Whole genome analyses reveal weak signatures of population structure and environmentally associated local adaptation in an important North American pollinator, the bumble bee *Bombus vosnesenskii*.

Running title: WGS analysis of *Bombus vosnesenskii*

Sam D. Heraghty*, Jason M. Jackson*, Jeffrey D. Lozier*

*Department of Biological Sciences, The University of Alabama, Tuscaloosa, AL, USA

Corresponding authors: Sam D. Heraghty

Mailing address: Box 870344, University of Alabama, Tuscaloosa, AL, 35487

Phone: 205-348-2754

Email: samd.heraghty@gmail.com

Abstract

Studies of species that experience environmental heterogeneity across their distributions have become an important tool for understanding mechanisms of adaptation and predicting responses to climate change. We examine population structure, demographic history, and environmentally associated genomic variation in *Bombus vosnesenskii*, a common bumble bee in the western U.S.A., using whole genome resequencing of populations distributed across a broad range of latitudes and elevations. We find that *B. vosnesenskii* exhibits minimal population structure and weak isolation by distance, confirming results from previous studies using other molecular marker types. Similarly, demographic analyses with Sequentially Markovian Coalescent (SMC) models suggest that minimal population structure may have persisted since the last interglacial period, with genomes from different parts of the species range showing similar historical effective population size (N_e) trajectories and relatively small fluctuations through time. Redundancy analysis revealed a small amount of genomic variation explained by bioclimatic variables, and environmental association analysis with latent factor mixed modeling (LFMM2) identified few outlier loci that were sparsely distributed throughout the genome. Some outlier loci were in genes with known regulatory relationships, suggesting the possibility of weak selection, although compared to other species examined with similar approaches, evidence for extensive local adaptation signatures in the genome was relatively weak. Overall, results indicate *Bombus vosnesenskii* is an example of a generalist with a high degree of flexibility in its environmental requirements that may ultimately benefit the species under periods of climate change.

Key words: *Bombus*, climate adaptation, genetic variation, PSMC, whole genome sequencing

1 | Introduction

Studying species that experience environmental heterogeneity is a key component in understanding phenomena like local adaptation and predicting species responses to climate change (Hoffmann, Weeks, & Sgro, 2021; Savolainen, Lascoux, & Merilä, 2013; Sears, Riddell, Rusch, & Angilletta, 2019). The degree of spatial and environmental heterogeneity across a species range can have important effects on gene flow and population structure (Y. Li et al., 2017; Manel, Schwartz, Luikart, & Taberlet, 2003). Spatially complex landscapes have features that might constrain gene flow, increasing population genetic structure of inhabitants compared to more homogeneous landscapes (Rahbek et al., 2019; Wang & Singh, 2019) and also harbor substantial abiotic variation that could drive local adaptation in populations from dissimilar environments (Andrews et al., 2022; Antoniou et al., 2023; Capblancq, Fitzpatrick, Bay, Exposito-Alonso, & Keller, 2020; Heraghty, Rahman, Jackson, & Lozier, 2022). Genomic data can be used to understand current and historical population structure as well as genome-environment associations that might signal local adaptation and can thus provide multiple types of information about evolutionary responses to complex landscapes in widespread species.

Although natural selection is expected to shape genetic variation in species distributed over large biogeographic gradients, not all species experience and respond to spatial or environmental heterogeneity in the same way (Gallegos, Hodgins, & Monro, 2023; Hartke et al., 2021; J. M. Jackson et al., 2020). For instance, species capable of long-distance dispersal or characterized by generalist individual phenotypes across a range of environmental conditions might show weak signals of local adaptation (Lenormand, 2002; Räsänen & Hendry, 2008). This phenomenon has been observed in a variety of taxa including marine invertebrates (Lal, Southgate, Jerry, Bosserelle, & Zenger, 2016), butterflies (Melero et al., 2022), and trees (De la

Torre, Wilhite, Neale, & Slotte, 2019). Landscape genomics provides a lens for assessing the challenges of a heterogenous environment by revealing specific environmental features that drive population structure (Capblancq & Forester, 2021), and has also become a widely used tool for detecting signals of environmental adaptation (Capblancq, Morin, Gueguen, Renaud, & Bazin, 2020; De la Torre et al., 2019; Heraghty et al., 2022; J. M. Jackson et al., 2020), especially as whole genome data becomes more easily obtained.

Demographic responses to past climatic fluctuations can provide another important dimension for understanding environmentally associated genomic variation. In widespread species, for example, distinct demographic histories (e.g. bottlenecks) across a distribution can influence on how populations adapt to their contemporary environment, and testing for parallel or dissimilar effective population size (N_e) trajectories through time can provide context for inferences of contemporary population structure or local adaptation (Ahrens et al., 2018). A common tool for inferring N_e is a class of methods known as Sequential Markovian Coalescence (SMC) models (Lozier, Strange, & Heraghty, 2023; Mather, Traves, & Ho, 2020; Nadachowska-Brzyska, Burri, Smeds, & Ellegren, 2016), which can provide estimates of temporal variation in N_e using genomes from single individuals or populations. SMC approaches can reveal how populations have changed in size over tens of thousands of years, and thus reveal whether past climate conditions may have affected genetic variation in similar or distinct ways across a species range (Lozier et al., 2023; Taylor et al., 2021). Such information may also be helpful in interpreting contemporary patterns of environmental association since complex histories could constrain genetic diversity which ultimately may limit the ability of organisms to adapt to their environment (Reed & Frankham, 2003).

