# Environmental drivers behind the genetic differentiation in mountain chickadees (*Poecile gambeli)*

Running title: Population genetics of mountain chickadees

Srikanthan P[1], TM Burg

Department of Biology, 4401 University Dr, University of Lethbridge, Lethbridge, AB, T1K 3M4, Canada

**1 Corresponding author**

Email: Srikanthan P. prahaladsrikanthan@gmail.com

Current address of corresponding author: Department of Biological Engineering, Indian Institute of Technology Madras, 600036 India

[1] Current address of corresponding author: Department of Biological Engineering, Indian Institute of Technology Madras, 600036 India

**Abstract**

Anthropogenic climate change has a large impact on wildlife populations and the scale of the impacts have been increasing. In this study, we utilised ddRAD sequence data to investigate genetic divergence and identify the environmental drivers of genetic differentiation between 12 populations of mountain chickadees, family Paridae, sampled across North America. To delineate populations and identify potential zones of hybridisation, we conducted a discriminant analysis of principal components (DAPC), admixture analysis, and calculated pairwise Fst values. The DAPC revealed four clusters: southern California, eastern Rocky Mountains, northwestern Rocky Mountains and Oregon/northern California. We then used BayeScEnv to highlight significant outlier SNPs associated with the five environmental variables. We identified over 150 genes linked to outlier SNPs associated with more than 15 pathways, including stress response and circadian rhythm. We also found a strong signal of isolation by distance. Local temperature was highly correlated with genetic distance. Maxent simulations showed a northward range shift over the next 50 years and a decrease in suitable habitat, highlighting the need for immediate conservation action.

**Keywords:** climate change; ddRAD; gene-environment interaction; *Poecile gambeli*; population genetics; species distribution models

## Introduction

Anthropogenic climate change is increasingly affecting ecosystems worldwide. The steady increase in temperature means and variability can affect the functioning of organisms and alter ecosystems that have existed on this planet for millennia. While most ecosystems are sensitive to changes in the environment, mountain ecosystems are particularly vulnerable to environmental changes (Beniston, 2003). Owing to increasing temperatures and reduced snow cover, organisms in mountainous regions may shift their elevational range to higher altitudes, potentially leading to population fragmentation and extinction (Calkins et al., 2012; McDonald & Brown, 1992; Parmesan, 2006; Wilson et al., 2007). Climate change also significantly affects phenology, such as the timing of flowering or breeding (Walther et al., 2002). The extent and effects of these changes are a current area of study, and the pace at which changes occur requires commensurate efforts in the form of conservation research to prevent the breakdown of ecosystems and extinction of organisms (Christmann & Menor, 2021; Payne et al., 2020).

The North American landscape is full of physical barriers that affect species dispersal, and therefore causing genetic divergence (Antonelli, 2017; Boutilier et al., 2014; Machado et al., 2018). In addition, the wide range of environmental conditions across the continent makes it an ideal area to study the effect of the environment on genetic differentiation. In particular, species with low migration rates could experience significant divergence (Keyghobadi et al., 1999). Therefore, it is crucial to study the population genetics of these species to investigate their resilience to climate change (Hanski et al., 2006). Despite being able to traverse long distances, birds are unable to pass through certain barriers, makes them a good model for studying the effects of physical barriers on divergence (Greenwood & Harvey, 1982).

The mountain chickadee, *Poecile gambeli*, is found across western North America and primarily occupies coniferous forests. As many as seven subspecies have been described and are divided into two mtDNA groups, eastern and western, with limited contemporary gene flow observed between them (Hindley et al., 2018; Spellman et al., 2007). A recent study using microsatellite data identified Washington as a potential contact zone between two clades (Hindley et al., 2018). Ubiquitous across the west, the mountain chickadee has limited migratory capabilities, and high philopatry makes it an ideal candidate for studying the effects of geography and climate on genetic differentiation.

In this study, we examined the population structure of mountain chickadees from populations across the range using double digest restriction-site associated DNA, ddRAD, data. In addition, we aimed to answer the following questions: (1) What are the environmental drivers of genetic differentiation? Previous studies using mtDNA have shown that isolation by distance influences differentiation (Hindley et al., 2018); however, more information is needed on the effects of other variables such as temperature and precipitation. (2) Which genes or pathways are undergoing selection across populations? Given the pace of climate change, it is essential to investigate the pathways undergoing selection to identify gene-environment interactions. (3) How much will the changing climate affect the habitat of mountain chickadees over the next 50 years? Although we expect a northward shift in habitat or a shift in elevational range, understanding how a common species is affected by climate change can provide valuable insights to inform conservation decisions on other co-distributed species.

## Methods

*Sampling and DNA extraction*

A total of 94 mountain chickadee samples were collected from 12 locations (Figure 1, Table 1). Birds were caught using mist nets and either blood or feather samples were collected. Following Aljanabi et al. (1997), we used ~40 ul of blood to extract DNA from samples. Library preparation for ddRAD sequencing was done at the University of Laval using the enzymes *Pst*l, *Nsi*l, and *Msp*l and sequencing was done at Genome Quebec on an Illumina NovaSeq 6000 S4 PE100.

*Bioinformatic pipeline*

Sabre was used to demultiplex the ddRAD data. Adapters were removed and fastq files were trimmed to 80 bp using Cutadapt (Martin, 2011). The fastq files with the trimmed sequences were aligned to a black-capped chickadee reference genome provided by Scott Taylor (University of Colorado, Boulder) with BWA-MEM using default settings (Branch et al., 2022; Li & Durbin, 2009; Wagner et al., 2020). The resulting bam files were used to identify SNPs with gstacks. The gstacks output was used to create a vcf file and calculate summary statistics with populations using the default parameters while limiting the SNPs to one per locus (Catchen et al., 2013). SNPs were subsequently filtered using Vcftools to keep bi-allelic SNPs with a minor allele frequency ≥ 0.05 (Danecek et al., 2011). SNPs with > 50% missing data and individuals with > 30% missing data were excluded. The filtered vcf files contained 28,600 sites and 61 individuals from 11 populations; no individuals from NM remained after filtering.

*Population analyses*

To analyse the population structure, we conducted a DAPC (discriminant analysis of principal components) in R using the packages ‘adegenet’,’ape’, ‘vcfR’, and ‘ade4’ (Dray & Dufour, 2007; Jombart, 2008; Knaus & Grünwald, 2016; Paradis & Schliep, 2018). DAPC accounts for both within- and between-group differences when clustering individuals. After clustering, we plotted the linear discriminants in 2d and 3d plots using the packages ‘ggplot2’ and ‘plotly’ (Sievert, 2020; Wilkinson, 2011).

To identify the number of ancestral populations, we used ADMIXTURE with default settings iteratively for K=1-7 (Alexander et al., 2009). Linkage disequilibrium pruning and vcf-to-bed conversion were performed using PLINK (Purcell et al., 2007).

Pairwise Fst values between populations were calculated with Arlequin using the default settings (Excoffier & Lischer, 2010). Vcf-to-arp conversion was done using PGDSpider2 (Lischer & Excoffier, 2011).

*Outlier loci and environmental variation*

To identify the candidate loci under selection, we used BayeScan (Foll & Gaggiotti, 2008). BayeScan is a Bayesian algorithm based on the multinomial-Dirchlet model, which identifies loci under selection using differences in allele frequencies between populations. We set the pr_odds value at 350, with a total of 1,00,000 iterations. The pr_odds value defines the likelihood of the neutral model compared with the selection model.

To identify divergence associated with environmental variation, we used BayeScEnv for five environmental variables: temperature, precipitation, altitude, temperature seasonality and

precipitation seasonality (Villemereuil & Gaggiotti, 2015). BayeScEnv, which is based on the F-model, detects local adaptation linked to a given environmental variable using Bayesian methods. In this study, we set the pr_jump value at 0.05 with a total of 50,000 runs. The pr_jump value is similar to the pr_odds value used in BayeScan. Using the 'sp' and 'raster' packages, we extracted 30s resolution from https://www.worldclim.org/ for the following environmental variables: Annual Mean Temperature (BIO1), Temperature Seasonality (stdev×100)(BIO4), Annual Precipitation (BIO11), Precipitation Seasonality (Coefficient of Variation) (BIO11), and Elevation from the SRTM data (Bivand et al., 2013; Fick & Hijmans, 2017; Hijmans, 2020). The environmental data were standardised before the BayeScEnv analysis.

To test for isolation by distance and the effect of the above environmental variables on genetic differentiation, we conducted Mantel and partial Mantel tests using GenAlEx and the R package 'vegan' (Dixon, 2003; Mantel, 1967; Peakall & Smouse, 2012). Mantel tests are commonly used in population genetics to test for correlation between two matrices, while partial Mantel tests use a third matrix to account for another variable (Diniz-Filho et al., 2013).

*Genes of interest*

We used the gff file for black-capped chickadees provided by Scott Taylor to identify genes under selection and investigate gene–environment interactions. We ran a custom R script to identify genes present within 100 kb of the outlier loci from the BayeScan/BayeScEnv analyses. Subsequently, to identify the pathways and functions associated with these genes,

we ran a gene ontology with ShinyGO using a human and zebra finch gene set. The genes were also analysed individually (Ge et al., 2019).

