetic approach. *Annual Review of Ecology and Systematics* **12**, 23-48.

Termignoni-García F, Jaramillo-Correa JP, Chablé-Santos J, *et al.* (2017) Genomic footprints of adaptation in a cooperatively breeding tropical bird across a vegetation gradient. *Molecular Ecology*.

Turner SD (2016) qqman: an R package for visualizing GWAS results using QQ and manhattan plots. bioRxiv. 2014. *DOI* **10**, 005165.

Uyeda JC, Arnold SJ, Hohenlohe PA, Mead LS (2009) Drift promotes speciation by sexual selection. *Evolution* **63**, 583-594.

Van Den Wollenberg AL (1977) Redundancy analysis an alternative for canonical correlation analysis. *Psychometrika* **42**, 207-219.

Villemereuil P, Gaggiotti OE (2015) A new FST-based method to uncover local adaptation using environmental variables. *Methods in Ecology and Evolution* **6**, 1248-1258.

Vincent B, Dionne M, Kent MP, Lien S, Bernatchez L (2013) Landscape genomics in Atlantic salmon (Salmo salar): searching for gene–environment interactions driving local adaptation. *Evolution* **67**, 3469-3487.

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

Warren DL, Glor RE, Turelli M (2008) Environmental Niche Equivalency Versus Conservatism: Quantitative Approaches to Niche Evolution. *Evolution* **62**, 2868-2883.

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

Wright S (1946) Isolation by distance under diverse systems of mating. *Genetics* **31**, 39.

Yates A, Akanni W, Amode MR, *et al.* (2016) Ensembl 2016. *Nucleic Acids Research*, gkv1157.

Zheng X (2012) SNPRelate: parrallel computing toolset for genome-wide association studies. *R package version* **95**.

Table 1. Oregon junco forms and number of genotyped individuals per locality. State abbreviations are the following: British Columbia (BC) in Canada; Oregon (OR) and California (CA) in the USA; and Baja California Norte (BCN) in Mexico. For additional locality information see Table S1 in the Supplementary Information.

Form	*State*	*Localities*	*Sequenced*
oreganus	BC	Banks Is., Porcher Is., Susan Is.	16
shufeldti	OR	Willamette N.F.	12
montanus	OR	Wallowa N.F.	15
northern *thurberi*	CA	Tahoe	18
southern *thurberi*	CA	Mount Laguna	17
pinosus	CA	Santa Cruz Mountains	16
pontilis	BCN	Sierra Juárez	16
townsendi	BCN	Sierra San Pedro Mártir	16
Total			126

Table 2. SNP data matrices used in each analysis. General filters included a coverage depth range from 4X to 50X, a genotyping phred quality score equal to 40, a MAF equal to 0.02 and a p-value for Hardy-Weinberg deviation of 10^{-4}. Coverage values correspond to the sequencing depth averaged over loci and individuals.

Analysis	**Samples**	**Missing data**	**LD corr.**	**BayeScan FDR**	**Coverage**	**SNPs**
PCANGSD	64	50%	-	-	6.79	255,163
STRUCTURE	64	25% (per pop.)	0.2	0.1	13.98	16,858
RDA:						
All loci	91	0%	-	-	16.59	15,252
BayeScEnv outliers	91	0%	-	-	15.59	49
PCA	91	0%	-	0.1	16.59	15,209
BayeScEnv scans	91	0%	-	-	16.59	15,252

Table 3. Set of environmental variables included in the initial step-forward selection method. Significant variables retained by the method are shown in bold.