In this study, we employ a whole genome resequencing (WGR) approach to evaluate population structure, demographic history, and potential targets of adaptation in a common bumble bee, *Bombus vosnesenskii* Radowski. *Bombus vosnesenskii* is one of the most common bumble bee species in the westernmost parts of North America, with a range extending from southern California, USA through British Columbia, Canada (Cameron et al., 2011; Fraser, Copley, Elle, & Cannings, 2012; J. Koch, Strange, & Williams, 2012; Stephen, 1957). The latitudinal range of *B. vosnesenskii* indicates that the species is capable of inhabiting a broad environmental niche, as well as occurring at a range of elevations from sea level to 2,400 m in elevation (Stephen, 1957). Like many bumble bees, *B. vosnesenskii* provides important pollination services (Greenleaf & Kremen, 2006; Strange, 2015; Velthuis & Doorn, 2006) and is the only native western US bumble bee commercially available for pollination. Thus, in addition to better revealing patterns of evolutionary adaptation to heterogeneous environments in widespread species, improved understanding of regional genomic variation in *B. vosnesenskii* could be of value for breeding or assessing risks associated with bees used in commercial contexts (Lozier & Zayed, 2016), and for understanding potential risks to pollination from range shifts under climate change (H. M. Jackson et al., 2022; Kerr et al., 2012).

Bombus vosnesenskii is among the most well-studied bumble bees in the United States in terms of its geographic range, habitat use, physiology, and ecology (Heinrich & Kammer, 1973; J. Koch et al., 2012; Mola, Miller, O'Rourke, & Williams, 2020a; Pimsler et al., 2020; Stephen, 1957), although no study to date has investigated the species using range-wide whole genome resequencing. Prior information leads to several hypotheses relating to the effects of landscape heterogeneity on genomic variation in *B. vosnesenskii*. Microsatellite and reduced representation (RADseq) single nucleotide polymorphism (SNP) datasets have largely found evidence of

minimal population structure and overall genetic homogeneity at the continental scale (J. M. Jackson et al., 2018; S. Jha, 2015; Lozier, Strange, Stewart, & Cameron, 2011). RADseq also revealed relatively few outlier loci associated with bioclimatic or elevational variables compared to a related narrower-niche montane species, *B. vancouverensis* (J. M. Jackson et al., 2020). Foraging in *B. vosnesenskii* appears uninhibited by certain complex landscapes like forests (Mola et al., 2020a) and can benefit from pulses in floral availability associated with wildfire disturbance (Mola, Miller, O'Rourke, & Williams, 2020b), potentially indicating that movement of reproductive castes, and thus gene flow, could also be unimpeded or facilitated by landscape heterogeneity. Consistent with the hypothesis of *B. vosnesenskii* genetic homogeneity, range-wide morphological analyses have also found little evidence of variation in functional traits like body size or wing loading across latitude or altitude compared to *B. vancouverensis* (Lozier et al., 2021). There is evidence, however, that some spatial or environmental landscape features can impact *B. vosnesenskii* demography by influencing nesting density (Shalene Jha & Kremen, 2013a) and introducing local or regional barriers to gene flow (J. M. Jackson et al., 2018; S. Jha, 2015; Shalene Jha & Kremen, 2013b). Physiological assays have also detected variation in critical thermal minima (CT_{min}) across latitudes and elevation that could indicate adaptive variation associated with local cold temperatures (Pimsler et al., 2020). Overall, *B. vosnesenskii* appears well-suited to occupying heterogenous landscapes, but with some potential for environmental features to shape aspects of the species' evolution that may be clarified using the large genetic marker sets from whole genomes.

Here, we generate WGR data from *B. vosnesenskii* workers sampled from diverse environments across California and Oregon to test several hypotheses relating to population structure and local adaptation. WGR data is a powerful tool for detecting adaptation and can

overcome potential shortcomings of previously used approaches such as reduced representation sequencing (e.g., RADseq) (Jackson et al. 2020), which may not detect all possible targets of selection when linkage blocks are small (Fuentes-Pardo & Ruzzante, 2017), as in bumble bees (Stolle et al., 2011). Whole genome data enables analyses that infer species' demographic histories which can complement contemporary landscape genomics (Beichman, Huerta-Sanchez, & Lohmueller, 2018; de Greef et al., 2022; Iannucci et al., 2021). First, based on prior results of near-panmixia from other genetic markers, we examine population structure and test for any influence of spatial-environmental variation that may not have been apparent in earlier studies. We also use SMC methods to test whether patterns of genetic structure or homogeneity detected in current populations is likewise reflected by divergent or parallel patterns of historical effective population size trajectories across the *B. vosnesenskii* range. Last, we expand upon previous RADseq data by performing genome scans at the fine resolution afforded by WGR data to identify potential signals of environmental adaptation across the *B. vosnesenskii* range.