*Species distribution models*

To understand the distribution of the species in the past and future, we created species distribution models for the current distribution, the last glacial maxima, and four Representative Concentration Pathways (RCP) in 2050 and 2070. RCPs describe different climatic futures depending on the emission of greenhouse gases with the four main RCPs being RCP2.6, RCP4.5, RCP6, and RCP8.6 (van Vuuren et al., 2011).

We used the Maxent algorithm in the 'Wallace' package (Kass et al., 2018). This algorithm uses occurrence data in conjunction with environmental layers to provide a potential distribution for a species. Presence data were obtained from iNaturalist observations over a 13 year period from 2010-2022 (GBIF.org, 2022) which we spatially thinned using a 10 km threshold. These points were then partitioned into six clusters using the random k-fold method (Aiello-Lammens et al., 2015). Using these 5330 points, we created a SDM with 14 environmental variables.

We tested 12 models: linear (L), quadratic (Q), hinge (H), and linear-quadratic-hinge (LQH), with a regularization multiplier between 1-3 using a step value of 1. Model selection was done using AICc and AUC values (Kass et al., 2021; Phillips, 2017). The trained models were projected onto the LGM and RCP scenarios using environmental layers obtained from http://www.paleoclim.org and CCSM4, respectively using a 10% minimum training presence threshold (Brown et al., 2018; Hill, 2015; Karger et al., 2017; Pisias & Moore, 1981).

## Results

*Population structure*

We looked at over 28,500 loci in 61 individuals from 11 populations after filtering. Prior to filtering, there were over 1,000,000 SNPs for 94 individuals from 12 populations. The DAPC, consistent with Hindley et al. (2018), revealed four distinct clusters: southern CA; northern CA/southern OR; eastern Rockies (MT, UT, SAB); and northwestern Rockies (WA, BCR, FTSJ, PG, NWBC) (Figure 2a). The first three linear discriminants (LD) explained 98.5% of the variance, while the fourth explained approximately 1.5%. Pairwise LD plots (Figure 2b) accounted for the effect of altitude on genetic differentiation.

As ADMIXTURE does not predict the K value, we ran the analysis for K=1-7 (Figure 3a-b). Figure 3(b) shows the CV errors for each K value, where K=1-3 are strongly supported (CV error 0.668-0.696). K=4-6 coincide with the DAPC despite having a slightly higher CV error (0.758-1.026). For K=4, the SCA population splits from the SOR/NCA population to form a separate cluster consistent with the DAPC analysis. At K=5 WA separates from the other northwestern Rocky populations and BCR individuals and three FTSJ individuals show admixture with WA.

The pairwise Fst values, consistent with the DAPC, showed four broad clusters, with the SCA population having the highest Fst values between every pair (Figure 4). The p-values were significant for most cases, except those that included NWBC, PG, or SAB populations. Within the northwestern Rockies group, WA was significantly different from all of the other populations except PG. Pairwise Fst values ranged from 0.02 to 0.37, the highest significant

value is seen between BCR-SCA while the smallest are between the BCR-FTSJ and UT-MT populations.

*Outlier loci and gene-environment analysis*

To identify areas of genomic divergence and their association with environmental factors, we used BayeScan and BayeScEnv, two commonly used genome scan programs (Foll & Gaggiotti, 2008; Villemereuil & Gaggiotti, 2015). Outlier loci associated with temperature (Temp), altitude (Alt), precipitation seasonality (PS), temperature seasonality (TS), and precipitation (Prec) were identified using BayeScEnv, whereas BayeScan was used to identify all possible outliers, regardless of their association with environmental variables.

We identified 2251 outlier loci at a false discovery rate (FDR) of 0.0001 using BayeScan, despite using a conservative model with a pr_odds value of 350 (Figure 5a). Several SNPs had a false discovery rate of 0, which was converted to $10e^{-10}$ for visualisation. Similarly, the BayeScEnv analysis, with an FDR of 0.0001, showed 1564 outlier SNPs associated with temperature, 2060 with temperature seasonality, 805 with precipitation, 1090 with precipitation seasonality, and 1606 with altitude (Figure 5b-f). Several SNPs had a false discovery rate of 0, which was converted to $10e^{-6}$ for visualisation. We used different values for visualisation between the BayeScan and BayeScEnv analysis due to several SNPs having FDRs of $10e^{-6}$ in the former.

To test whether the large number of outlier loci was due to the isolated SCA population, we ran the same tests for pairs of the identified DAPC clusters and with all populations except for SCA. We found a significant decrease in the number of outlier loci identified in the pairwise

cluster tests. The decrease differed amongst clusters, but the number of outlier loci ranged from 4-800. In the case of precipitation, the outliers ranged from 30-500 compared to 805. However, there was no change when we repeated the analysis with every population excluding the SCA population.

We performed Mantel and partial Mantel tests to investigate the possibility of isolation by distance and the effect of the aforementioned environmental variables on divergence. Geographic distance ($R_{xy}$= 0.54) was most highly correlated with genetic distance, followed by temperature ($R_{xy}$= 0.40) (Figure 6a-b). All the results were statistically significant ($p < 0.0003$). The partial Mantel tests highlighted a similar pattern; however, when the z-axis was either temperature, temperature seasonality, or geographic distance; the results for precipitation seasonality did not meet the cut-off criterion of $p = 0.05$. In addition, the results of temperature seasonality when the z-axis was geographic distance did not meet the cut-off criteria.

*Genes of interest*

We identified 181 genes within 100 kb of temperature-associated SNPs and 189 genes from the BayeScan outlier SNPs. Genes of interest were analysed only for these two cases because of the overlap in SNPs across environmental variables. ShinyGO identified 15 significant pathways (Supplementary, gene_SuppInfo.xlsx) for the BayeScan genes, which included thermogenesis ($p = 4.58e^{-5}$).

Subsequently, we grouped the genes based on their biological functions using ShinyGO (Ge et al., 2019). The BayeScan analysis revealed 37 genes associated with response to stress, 22 genes related to immune system processes, and three genes associated with circadian rhythm.

CSNK1D, a clock gene, was one of the genes associated with circadian rhythm, which has been shown to play a role in migration (Steinmeyer et al., 2009). In addition, several genes affecting immune response, response to external stimuli, and growth were found to be associated with outlier SNPs. We also identified genes associated with the reproductive system in all the analyses. The JAM3 gene is known to play a role in spermatogenesis and the RNF212 gene has been shown to influence meiotic recombination. A similar grouping of genes was observed using the BayeScEnv temperature-associated SNPs, with the number of genes associated with each function differing.

We obtained similar results when using the zebra finch genome, although the number of unmapped genes was higher than that in the human genome (BayeScan – 137 vs 87 unmapped, BayeScEnv – 108 vs 64 unmapped).

*Species distribution models*

We created species distribution models to understand the effects of climate change on mountain chickadee habitat (Figure 7). We ran 12 models with a regularisation multiplier between 1-3 with a step value of one. While all models had an AUC value of 0.93 or above, the hinge model with regularisation multiplier 1 was chosen as the best model based on AICc and AUC values.

The SDMs show a significant decrease in suitable habitat over the next five decades across all RCP scenarios except RCP2.6, which is an ideal projection of future climate. In addition, a northward shift in habitat is also observed. The SDM for the last glacial maximum is consistent with previous studies, with populations present near the coast and extending into

Mexico (Manthey et al., 2012) (Figure 7d). The SDMs for other RCPs, response curves, and model statistics are available in the supplementary material.

## Discussion

In this study, apart from delineating population structure, we aimed to answer the following questions: (1) What are the environmental drivers of genetic differentiation? (2) Which genes or pathways are undergoing selection across populations? (3) How much will climate change affect the habitat of mountain chickadees over the next 50 years?

DAPC revealed four primary clusters, which coincided with our ADMIXTURE and pairwise Fst values. The eastern Rocky Mountain cluster (UT, MT, SAB) is separated from the northwestern Rocky Mountain populations by the Rocky Mountains. These populations could have diverged because of the presence of physical barriers or differences in habitat on either side of the mountains. This is evident from the high, but not significant, pairwise Fst values between the BCR and SAB populations, which are geographically close (300 km), but separated by the Rockies. In addition, the SCA population is of particular interest because it is genetically isolated from all other populations. As indicated by the Mantel tests and previous studies, this could be due to its distance from other populations or due to geographic/environmental features within its habitat (Spellman et al., 2007; Hindley et al., 2018).

Our outlier SNP analysis revealed several loci under selection associated with environmental conditions. Despite the use of conservative models, we identified several outlier loci

highlighting the genetic diversity of the species. However, because the large number of loci could have been due to the isolated SCA population, we conducted the same analyses by (1) excluding the SCA population and (2) between the DAPC clusters. There were no significant differences in the first case, whereas we observed a decrease in the second case. This decrease can be attributed to the low genetic distances between populations within the same cluster. Additionally, the overlap of SNPs across analyses and the existence of several SNPs with an FDR of zero highlights the genetic diversity present in the species. Divergence across populations is expected because of the reduced gene flow among them due to the non-migratory nature of the species (Templeton, 2006; Eckert et al., 2008; McCallum et al., 2020).

Consistent with mtDNA and microsatellite studies, genetic diversity is highly influenced by geographic distance (Hindley et al., 2018). The strong correlation between geographic and genetic distance ($R_{xy}$ = 0.55) indicates isolation by distance. However, we cannot discount the role of habitat differences, given that local temperature ($R_{xy}$ = 0.4), temperature seasonality ($R_{xy}$ = 0.23), and altitude ($R_{xy}$ = 0.2) are also significantly correlated with genetic distance. The rapid increase in global temperatures could affect the genetic isolation in the coming years. Additionally, an increase in temperature forces species to shift their elevational range. Precipitation and precipitation seasonality are predicted to increase with rising temperatures, resulting in more extreme climate scenarios (Boer, 2009; Pendergrass et al., 2017). As a result, despite their weak correlation with divergence, these factors could play a major role in the future of these species.