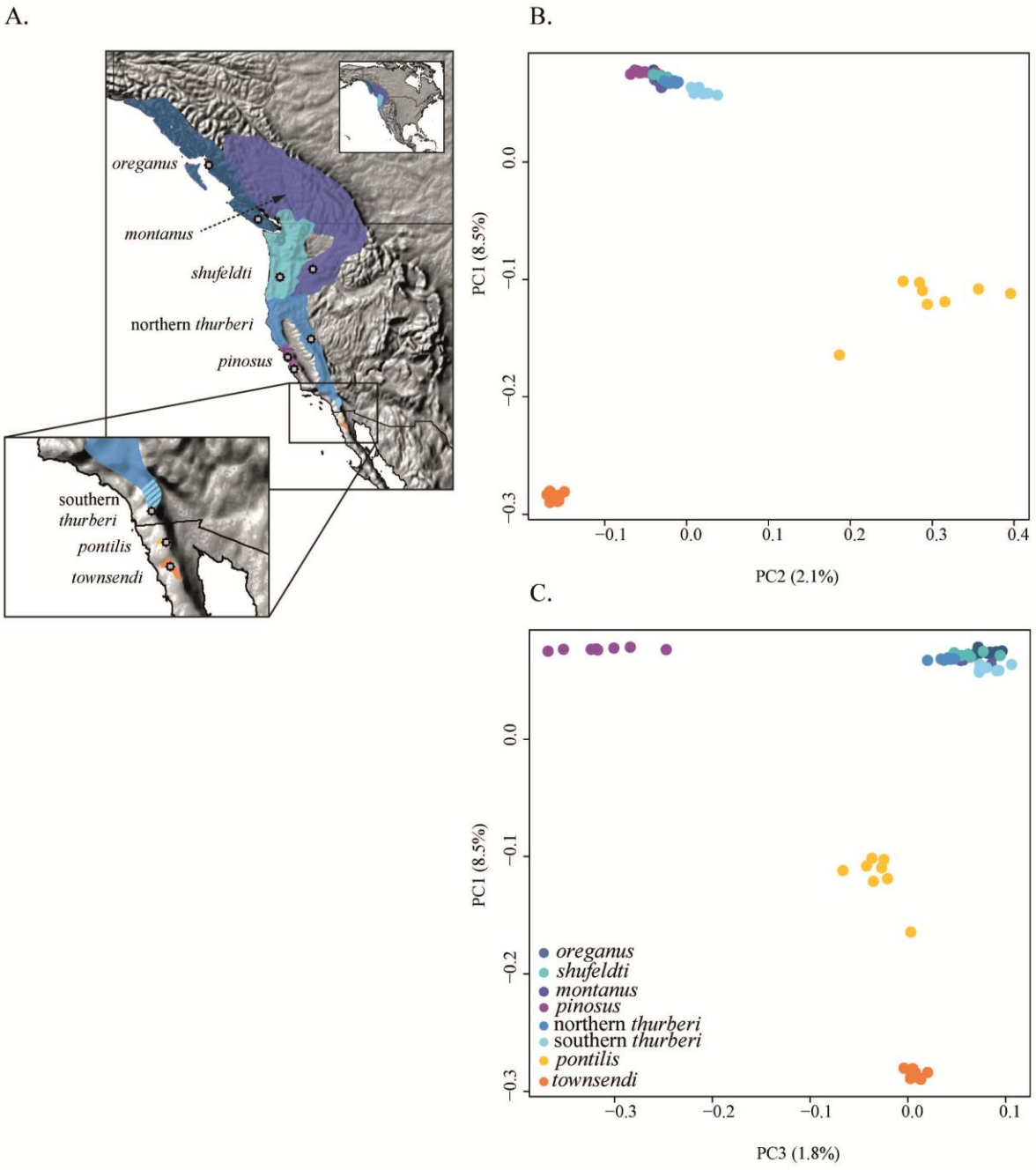
Variable	Description
BIO1	Annual mean temperature
BIO3	Isothermality
BIO4	Temperature seasonality (standard deviation *100)
BIO10	Mean temperature of the warmest quarter
BIO12	Annual precipitation
BIO15	Precipitation seasonality (coefficient of variation)
BIO18	Precipitation of the warmest quarter
NDVI	Normalized Difference Vegetation Index (greenness)
NDVI SD	Annual NDVI standard deviation (greenness seasonality)
TREE	Percent tree cover
SRTM	Elevation from the NASA Shuttle Radar Topographic Mission