2 | Materials and Methods

2.1 | Sample Collection, DNA extraction and Sequencing

Bombus vosnesenskii workers (diploid females) were selected for whole genome resequencing from previously collected samples (J. M. Jackson et al., 2018, 2020) that are representative of elevational extremes sampled across a range of latitudes in California and Oregon (36.5°N – 45.3°N latitude and 49m – 2797m above sea level) (Fig 1, Table 1). Our sampling represents a series of six relatively low and high elevation site-pairs nested across a range of latitudes, which should cover the breadth of environmental conditions experienced by *B. vosnesenkii*, such as mean annual temperature from 3°C- 17°C and annual precipitation from 369

mm – 2,177 mm (from WorldClim v2 (Fick & Hijmans, 2017). Briefly, samples were collected at each site via sweep netting, placed on ice for identification (being especially careful to exclude the phenotypically similar *B. caliginosus*), and then placed in 100% ethanol on dry ice, before final storage in ethanol at -80°C. Based on prior estimation of relatedness with RADseq data, selected workers should represent independent colonies. See Jackson et al. 2018 for a more complete description of sampling and the study region.

DNA was extracted using the Qiagen DNeasy Blood and Tissue kit (Hilden, N.R.W., Germany). Some samples were sequenced using whole genome shotgun libraries prepared via the NEBnext Ultra II FS DNA kit (Ipswich, MA, USA) with subsequent 150 bp paired-end sequencing using Illumina NovaSeq 6000 technology (Psomagen, Rockville MD). The remaining sample libraries were prepared by the Genomics and Cell Characterization Core Facility at the University of Oregon and sequenced across three separate lanes of an Illumina Hiseq 4000 instrument.

2.2 | Read mapping, filtering, and variant calling

Sequencing reads were processed with bbduk v37.32 (Bushnell 2020) to remove adaptors, trim low quality bases, and remove short reads (ktrim=r k=23 mink=11 hdist=1 tpe tbo ftm=5 qtrim=rl trimq=10 minlen=25). Read quality was evaluated using FastQC v0.11.5 (Andrews 2010) before being mapped to the *B. vosnesenskii* reference genome (NCBI RefSeq ID: GCF_011952255.1) (Heraghty et al., 2020) using BWA mem v0.7.15-r1140 (H. Li & Durbin, 2009). Samtools v1.10 (H. Li et al., 2009) was used to convert the SAM files to BAM files and Picard tools v2.20.4 (Broad Institute 2019) was used to sort, mark duplicates, and index the BAM files. Single nucleotide polymorphisms (SNPs) were called using freebayes v1.3.2 (Garrison & Marth, 2012) and filtered following Heraghty et al. (2022). An initial round of

filtering was performed to remove low-quality variants using vcftools v0.1.13 (Danecek et al., 2011) with the following flags: --remove-indels --min-alleles 2 --max-alleles 2 --minQ 20 --minDP 4 --max-missing 0.75. Subsequent filtering was designed to remove SNPs with unusually high coverage (> 2x average coverage), excess heterozygosity (--hardy flag in vcftools), located on small scaffolds (< 100 kb in size), and with a minor allele frequency (MAF) ≤ 0.05 to filter SNPs that might have arouse from repeat regions or sequencing artefacts and limit the effects of low frequency variants. A final round of filtering was done to remove differences that might arise from using data generated on different sequencing platforms, as recommended by (De-kayne et al., 2021), using a higher SNP quality filter (min Q of 30 and min GQ of 20).

2.3 | Environmental Variable Selection and Population Structure

Environmental variables (19 Bioclim variables) at a 0.5 arcminute resolution were obtained from WorldClim v2 (Fick & Hijmans, 2017). To reduce correlation between variables we used an item clustering analysis (*iclust* function with default settings) from the *psych* v.2.1.9 package (Revelle 2020) in R v 4.2.0 (R Core Team 2022) and retained a single variable per correlated variable cluster. Retained variables were: BIO1 – annual mean temperature, BIO3 – isothermality, and BIO12 – annual precipitation. Elevation was also included as a variable of interest despite being partly correlated with other environmental variables as it may capture unique environmental information that could be relevant for bumble bees such as air density and oxygen availability (Cheviron & Brumfield, 2012; Dillon, 2006; Heraghty et al., 2022).

To assess population structure, the fully filtered vcf file was converted into a genlight object via the *vcfR2genlight* function in the *vcfR* v1.12.0 package in R (Knaus & Grünwald, 2017). Global F_{ST} (Weir & Cockerham 1984 method (Weir & Cockerham, 1984)) was calculated using *SNPRelate* 1.30.1 (Zheng et al., 2012) in R. Pairwise F_{ST} was calculated between all

sampling localities using the stamppFst function from the *dartR* v1.9.9.1 package in R (Gruber, Unmack, Berry, & Georges, 2018). A geographic distance matrix for all sampling points was created using the *distm* function from the *geosphere* v1.5-14 package in R using coordinates of each site. A Mantel test (statistical testing using 1000 permutations) was performed on pairwise F_{ST} and geographic distance matrices to test for signatures of isolation by distance (Slatkin, 1993) using the mantel function from the *vegan* v2.5-7 package in R (Oksanen et al 2020). To visualize population structure, the *gl.pcoa* function from the *dartR* v1.9.9.1 package in R (Gruber et al., 2018) was used to conduct a Pearson PCA. Nucleotide diversity (π) was calculated for each population using the – site-pi flag in vcftools v0.1.13 (Danecek et al., 2011).