We identified genes associated with SNPs undergoing selection. Over 30 genes linked to stress response were found to be near with SNPs associated with temperature. We also found

over 19 genes linked to response to external stimuli, and ShinyGO analysis revealed that the thermogenesis pathway had a significant number of genes involved. Previous studies have shown that increasing temperatures are linked to the activation of the stress response in birds in the form of thermoregulatory strategies such as panting and increased glucocorticoid levels (Bohler et al., 2021; Mentesana & Hau, 2022; Siegel, 1980). This is of particular concern because of rising global temperatures, which could lead to negative consequences for the species. In addition, the selection of a clock gene, CSNK1D, could imply a change in phenology in response to rising temperatures (Milligan et al., 2009).

We observed a northward shift in suitable habitat over the next five decades with the SDM. This pattern was observed for all RCP scenarios, except for RCP2.6. However, RCP2.6 is an ideal scenario where all expected climate change goals are fulfilled, and temperatures increase by 1°C above pre-industrial levels by the year 2050 and remain the same in 2070 (van Vuuren et al., 2011). While shifting to cooler habitats is a normal thermoregulatory response, the massive decrease in suitable habitats for such a common species is worrying (Siegel, 1980). Additionally, because the model does not account for other factors, such as human-induced habitat loss, competition, and invasion; the amount of suitable habitat could be much less than predicted. This could lead to a further population decline in this species.

## Conclusions

Mountain chickadee genetic distance is highly correlated with geographical distance and temperature. Genes affecting several essential functions associated with outlier SNPs were identified, highlighting the genetic diversity and selection pressure faced by the species. The

identification of genes related to circadian rhythm may underlie changes in phenology. In addition, the large decrease in suitable habitat over the next five decades for a common species highlights the need for immediate action to protect this species and other species from extinction.

**Acknowledgements**

We thank NSERC Discovery Grant, Alberta Innovates and Mitacs for funding and Burke Museum of Natural History, University of Michigan Museum of Natural History, University of California, Berkeley Museum of Vertebrate Zoology, Louisiana State University Museum of Natural Science, and Smithsonian Museum of Natural Science for providing some of the samples used in this study. Samples were also collected by members of the Burg Lab at the University of Lethbridge including John Hindley, Brendan Graham, Linda Lait, Paulo Pulgarin, Kim Dohms and numerous field assistants. We also thank the National Park Service, USFWS, Parks Canada and Canada Wildlife Service for their assistance.

## Literature cited


Adams, R. V, & Burg, T. M. (2015). Influence of ecological and geological features on rangewide patterns of genetic structure in a widespread passerine. *Heredity*, *114*(2), 143–154. https://doi.org/10.1038/hdy.2014.64

Aiello-Lammens, M. E., Boria, R. A., Radosavljevic, A., Vilela, B., & Anderson, R. P. (2015). spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography*, *38*(5), 541–545. https://doi.org/10.1111/ecog.01132

Alan R. Templeton. (2006). Gene Flow and Population Subdivision. In *Population Genetics and Microevolutionary Theory* (pp. 169–236). Wiley. https://doi.org/10.1002/9781119836070.ch6

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*(9), 1655–1664. https://doi.org/10.1101/gr.094052.109

Antonelli, A. (2017). Biogeography: Drivers of bioregionalization. *Nature Ecology & Evolution*, *1*(4), 0114. https://doi.org/10.1038/s41559-017-0114

Beniston, M. (2003). Climatic change in mountain regions: A review of possible impacts. In *Advances in Global Change Research* (pp. 5–31). Springer Netherlands. https://doi.org/10.1007/978-94-015-1252-7_2

Bivand, R. S., Pebesma, E., & Gómez-Rubio, V. (2013). Classes for spatial data in R. In *Applied Spatial Data Analysis with R* (pp. 21–57). Springer New York. https://doi.org/10.1007/978-1-4614-7618-4_2

Boer, G. J. (2009). Changes in interannual variability and decadal potential predictability under global warming. *Journal of Climate*, *22*(11), 3098–3109. https://doi.org/10.1175/2008JCLI2835.1

Bohler, M. W., Chowdhury, V. S., Cline, M. A., & Gilbert, E. R. (2021). Heat stress responses in birds: A review of the neural components. *Biology*, *10*(11), 1095. https://doi.org/10.3390/biology10111095

Boutilier, S. T., Taylor, S. A., Morris-Pocock, J. A., Lavoie, R. A., & Friesen, V. L. (2014). Evidence for genetic differentiation among Caspian tern (*Hydroprogne caspia*) populations in North America. *Conservation Genetics*, *15*(2), 275–281. https://doi.org/10.1007/s10592-013-0536-1

Brown, J. L., Hill, D. J., Dolan, A. M., Carnaval, A. C., & Haywood, A. M. (2018). PaleoClim, high spatial resolution paleoclimate surfaces for global land areas. *Scientific Data*, *5*(1), 180254. https://doi.org/10.1038/sdata.2018.254

Calkins, M. T., Beever, E. A., Boykin, K. G., Frey, J. K., & Andersen, M. C. (2012). Not-so-splendid isolation: Modeling climate-mediated range collapse of a montane mammal *Ochotona princeps* across numerous ecoregions. *Ecography*, *35*(9), 780–791. https://doi.org/10.1111/j.1600-0587.2011.07227.x

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124–3140. https://doi.org/10.1111/mec.12354

Christmann, T., & Menor, I. O. (2021). A synthesis and future research directions for tropical mountain ecosystem restoration. *Scientific Reports*, *11*(1), 23948. https://doi.org/10.1038/s41598-021-03205-y

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Diniz-Filho, J. A. F., Soares, T. N., Lima, J. S., Dobrovolski, R., Landeiro, V. L., Telles, M. P. de C., Rangel, T. F., & Bini, L. M. (2013). Mantel test in population genetics. *Genetics and Molecular Biology*, *36*(4), 475–485. https://doi.org/10.1590/s1415-47572013000400002

Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science*, *14*(6), 927–930. https://doi.org/10.1111/j.1654-1103.2003.tb02228.x

Dray, S., & Dufour, A.-B. (2007). Theade4Package: Implementing the duality diagram for ecologists. *Journal of Statistical Software*, *22*(4). https://doi.org/10.18637/jss.v022.i04

Eckert, C. G., Samis, K. E., & Lougheed, S. C. (2008). Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. *Molecular Ecology*, *17*(5), 1170–1188. https://doi.org/10.1111/j.1365-294X.2007.03659.x

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*(3), 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1□km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315. https://doi.org/10.1002/joc.5086

Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, *180*(2), 977–993. https://doi.org/10.1534/genetics.108.092221

GBIF.org. (2022). *GBIF Download*. https://doi.org/10.15468/dl.tgh2v7

Ge, S. X., Jung, D., & Yao, R. (2019). ShinyGO: A graphical gene-set enrichment tool for animals and plants. *Bioinformatics*, *36*(8), 2628–2629. https://doi.org/10.1093/bioinformatics/btz931

Greenwood, P. J., & Harvey, P. H. (1982). The natal and breeding dispersal of birds. *Annual Review of Ecology and Systematics*, *13*(1), 1–21. https://doi.org/10.1146/annurev.es.13.110182.000245

Hanski, I., Saastamoinen, M., & Ovaskainen, O. (2006). Dispersal-related life-history trade-offs in a butterfly metapopulation. *Journal of Animal Ecology*, *75*(1), 91–100. https://doi.org/10.1111/j.1365-2656.2005.01024.x

Hijmans, R. J. (2020). raster: Geographic analysis and modeling with raster data. *R Package Version 3.1-5*, *3.1*(5).

Hill, D. J. (2015). The non-analogue nature of Pliocene temperature gradients. *Earth and Planetary Science Letters*, *425*, 232–241. https://doi.org/10.1016/j.epsl.2015.05.044

Hindley, J. A., Graham, B. A., Pulgarin-R., P. C., & Burg, T. M. (2018). The influence of latitude, geographic distance, and habitat discontinuities on genetic variation in a high latitude montane species. *Scientific Reports*, *8*(1), 11846 https://doi.org/10.1038/s41598-018-29982-7

IUCN Red List. (n.d.). *IUCN. 2022. The IUCN Red List of Threatened Species. Version 2022-2. https://www.iucnredlist.org. Accessed on 28 December 2022.*

Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P., & Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. *Scientific Data*, *4*(1), 170122. https://doi.org/10.1038/sdata.2017.122

Kass, J. M., Muscarella, R., Galante, P. J., Bohl, C. L., Pinilla□Buitrago, G. E., Boria, R. A., Soley□Guardia, M., & Anderson, R. P. (2021). ENMeval 2.0: Redesigned for customizable and reproducible modeling of species' niches and distributions. *Methods in Ecology and Evolution*, *12*(9), 1602–1608. https://doi.org/10.1111/2041-210X.13628

Kass, J. M., Vilela, B., Aiello□Lammens, M. E., Muscarella, R., Merow, C., & Anderson, R. P. (2018). Wallace: A flexible platform for reproducible modeling of species niches and distributions built for community expansion. *Methods in Ecology and Evolution*, *9*(4), 1151–1156. https://doi.org/10.1111/2041-210X.12945

Keyghobadi, N., Roland, J., & Strobeck, C. (1999). Influence of landscape on the population genetic structure of the alpine butterfly *Parnassius smintheus* (Papilionidae). *Molecular Ecology*, *8*(9), 1481–1495. https://doi.org/10.1046/j.1365-294x.1999.00726.x

Knaus, B. J., & Grünwald, N. J. (2016). vcfr: A package to manipulate and visualize variant call format data in R. *Molecular Ecology Resources*, *17*(1), 44–53. https://doi.org/10.1111/1755-0998.12549

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Lischer, H. E. L., & Excoffier, L. (2011). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, *28*(2), 298–299. https://doi.org/10.1093/bioinformatics/btr642

Machado, A. P., Clément, L., Uva, V., Goudet, J., & Roulin, A. (2018). The Rocky Mountains as a dispersal barrier between barn owl (*Tyto alba*) populations in North America. *Journal of Biogeography*, *45*(6), 1288–1300. https://doi.org/10.1111/jbi.13219

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Res*, *27*(2), 209-220.