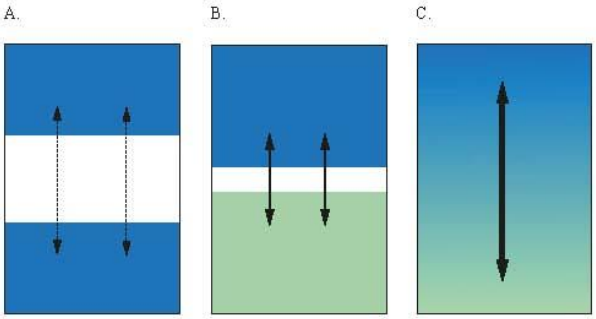
Table 4. RDA loads of the constraining variables in the first two axes for each of the RDA models. The total explained variance (adjusted R^2) is shown for each analysis. In all of the three analyses, the p-value for the full models was below 0.001.

	Simple RDA		**Partial RDA**		**BayeScEnv outliers**	
Total explained variance (Adjusted R2)	6.48%		1.17%		31.07%	
Variables	RDA1	RDA2	RDA1	RDA2	RDA1	RDA2
Isothermality	0.147	0.196	0.813	-0.438	-0.197	-0.706
Mean Temperature of Warmest Quarter	0.126	-0.624	0.329	-0.682	-0.295	-0.309
Mean Precipitation of Warmest Quarter	0.012	0.033	-0.416	0.481	0.070	0.981
Elevation	0.655	-0.002	-0.311	-0.291	-0.619	-0.543
Greenness	-0.656	0.581	-0.295	0.165	0.829	0.252
Tree cover	-0.720	0.310	-0.264	0.287	0.799	0.364

Table 5. Results from the niche divergence test for *pinosus* vs. northern *thurberi*, northern vs. southern *thurberi* and southern *thurberi* vs. *townsendi*. Variables showing significant divergence (Diverged) or conservatism (Conserved) are shown in bold (p-value < 0.05). e.d.b.b.: expected divergence based on background. See Table 3 for variable definitions.

	Variable	Result	p-value	Observed mean difference	Background mean difference
pinosus vs.	TREE	e.d.b.b.	0.636	3.95	7.17
northern *thurberi*	BIO10	e.d.b.b.	0.718	20.40	17.34
	BIO3	**Diverged**	**0.000**	**16.78**	**9.45**
	BIO18	**Diverged**	**0.000**	**41.92**	**27.24**
	SRTM	**Diverged**	**0.032**	**1329.81**	**978.26**
	NDVI	e.d.b.b.	0.550	192.67	428.18
northern vs.	TREE	e.d.b.b.	0.580	2.11	4.52
southern *thurberi*	**BIO10**	**Diverged**	**0.030**	**41.85**	**23.87**
	BIO3	**Conserved**	**0.002**	**1.37**	**2.87**
	BIO18	e.d.b.b.	0.860	15.02	14.25
	SRTM	e.d.b.b.	0.474	202.60	140.29
	NDVI	e.d.b.b.	0.144	258.89	1019.97
southern *thurberi*	TREE	e.d.b.b.	0.536	15.01	12.25
vs. t*ownsendi*	BIO10	e.d.b.b.	0.412	47.88	43.03
	BIO3	e.d.b.b.	0.472	2.47	2.89
	BIO18	e.d.b.b.	0.188	74.77	67.26
	SRTM	e.d.b.b.	0.786	750.22	717.21
	NDVI	**Diverged**	**0.012**	**1597.62**	**574.28**





	A.	B.	C.
Habitat types:	ONE	TWO	GRADIENT
Gene flow:	LOW	MODERATE	HIGH
Neutral divergence:	HIGH	MODERATE	LOW
Adaptive divergence:	LOW	HIGH	HIGH