We also evaluated the relative contribution of spatial and environmental factors on the partitioning of genetic variation in *B. vosnesenskii* using partial redundancy analysis (pRDA) (Capblancq & Forester, 2021). A full RDA model was created with *rda* function in the *vegan* v2.5-7 package in R using the selected environmental variables BIO1, BIO3, BIO12, elevation, and using latitude as a proxy for geographic effects. The pRDA model accounting for environmental effects (hereafter referred to as "Environmental pRDA") used the following model (BIO1 + BIO3 + BIO12 + elevation | latitude) and the pRDA model accounting for geographic distance effects (hereafter referred to as "Geography pRDA") used the following model (latitude | BIO1 + BIO3 + BIO12 + elevation). We also examined models with just environmental variables or latitude separately (i.e. without the partial effects), which we refer to as the "Environmental RDA" and "Geography RDA", respectively. All models were assessed for significance using the *anova* function in R following the protocol from (Capblancq & Forester, 2021).

2.4 | Demographic Inference

To test whether populations from different parts of the *B. vosnesenskii* range exhibit parallel or distinct recent evolutionary histories, we inferred historical N_e trends using SMC analyses of whole genomes (H. Li & Durbin, 2011; Mather et al., 2020). SMC methods use the distribution of heterozygous and homozygous sites within loci to calculate coalescence rates that provide insight into past demographic events (Beichman et al., 2018; H. Li & Durbin, 2011). This approach is a common tool used when WGR data are available, even for small numbers of individuals, as methods can make use of even single diploid genomes (Mather et al., 2020; Nadachowska-Brzyska et al., 2016). We used PSMC (H. Li & Durbin, 2011) for demographic analyses, which is the original but still widely employed (Lozier et al., 2023; Morin et al., 2021; Nadachowska-Brzyska et al., 2016; Patil et al., 2021; Skovrind et al., 2021) SMC method, and generally followed methods in Lozier et al. (2023). PSMC uses whole genome data from single individual diploid samples to estimate coalescence rates to track N_e history. Because SMC methods generally perform best with high coverage and long scaffolds (Mather et al., 2020), we only ran PSMC for *B. vosnesenskii* workers with ≥18x mean coverage (Nadachowska-Brzyska et al., 2016) and for reference genome scaffolds ≥ 500kb in length (Gower et al., 2018). Following methods described in the PSMC manual, the PSMC input files were constructed from the filtered BAM files using Samtools v1.10 mpileup with base and mapping quality set to 30. The consensus sequence was generated using bcftools call, converted to fastq format with Samtools vcfutils.pl vcf2fq, and converted to the psmcfa format for demographic inference using the psmc program. The default p parameter describing temporal intervals of “4+25*2+4+6" was used, which is typically useful for a range of taxa (Patil & Vijay, 2021). Results were plotted using the plot_psmc.py script provided with PSMC. To convert the results into real time, we used the

direct mutation rate estimate of 3.6×10^{-9} per site per year from the bumble bee *Bombus terrestris* (Liu et al., 2017) and a generation time of 1 year.

2.5 | Identifying environmentally associated genomic loci

Environmental Association Analysis (EAA) is often used to detect genetic variants with unusual allele frequency distributions that suggest a possible role in local adaption (e.g., "outlier" loci)(Ahrens et al., 2018). Here we use a common EAA method, latent factor mixed modeling with *LFMM2* from the *LEA* v3.0.0. R package (Gain & François, 2021), which uses a least-squares approach to identify SNPs with a significant association with a given variable while controlling for population structure. The likely number of population clusters (k) used for background population structure control was determined with the *sMNF* function from the *LEA* v3.0.0. R package (Gain & François, 2021), with k representing the value that had the smallest cross-entropy selected from a range of $k = 1 – 10$. To correct for multiple testing with the large number of genetic markers, raw p-values from LFMM2 were transformed into q-values using the *q-value* v2.20.0 R package (Storey et al. 2022). SNPs were considered significantly associated with a given variable at a threshold of $q \leq 0.05$.

3 | Results

3.1 | Data summary

A total of 111 individuals were sequenced to an average of 21,053,016 reads per individual and variant calling yielded an initial dataset of 22,105,684 SNPs. The final dataset consisted of 1,091,021 SNPs at a 5% minor allele frequency with an average coverage of 13.39x SNP^{-1} $individual^{-1}$ and < 25% missing data SNP^{-1} (98.02% complete data across all samples).

3.2 | Population structure and demographic history

Consistent with prior results in *B. vosnesenskii*, there was minimal population structure, with global F_{ST} = 0.001. The PCA also revealed little population structure and weak spatial clustering, but with a trend towards individual loadings related to the latitude at which the samples were collected (Fig 2a). This trend likely reflects the weak but significant isolation by distance at a range-wide scale (Mantel r: 0.374, P = 0.006; Fig 2b). The RDA models also supported minimal population structure, with all models explaining little variation, with the full model explaining only ~5% of the variation (Table 2). Neither of the partial RDA models were significant, but all three RDA models without partial effects (Full model, Environment RDA, and Geography RDA) were significant. The Geography RDA indicates that most variation is likely attributable to weak isolation by distance as suggested by above analyses, with samples clustering together by approximate sampling latitude along the x-axis (Fig. 2), but with a significant amount of unexplained variation (captured by the y-axis).