Manthey, J. D., Klicka, J., & Spellman, G. M. (2012). Is gene glow promoting the reversal of Pleistocene divergence in the mountain chickadee (*Poecile gambeli*)? *PLoS ONE*, *7*(11), e49218. https://doi.org/10.1371/journal.pone.0049218

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, *17*(1), 10. https://doi.org/10.14806/ej.17.1.200

McCallum, D. A., Grundel, R., & Dahlsten, D. L. (2020). Mountain Chickadee (*Poecile gambeli*). In *Birds of the World*. Cornell Lab of Ornithology. https://doi.org/10.2173/bow.mouchi.01

McDonald, K. A., & Brown, J. H. (1992). Using montane mammals to model extinctions due to global change. *Conservation Biology*, *6*(3), 409–415. https://doi.org/10.1046/j.1523-1739.1992.06030409.x

Mentesana, L., & Hau, M. (2022). Glucocorticoids in a warming world: Do they help birds to cope with high environmental temperatures? *Hormones and Behavior*, *142*, 105178. https://doi.org/10.1016/j.yhbeh.2022.105178

Milligan, S. R., Holt, W. V., & Lloyd, R. (2009). Impacts of climate change and environmental factors on reproduction and development in wildlife. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1534), 3313–3319. https://doi.org/10.1098/rstb.2009.0175

Paradis, E., & Schliep, K. (2018). ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*(3), 526–528. https://doi.org/10.1093/bioinformatics/bty633

Parmesan, C. (2006). Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution, and Systematics*, *37*(1), 637–669. https://doi.org/10.1146/annurev.ecolsys.37.091305.110100

Payne, D., Spehn, E. M., Prescott, G. W., Geschke, J., Snethlage, M. A., & Fischer, M. (2020). Mountain biodiversity is central to sustainable development in mountains and beyond. *One Earth*, *3*(5), 530–533. https://doi.org/10.1016/j.oneear.2020.10.013

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research--an update. *Bioinformatics*, *28*(19), 2537–2539. https://doi.org/10.1093/bioinformatics/bts460

Pendergrass, A. G., Knutti, R., Lehner, F., Deser, C., & Sanderson, B. M. (2017). Precipitation variability increases in a warmer climate. *Scientific Reports*, *7*(1), 17966. https://doi.org/10.1038/s41598-017-17966-y

Phillips, S. (2017). Maxnet: Fitting 'Maxent' Species Distribution Models with "glmnet." *R Package*.

Pisias, N. G., & Moore, T. C. (1981). The evolution of Pleistocene climate: A time series approach. *Earth and Planetary Science Letters*, *52*(2), 450–458. https://doi.org/10.1016/0012-821X(81)90197-7

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for

whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

Siegel, H. S. (1980). Physiological stress in birds. *BioScience*, *30*(8), 529–534. https://doi.org/10.2307/1307973

Sievert, C. (2020). *Interactive Web-Based Data Visualization with R, plotly, and shiny*. Chapman and Hall/CRC. https://doi.org/10.1201/9780429447273

Spellman, G. M., Riddle, B., & Klicka, J. (2007). Phylogeography of the mountain chickadee (*Poecile gambeli*): Diversification, introgression, and expansion in response to Quaternary climate change. *Molecular Ecology*, *16*(5), 1055–1068. https://doi.org/10.1111/j.1365-294x.2007.03199.x

Steinmeyer, C., Mueller, J. C., & Kempenaers, B. (2009). Search for informative polymorphisms in candidate genes: clock genes and circadian behaviour in blue tits. *Genetica*, *136*(1), 109–117. https://doi.org/10.1007/s10709-008-9318-y

van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., Lamarque, J.-F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S. J., & Rose, S. K. (2011). The representative concentration pathways: an overview. *Climatic Change*, *109*(1–2), 5–31. https://doi.org/10.1007/s10584-011-0148-z

Villemereuil, P., & Gaggiotti, O. E. (2015). A new FST based method to uncover local adaptation using environmental variables. *Methods in Ecology and Evolution*, *6*(11), 1248–1258. https://doi.org/10.1111/2041-210x.12418

Walther, G-R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T. J. C., Fromentin, J-M., Hoegh-Guldberg, O., & Bairlein, F. (2002). Ecological responses to recent climate change. *Nature*, *416*(6879), 389–395. https://doi.org/10.1038/416389a

Wilkinson, L. (2011). ggplot2: Elegant graphics for data analysis by Wickham, H. *Biometrics*, *67*(2), 678–679. https://doi.org/10.1111/j.1541-0420.2011.01616.x

Wilson, R. J., Gutiérrez, D., Gutiérrez, J., & Monserrat, V. J. (2007). An elevational shift in butterfly species richness and composition accompanying recent climate change. *Global Change Biology*, *13*(9), 1873–1887. https://doi.org/10.1111/j.1365-2486.2007.01418.x

## Data accessibility and benefit sharing

*Data accessibility*

Multiplexed ddRAD sequence data will be uploaded to a data repository prior to publication.

*Benefit sharing*

A research collaboration was developed between researchers from India and Canada with funding from Mitacs. Our research also addresses a priority concern, climate change, and provides benefits by the sharing of our data and results on public databases.

## Author contributions

The authors confirm contribution to the paper as follows: Study conception and design: Srikanthan and Burg. Data collection: Burg. Data analyses: Srikanthan. Wrote the paper: Srikanthan and Burg. All authors reviewed the results and approved the final version of the manuscript.

## Tables

Table 1 List of sampling sites and sample sizes for mountain chickadees used in this study. See Fig 1 for location.

| Location | Abbreviation | Sample size |
|---|---|---|
| Revelstoke, BC | BCR | 4 |
| Fort St James, BC | FTSJ | 8 |
| Northwest British Columbia | NWBC | 2 |
| Prince George, BC | PG | 2 |
| Washington | WA | 8 |
| Montana | MT | 7 |
| Utah | UT | 10 |
| Southern Alberta | SAB | 2 |
| Southern Oregon | SOR | 10 |
| Northern California | NCA | 5 |
| Southern California | SCA | 3 |

## Figure Legends:

Figure 1 Mountain chickadee sampling sites in western North America along with the range map from the IUCN Red List (pink)

Figure 2 (a) 3d DAPC plot showing four distinct clusters(b) LD1 vs LD2 (c) LD1 vs LD3 (d) LD1 vs LD4

Figure 3 (a) Admixture plot for K=2-7 (b) Cross validation errors for K=1-7

Figure 4 Pairwise Fst values for 11 populations. Asterisks indicate non-significant observations (p-values >0.05).

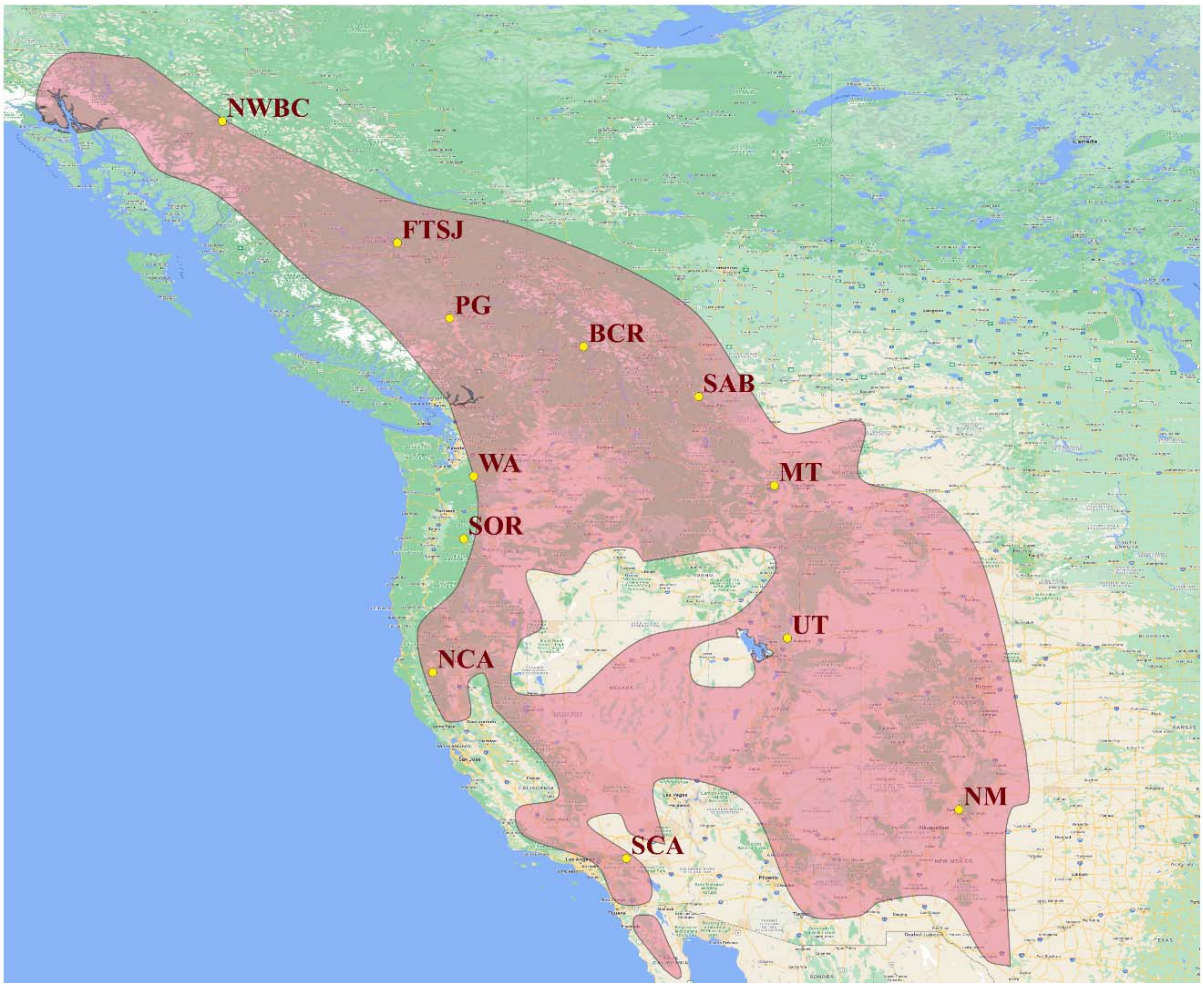
Figure 5 BayeScan and BayScEnv plots with correlation q-values for genetic divergence where q-value = -log(FDR). (a) BayeScan plot (N=2251, FDR=0.001). (b) BayScEnv Temperature plot (N=1564, FDR=$10^{-4}$). (c) BayScEnv Temperature Seasonality plot (N=2060, FDR=$10^{-4}$). (d) BayScEnv Precipitation plot (N=805, FDR=$10^{-4}$). (e) BayScEnv Precipitation Seasonality plot (N=1090, FDR=$10e^{-4}$). (f) BayScEnv altitude plot (N=1606, FDR=$10e^{-4}$).

Figure 6 (a) Mantel test $R_{xy}$ and p-values. Geographic distance-Dist, Precipitation-Prec, Precipitation seasonality-PS, Temp -Temperature, Temperature Seasonality-TS, Altitude-Alt (b) Geographic vs genetic distance plot.

Figure 7 Species distribution models of mountain chickadees using the $10^{th}$ percentile training presence threshold. Legend: green-presence, brown-absence (a) Current distribution. (b) 2050 distribution under RCP 8.5 (c) 2070 distribution under RCP 8.5. (d) Distribution during the last glacial maxima 21 kya.