Based on the major latitudinal pattern in population structure we tested for an effect of latitude on genetic variation, and found a significant decline in π with latitude (linear regression, $F_{1,10}$ = 9.465, P = 0.012, Fig 3a.). Despite this latitudinal trend, however, PSMC indicated similar N_e trajectories across samples that indicate a shared demographic history for genomes across the species range, consistent with the minimal overall population structure in the species (Fig 3b). Inferred population sizes were fairly stable over time, being smallest at the oldest time scales (before the last interglacial period), changing little from the Last Interglacial period to the Last Glacial Maxima aside from a modest increase just before the LGM and decline during or immediately following the LGM (Fig 3). All populations showed some degree of recent increases in N_e, although the variation among genomes in the most recent time interval is likely

in part driven by challenges to modeling the most recent time segments with SMC methods (Beichman et al., 2018).

3.3 | Environmentally associated outlier SNPs

Only a single population cluster was identified as optimal with sMNF based on the lowest cross entropy, so *k*=1 was used for LFMM2. A total of 81 loci in 72 genes were found to be significantly associated with at least one environmental variable. Significant SNPs were sparsely spread across the genome with few genes having multiple outlier SNPs and no clear peaks of association (Table 2, Fig 4). Most outlier SNPs fell within genes (60 of 81) and were principally associated with annual precipitation (n = 66, BIO12), with smaller numbers of SNPs associated with annual mean temperature (n = 4, BIO1) and isothermality (n = 16, BIO3). No outliers were associated with elevation. BIO3 and BIO12 shared 9 outlier SNPs, but BIO1 outliers were all unique.

Several notable genes from the LFMM2 results included LOC117231927 (*calcium/calmodulin-dependent protein kinase type II alpha chain*, homologous to *CaMKII* in *D. melanogaster*) which had the most outlier SNPs (n=3; two with BIO3 and one with BIO12). This gene has roles in neuronal growth, calcium signaling, as well as learning associated with appetite (Gillespie & Hodge, 2013). There were also several other genes with functions relating to neuronal/neuromuscular function such as LOC117231202 (*hemicentin-1*, homologous to *nrm* in *D. melanogaster*) (Kania, Han, Kim, & Bellen, 1993) and LOC117241676 (*bone morphogenetic protein 1-like*, homologous to *tok* in *D. melanogaster*) (Serpe, Ralston, Blair, & O'Connor, 2005). Interestingly, there were also several outliers in gene pairs with some form of known interaction; LOC117231348 (*potassium voltage-gated channel protein eag*, homologous to *eag* in *D. melanogaster*) is regulated by LOC117231927 (*CaMKII* in *D. melanogaster*) (Bronk et al.,

2018) and LOC117233891 (*dorsal-ventral patterning protein sog*, homologous to *sog* in *D. melanogaster*) has its product cleaved by LOC117241676 (*tok* in *D. melanogaster*) (Serpe et al., 2005).

4 | Discussion

Whole-genome resequencing of *B. vosnesenskii* across large latitude and altitude gradients in the western USA confirmed the presence of weak overall population structure and patterns of environmentally associated genomic variation that match prior results with other genetic markers (J. M. Jackson et al., 2018, 2020; Lozier et al., 2011). However, WGR data did provide some new insights, indicating that populations appear to have shared fairly concordant demographic histories (reflected by N_e) over time, suggesting that the minimal population structure apparent today has persisted for thousands of years. Gene flow appears extensive across the study area and may be a likely explanation for the relatively small amount of genomic variation associated with the environmental variables and homogenous evolutionary histories reflected in genomes from different regions. Although there were relatively few and sparsely distributed environmentally associated outlier SNPs, there were some notable patterns highlighting the potential importance of ion homeostasis and neuromuscular function that is consistent with other bumble bee studies (Heraghty et al., 2022; Huml et al., 2023; Sun et al., 2020). Overall, the results suggest that *B. vosnesenskii* may be an example of a species with minimal population structure impeding adaptation or with weak selection over gene networks that will require much more extensive sampling to resolve.