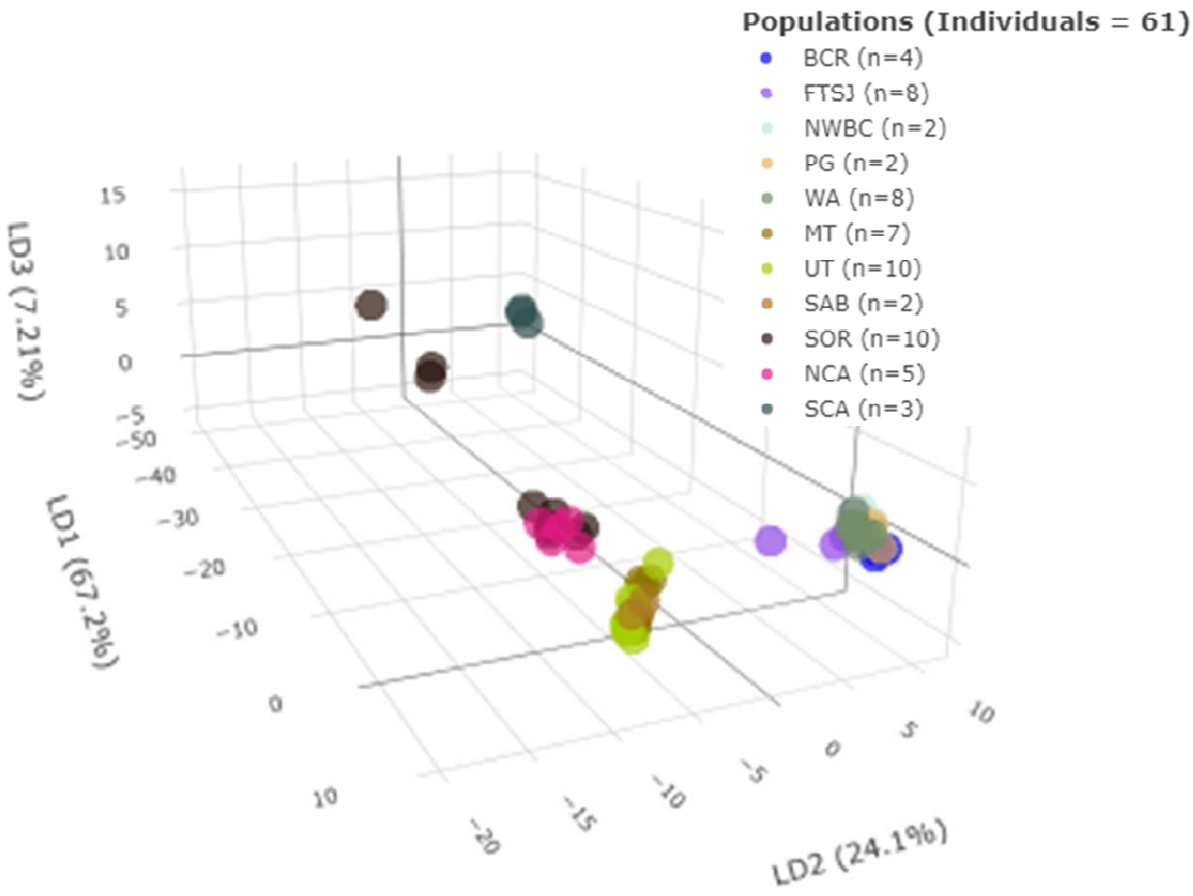
**Figures:**

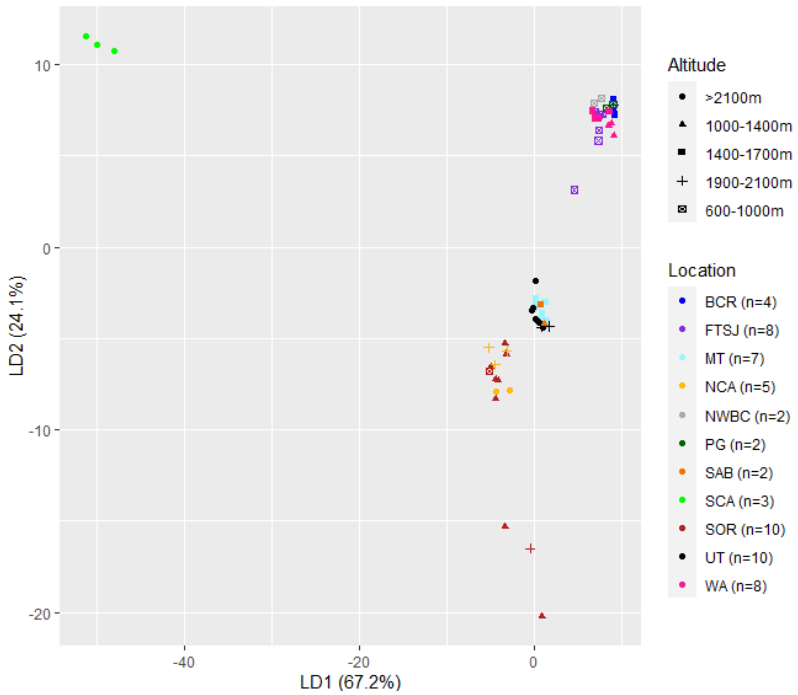
Figure 1

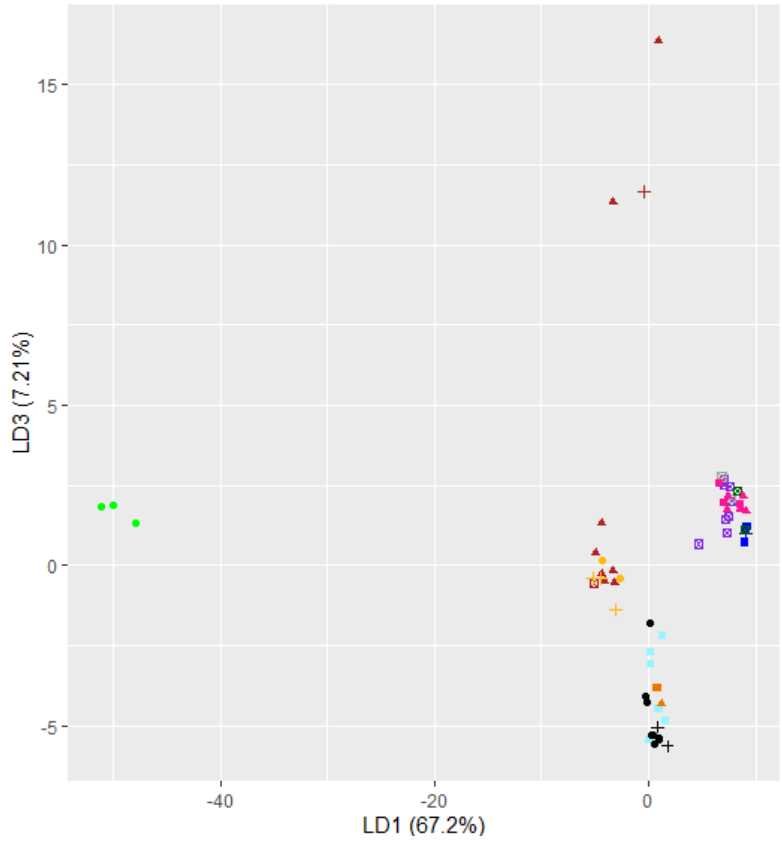


Figu

re 2


Populations (Individuals = 61)
BCR (n=4)
FTSJ (n=8)
NWBC (n=2)
PG (n=2)
WA (n=8)
MT (n=7)
UT (n=10)
SAB (n=2)
SOR (n=10)
NCA (n=5)
SCA (n=3)
LD3 (7.21%)
LD1 (67.2%)
LD2 (24.1%)


Altitude
>2100m
1000-1400m
1400-1700m
1900-2100m
600-1000m
Location
BCR (n=4)
FTSJ (n=8)
MT (n=7)
NCA (n=5)
NWBC (n=2)
PG (n=2)
SAB (n=2)
SCA (n=3)
SOR (n=10)
UT (n=10)
WA (n=8)
LD2 (24.1%)
LD1 (67.2%)
10
0
-10
-20
-40
-20
0


LD3 (7.21%)
LD1 (67.2%)
15
10
5
0
-5
-40
-20
0

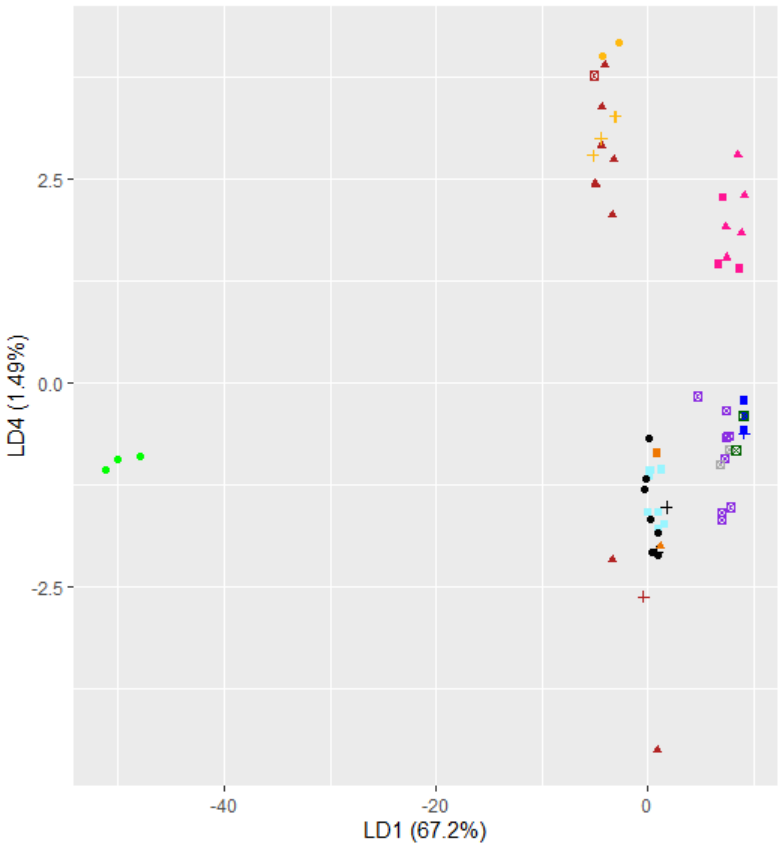
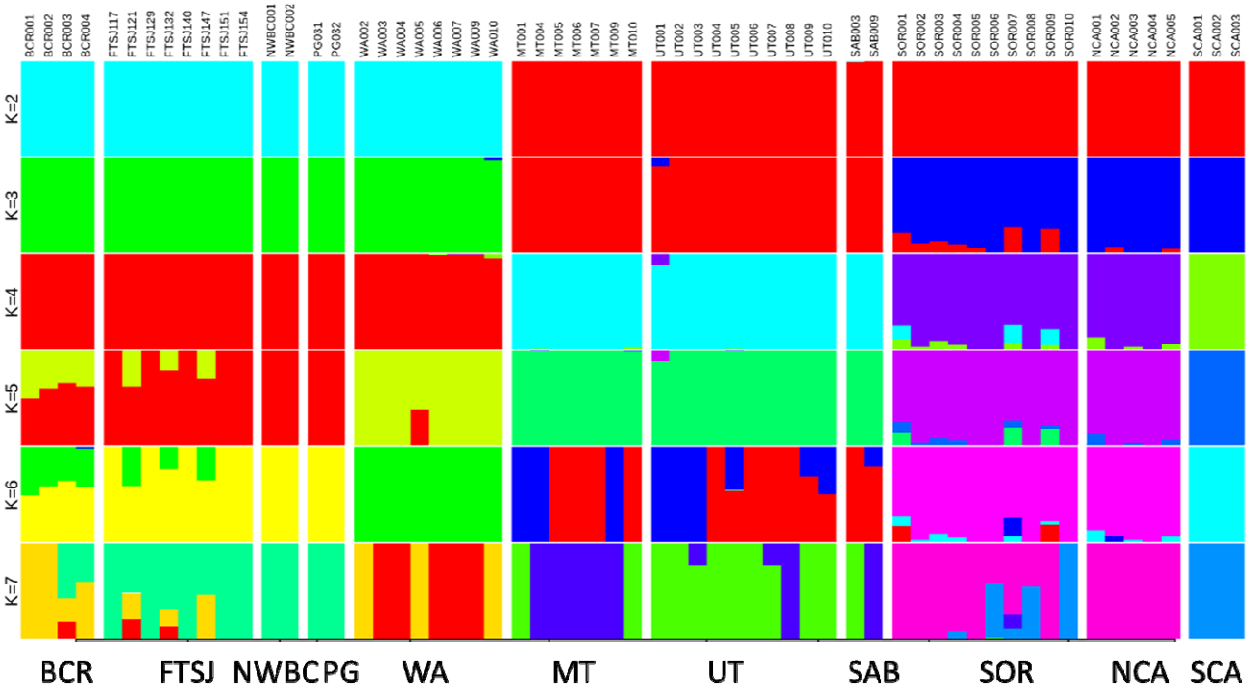

2.5
0.0
-2.5
LD4 (1.49%)
-40
-20
0
LD1 (67.2%)

Figure 3


BCR001
BCR002
BCR003
BCR004
FTSJ117
FTSJ121
FTSJ129
FTSJ132
FTSJ140
FTSJ147
FTSJ151
FTSJ154
NWBC001
NWBC002
PG031
PG032
WA002
WA003
WA004
WA005
WA006
WA007
WA009
WA010
MT001
MT004
MT005
MT006
MT007
MT009
MT010
UT001
UT002
UT003
UT004
UT005
UT006
UT007
UT008
UT009
UT010
SAB003
SAB009
SOR001
SOR002
SOR003
SOR004
SOR005
SOR006
SOR007
SOR008
SOR009
SOR010
NCA001
NCA002
NCA003
NCA004
NCA005
SCA001
SCA002
SCA003
K=2
K=3
K=4
K=5
K=6
K=7
BCR
FTSJ
NWBC
PG
WA
MT
UT
SAB
SOR
NCA
SCA

| Number of ancestral populations (K) | CV error |
|---|---|
| 1 | 0.668 |
| 2 | 0.672 |
| 3 | 0.696 |
| 4 | 0.758 |
| 5 | 0.830 |
| 6 | 0.941 |
| 7 | 1.026 |

Figure 4

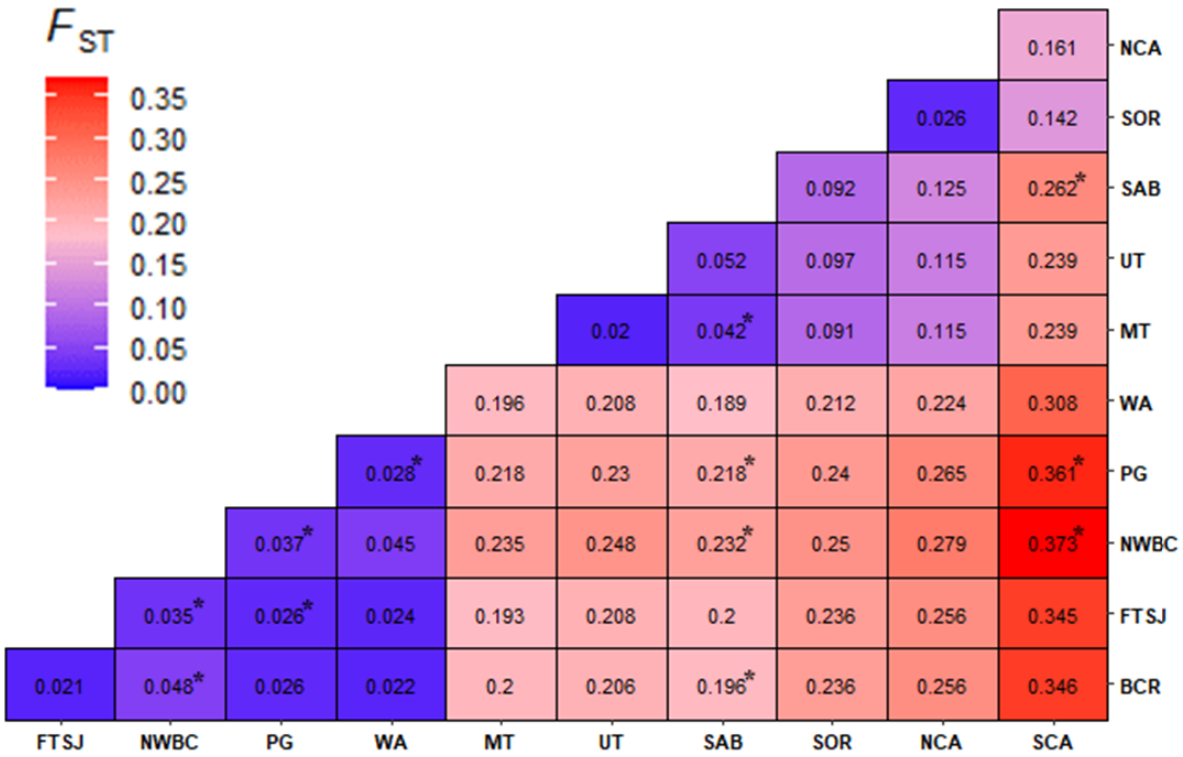
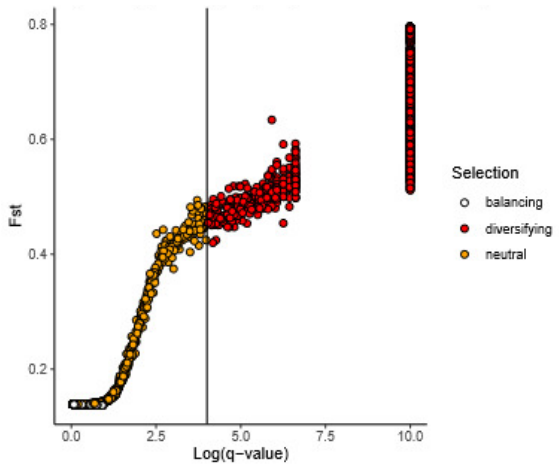
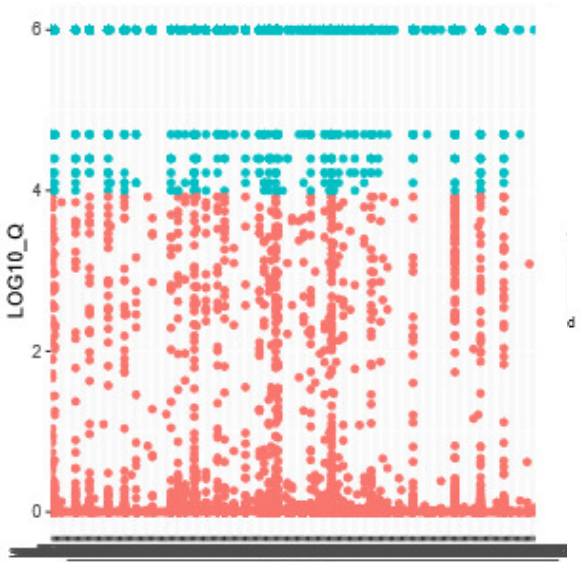
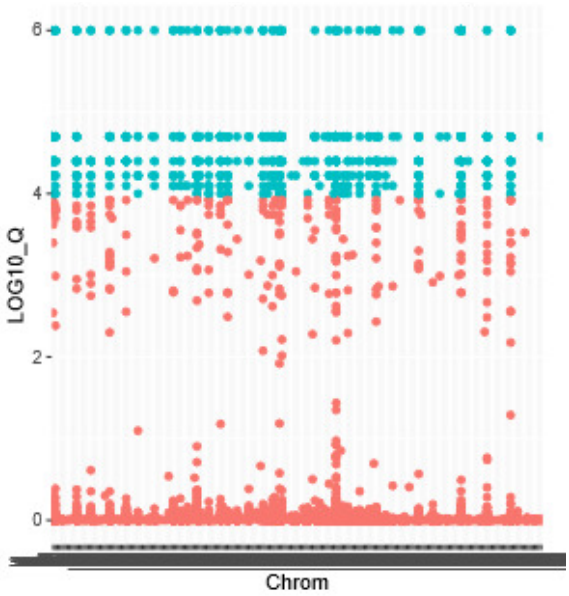
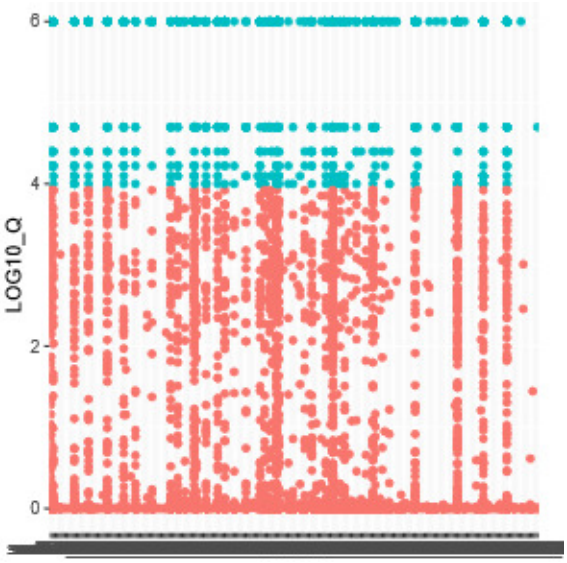

$F_{ST}$
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0.161 NCA
0.026 0.142 SOR
0.092 0.125 0.262* SAB
0.052 0.097 0.115 0.239 UT
0.02 0.042* 0.091 0.115 0.239 MT
0.196 0.208 0.189 0.212 0.224 0.308 WA
0.028* 0.218 0.23 0.218* 0.24 0.265 0.361* PG
0.037* 0.045 0.235 0.248 0.232* 0.25 0.279 0.373* NWBC
0.035* 0.026* 0.024 0.193 0.208 0.2 0.236 0.256 0.345 FTSJ
0.021 0.048* 0.026 0.022 0.2 0.206 0.196* 0.236 0.256 0.346 BCR
FTSJ NWBC PG WA MT UT SAB SOR NCA SCA

Figure 5


Fst
0.8
0.6
0.4
0.2
0.0
2.5
5.0
7.5
10.0
Log(q−value)
Selection
balancing
diversifying
neutral


LOG10_Q
6
4
2
0


6
4
2
0
LOG10_Q
Chrom


6
4
2
0
LOG10_Q

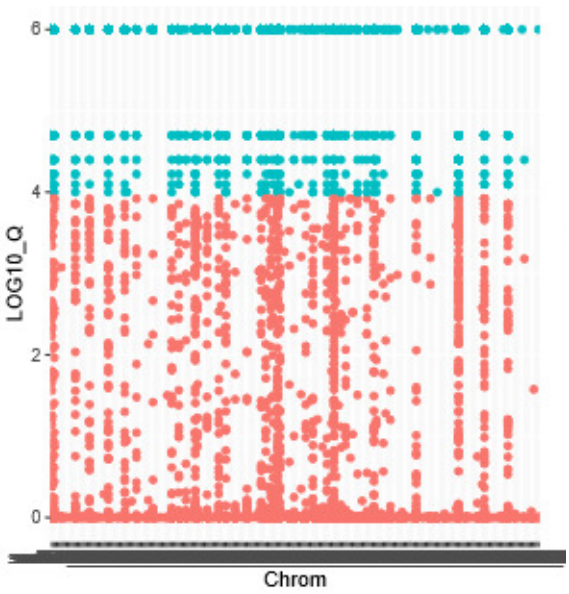

6
4
2
0
LOG10_Q
Chrom

Figure 6

| Mantel Test | | |
|---|---|---|
| **X-axis** | **Rxy** | **P_Value** |
| **Dist** | 0.547 | 0.0001 |
| **Prec** | 0.197 | 0.0002 |
| **PS** | 0.184 | 0.0003 |
| **Temp** | 0.402 | 0.0001 |
| **TS** | 0.233 | 0.0001 |
| **Alt** | 0.206 | 0.0001 |