Analysis of population structure suggests *Bombus vosnesenskii* is nearly but not completely panmictic, with weak isolation by distance and much of the explainable genetic variation associated with latitudinal separation, which is consistent with previous studies

(Cameron et al., 2011; J. M. Jackson et al., 2018; Lozier et al., 2011). We were interested in determining whether the greater resolution afforded by WGR data would provide additional evidence for gene flow barriers across the sampled regions, but no major differences emerged compared to prior studies. Low genetic differentiation over large spatial scales, is often observed in bumble bee species in the absence of obvious physical dispersal barriers (Christmas et al., 2022; Heraghty et al., 2022; J. B. Koch, Looney, Sheppard, & Strange, 2017; Lozier et al., 2011). Although long distance dispersal of individual reproductives (gynes or drones) may be rare (Williams et al. 2022), stepping stone dispersal through suitable habitat is likely in bumble bees (Williams et al., 2022; Williams, Lobo, & Meseguer, 2018) and would contribute to weak genetic structure when species inhabit continuous geographic ranges. The isolation by distance observed here is consistent with such a model of stepping-stone population structure. *Bombus vosnesenskii* is abundant and can be found at low and high elevations throughout California and Oregon, and thus its range in these regions is composed almost entirely of contiguous suitable habitat. Not all bumble bees show this pattern, however, and data from other species in the region have higher levels of population structure. For example, *B. vancouverensis*, a species in the same subgenus (Williams, Cameron, Hines, Cederberg, & Rasmont, 2008) with a similar latitudinal range, but more narrow higher elevation niche at any given latitude (J. Koch et al., 2012), has much greater range-wide F_{ST} based on whole genome and RADseq data (Heraghty et al., 2022; J. M. Jackson et al., 2018). Importantly, however, our sampling did not include strongly isolated populations, such as those on islands, or populations from many coastal regions, which may harbor unique genetic variation (S. Jha, 2015). Incorporating whole genome data from such populations may provide evidence of the potential for greater population structure in *B. vosnesenskii*.

PSMC historical demographic analyses were also consistent with contemporary patterns of population structure. All individuals showed highly similar historical N_e trajectories, likely indicating a relatively high degree of homogeneous population structure through much of the recent past. Again, results in *B. vosnesenskii* can be compared to *B. vancouverensis*, where SMC analyses inferred markedly different bottleneck and expansion magnitudes across the species range that were linked to major climate fluctuations and consequent changes in suitable habitat areas over time that may have promoted genetic divergence (Lozier et al., 2023). For *B. vosnesenskii*, there were relatively small signatures of genomic bottlenecks or other major N_e fluctuation associated with glacial-interglacial periods, possibly suggesting that flexibility in environmental requirements may facilitate population stability and near-panmixia over time. This would also be consistent with the minimal environmentally associated genomic variation observed here and in earlier RADseq results (Jackson et al. 2020). Despite, the minimal changes in N_e associated with changing historical conditions, there is subtle yet significant decline in nucleotide diversity with latitude, a pattern that was previously observed in microsatellite data (Lozier et al. 2011). Such patterns may provide some evidence for post-glacial northward expansion that may be shaping variation in *B. vosnesenskii*, despite the similarity in PSMC trajectories among samples, but would also be consistent with the subtle differences in the strength of very recent N_e increases observed in several of the genomes (Fig 3b).

4.1 | Environmentally associated outliers

The lack of major population structure and parallel N_e histories across the *B. vosnesenskii* range is mirrored by a lack of clear influence of environmental variables on overall population structure or on allele frequencies at individual loci that might be produced by local adaptation. Interestingly, most of the detected outliers were associated with precipitation, with smaller

numbers of loci associated with other variables, and none with elevation. Although, the outlier loci were sparsely spread throughout the genome (Table 2, Fig 4), there were several interesting cases where outlier-containing genes interacted, possibly suggesting the potential for weak selection on members of gene networks. This is supported by several instances where outlier genes have known regulatory interactions (e.g., LOC117231927 and LOC117231348 or LOC117233891 and LOC117241676). The overall effect of selection across networks can be evaluated by looking at the roles of the individual genes with detected outliers, such as the number of outlier-containing genes related to neuromuscular function.

Given the broad latitudinal and elevational range sampled for this study, the scarcity of outlier SNPs associated with thermal variables (BIO1 and BIO3) in these whole genomes was somewhat unexpected but does match the less dense RADseq SNP set results from Jackson et al. (2020) that included even more individuals and populations. One possible explanation is that *B. vosnesenskii* is a noted elevation generalist (J. Koch et al., 2012; Lozier et al., 2021; Thorp, Horning Jr., & Dunning, 1983), which, combined with high gene flow, may not produce strong or consistent signatures adaptations for elevation or temperature stressors like isothermality across the species range. The scarcity of outliers tied to elevation or temperature is similar to the absence of correlations between relevant functional morphological traits (e.g., body size and wing loading) with temperature or elevation observed in this species (Lozier et al., 2021). However, a lack of genome regions that show clear associations with temperature is still surprising given recent work on thermal tolerance physiology in *B. vosnesenskii* (Pimsler et al 2020) that found populations from cold, high elevation habitats had significantly lower critical thermal minima (CT_{min}) compared to populations from warmer localities (Pimsler et al., 2020). It is possible that selection is weak on individual genes associated with a complex trait like thermal

tolerance (Yang, Crossley, Schrader, & Dubovskiy, 2022), that different genes may be targeted by selection in different regions (Yeaman, 2022), or that another non sequence-based mechanism (e.g., epigenetic) could facilitate variation in cold tolerance among populations (Mccaw, Stevenson, & Lancaster, 2020). Detecting selection under the high levels of gene flow that seems prevalent in *B. vosnesenskii* will also be challenging, and more independent studies, additional sampling, and experimentation will be needed to fully understand any mechanisms of genetic adaptation in this species.