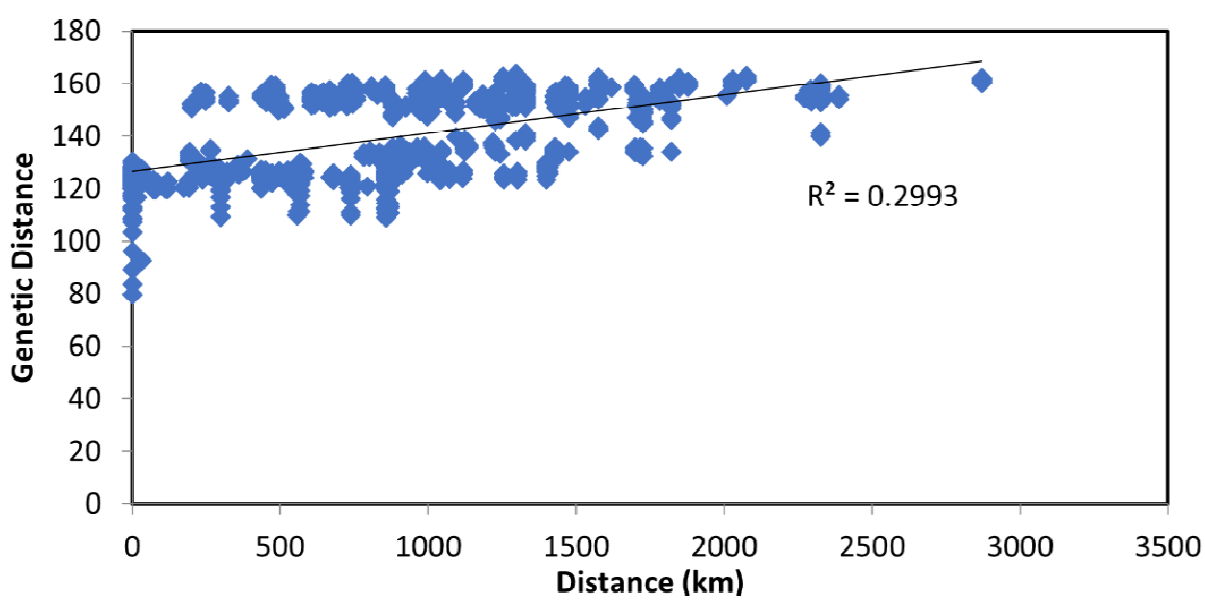
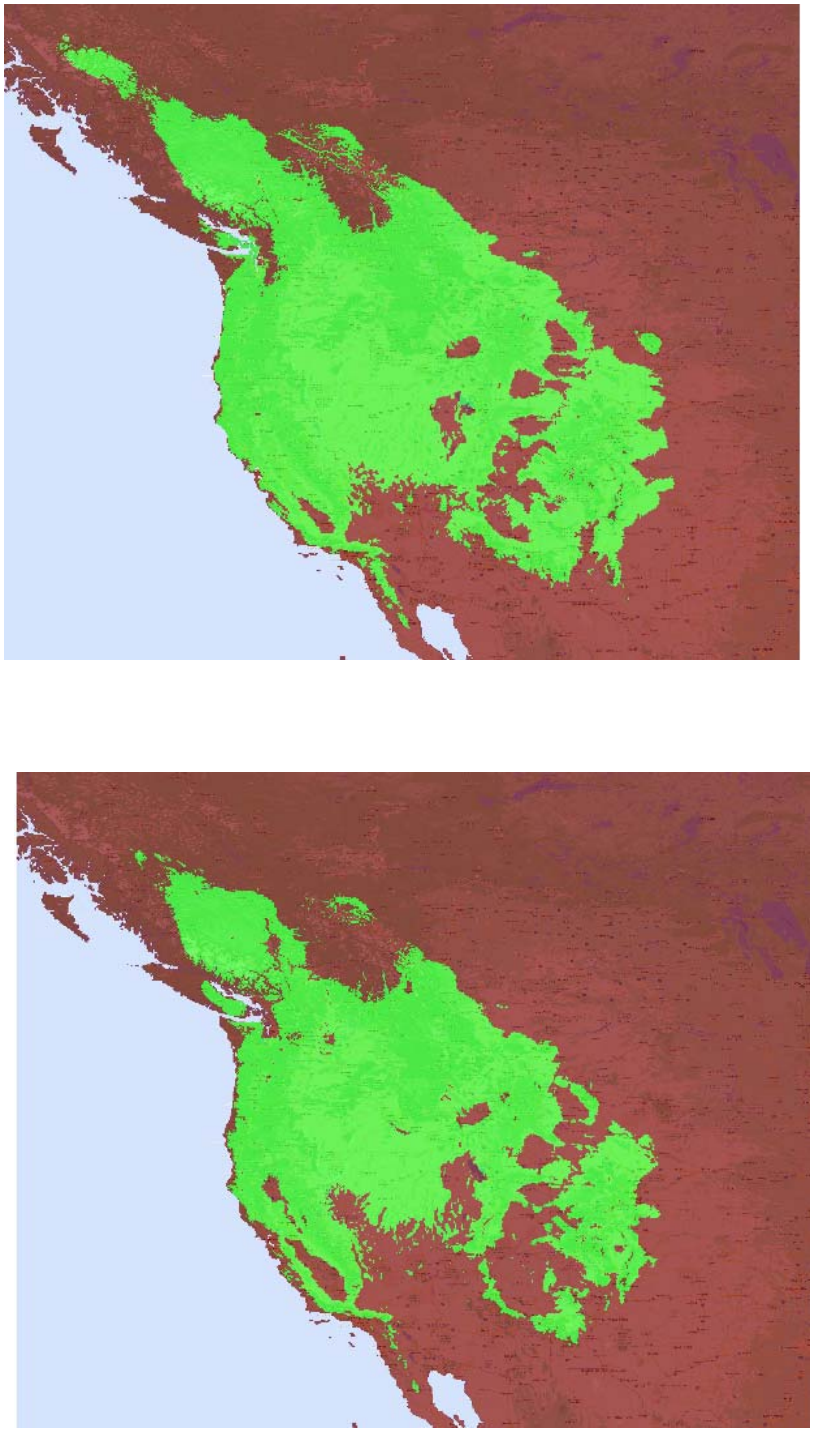
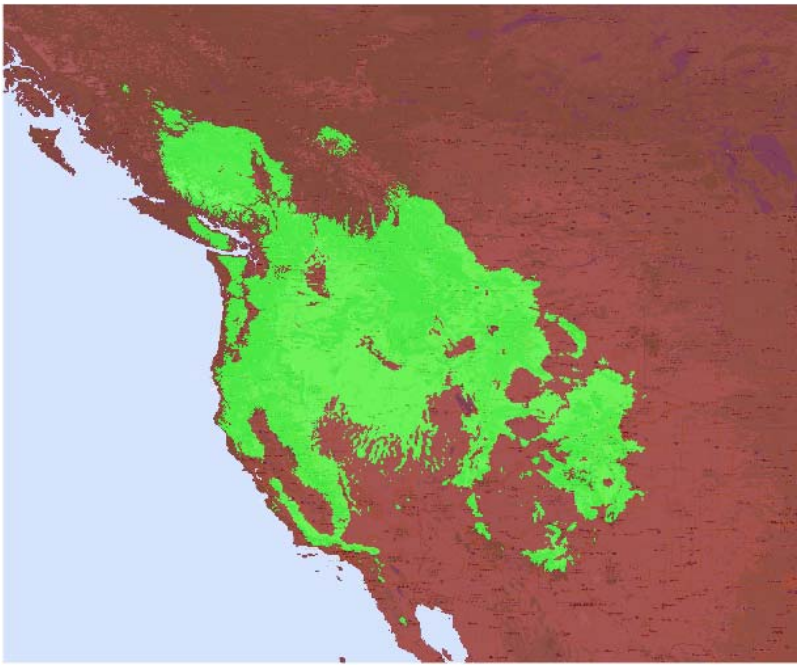
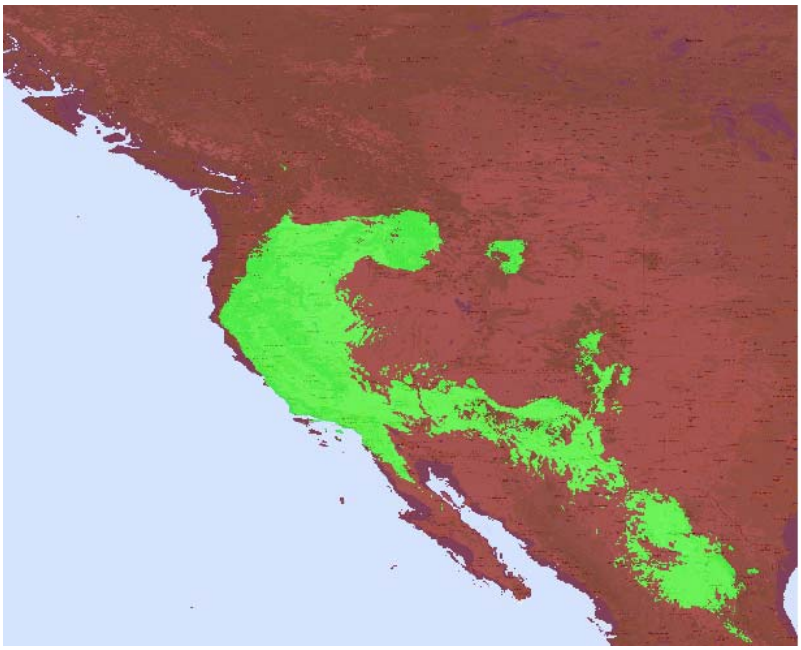



Figure 7