In addition to the general challenges of sample size when selection is weak on individual genes, there are several other important considerations when evaluating adaptive potential of individual outliers. First, environmental variation is spatially correlated, so models need to consider species demography (Hoban et al., 2016). Here, population structure is weak and the statistical method employed (LFMM2) accounts for population structure, which may limit major effects here, but at the same time LFMM2 is a relatively conservative method (Luo, Tang, Schoville, & Zhu, 2021), which might contribute to reduced power to detect outliers with weak environmental associations with our modest sample size (Ahrens et al., 2018). Second, correlation between environmental variables can complicate linking environmentally associated loci to driving pressures. For instance, annual precipitation, which had the most outlier loci, could be associated with a number of different potential pressures such as desiccation tolerance, changes in biotic interactions (e.g., flowering phenology), or a number of other factors (see (Chown, Sørensen, & Terblanche, 2011) for a more thorough discussion). The presence of outliers that are common to multiple environmental variables (e.g., BIO12 and BIO3) highlights this challenge, and indicates that even a strong signal of possible selection could be associated with multiple (even unsampled) variables.

The outlier signatures in *B. vosnesesnskii* can be summarized by once again comparing to *B. vancouverensis* from similar geographic regions (Heraghty et al., 2022; J. M. Jackson et al., 2020). From both whole genome resequencing and RADseq SNPs, *B. vancouverensis* has much stronger signals of environmental association across its genome, with several large peaks of genomic divergence in genes with putatively relevant functions (Heraghty et al., 2022). Thus, in comparison to *B. vancouverensis*, *B. vosnesenskii* has less substantial population structure, a lack of population-specific demographic histories from SMC analyses, and much sparser evidence for local adaptation. That said, our results here do reveal a few parallels between these species with respect to the functions of genes with identified outlier SNPs. For instance, the gene LOC117231202 (*hemicentin-1*) was identified as containing outlier SNPs in both species. This gene reflects the shared trend towards genes involved in neuromuscular function (Kania et al., 1993) in outlier sets for both species, which has also been observed in outlier analysis of other bumble bees (Huml et al., 2023). Although elevation and thermal variables were not as associated with outliers in *B. vosnesenskii*, given correlations among environmental variables discussed above, such processes could nonetheless reflect shared selection pressures on traits such as thermal adaptation and flight. Bumble bee flight is crucial for foraging, dispersal, and overall colony fitness (Mola et al., 2020a), while the capacity for thermogenesis by shivering their flight muscles is well-recognized in *Bombus* (Heinrich 1975). These overlaps may suggest there may be some common tools important for bumble bees that occupy spatial-environmental heterogeneity across their ranges, even if such species exhibit different intensities in evidence for selection, which would be well worth additional research.

4.2 | Broader Implications

As one of the most common bees in the western U.S.A., *B. vosnesenskii* is an important pollinator in natural ecosystems as well as agricultural systems in California and Oregon (Fisher, Watrous, Williams, Richardson, & Woodard, 2022; Greenleaf & Kremen, 2006). Thus, ensuring the health of *B. vosnesenskii* populations in a changing environment is of great interest. Our results contribute to the current understanding of how *B. vosnesenskii* may fare under climate change. Contemporary analysis of *B. vosnesenskii* suggests that this species has remained stable over recent decades (Cameron et al., 2011). The historical demographic inferences from SMC modeling suggest that populations sizes across our study region have likewise been minimally affected by the change in environmental conditions from the LIG to the LGM (Fig 3). During the LIG, conditions were warmer than present by ~1.2 °C relative to the 1998-2016 average (Lüning & Vahrenholt, 2017) whereas LGM conditions were colder by ~6 °C (Von Deimling, Ganopolski, Held, & Rahmstorf, 2006). The species' success under warmer conditions is particularly encouraging since it may indicate stability in future climates (IPCC 2021). Recent work suggests that higher temperatures may even be beneficial to *B. vosnesenskii* and a possible factor in range expansions into British Columbia (Fraser et al., 2012; H. M. Jackson et al., 2022), with *B. vosnesenskii* representing a possible "climate winner", at least to date (Jackson et al. 2022). In addition to temperature, precipitation patterns will also shift in future climate change (IPCC 2021). Our results suggest that annual precipitation is associated with most genomic selection signatures, and prior results found precipitation to be a key predictor of *B. vosnesenskii* range limits in species distribution models (Jackson et al. 2018). Although recent work has indicated that changes in precipitation are less likely to have an impact on future distributions of North American bumble bees, there are some exceptions (H. M. Jackson et al., 2022), and our

genomic data indicate that water limitation may represent an important consideration under future climates that may warrant further study.

In summary, we have confirmed previous findings indicating there is relatively little population structure in *B. vosnesenskii* (J. M. Jackson et al., 2018) and identified relatively few environmental associated outliers which could underlie adaptation in this species. The lack of outliers could be associated with a gene swamping effect resulting from the high observed gene flow (Lenormand, 2002), although contradicts data that suggests biogeographic variation in key thermal performance metrics (Pimsler et al. 2021). For the environmentally associated outliers that were identified, future work should be conducted to evaluate their potential role in local adaptation. Additionally, future work should be aimed at teasing apart the environmental variation that is captured by precipitation (BIO12) and how that might impact adaptation and future species distributions, given changes in precipitation patterns forecasted under climate change. Finally, more work is needed to examine other avenues for adaptation in this species, such as epigenetic modifications which may explain observed differences in cold tolerance across the range.

Acknowledgements

We thank the university of Alabama College of Arts and Sciences and the National Science Foundation (DEB-1457645 and URoL 1921585 to J.D.L.) for support related to this project. We also thank the other members of the Lozier lab for useful feedback and discussion.

References:

Ahrens, C. W., Rymer, P. D., Stow, A., Bragg, J., Dillon, S., Umbers, K. D. L., & Dudaniec, R.

Y. (2018). The search for loci under selection: trends, biases and progress. *Molecular Ecology*, *27*(6), 1342–1356. https://doi.org/10.1111/mec.14549

Andrews, K. R., Seaborn, T., Egan, J. P., Fagnan, M. W., New, D. D., Chen, Z., … Narum, S. R. (2022). Whole genome resequencing identifies local adaptation associated with environmental variation for redband trout. *Molecular Ecology*, *32*(4), 800–818. https://doi.org/10.1111/mec.16810

Andrews, S. 2010. FastQC: A quality control tool for high throughput sequence data http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Antoniou, A., Manousaki, T., Ram, F., Cariani, A., Cannas, R., Kasapidis, P., … Ciencies, I. De. (2023). Sardines at a junction: seascape genomics reveals ecological and oceanographic drivers of variation in the NW Mediterranean Sea. *Molecular Ecology*, *0*, 1–21. https://doi.org/10.1111/mec.16840

Beichman, A. C., Huerta-Sanchez, E., & Lohmueller, K. E. (2018). Using genomic data to infer historic population dynamics of nonmodel organisms. *Annual Review of Ecology, Evolution, and Systematics*, *49*, 433–456. https://doi.org/10.1146/annurev-ecolsys-110617-062431

Broad Institute. 2019. http://broadinstitute.github.io/picard/

Bronk, P., Kuklin, E. A., Gorur-Shandilya, S., Liu, C., Wiggin, T. D., Reed, M. L., … Griffith, L. C. (2018). Regulation of Eag by Ca2+/calmodulin controls presynaptic excitability in *Drosophila*. *Journal of Neurophysiology*, *119*(5), 1665–1680. https://doi.org/10.1152/jn.00820.2017

Bushnell, B. 2020. BBMap. http://sourceforge.net/ projects/bbmap/

Cameron, S. A., Lozier, J. D., Strange, J. P., Koch, J. B., Cordes, N., Solter, L. F., & Griswold,

T. L. (2011). Patterns of widespread decline in North American bumble bees. *Proceedings of the National Academy of Sciences*, *108*(2), 662–667. https://doi.org/10.1073/pnas.1014743108

Capblancq, T., Fitzpatrick, M. C., Bay, R. A., Exposito-Alonso, M., & Keller, S. R. (2020). Genomic prediction of (mal)adaptation across current and future climatic landscapes. *Annual Review of Ecology, Evolution, and Systematics*, *51*, 245–269. https://doi.org/10.1146/annurev-ecolsys-020720-042553

Capblancq, T., & Forester, B. R. (2021). Redundancy analysis: A Swiss Army Knife for landscape genomics. *Methods in Ecology and Evolution*, *12*(12), 2298–2309. https://doi.org/10.1111/2041-210X.13722

Capblancq, T., Morin, X., Gueguen, M., Renaud, J., & Bazin, E. (2020). Climate-associated genetic variation in *Fagus sylvatica* and potential responses to climate change in the French Alps. *The Journal of Evolutionary Biology*, *33*(6), 783–796. https://doi.org/10.1111/jeb.13610

Cheviron, Z. A., & Brumfield, R. T. (2012). Genomic insights into adaptation to high-altitude environments. *Heredity*, *108*(4), 354–361. https://doi.org/10.1038/hdy.2011.85

Chown, S. L., Sørensen, J. G., & Terblanche, J. S. (2011). Water loss in insects: An environmental change perspective. *Journal of Insect Physiology*, *57*(8), 1070–1084. https://doi.org/10.1016/j.jinsphys.2011.05.004

Christmas, M. J., Jones, J. C., Olsson, A., Wallerman, O., Bunikis, I., Kierczak, M., … Webster, M. T. (2022). A genomic and morphometric analysis of alpine bumblebees: Ongoing reductions in tongue length but no clear genetic component. *Molecular Ecology*, *31*(4), 1111–1127. https://doi.org/10.1111/mec.16291

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

De-kayne, R., Frei, D., Greenway, R., Mendes, S. L., Retel, C., & Feulner, P. G. D. (2021). Sequencing platform shifts provide opportunities but pose challenges for combining genomic data sets. *Molecular Ecology Resources*, *21*(3), 653–660. https://doi.org/10.1111/1755-0998.13309

de Greef, E., Einfeldt, A. L., Miller, P. J. O., Ferguson, S. H., Garroway, C. J., Lefort, K. J., … Feyrer, L. J. (2022). Genomics reveal population structure, evolutionary history, and signatures of selection in the northern bottlenose whale, *Hyperoodon ampullatus*. *Molecular Ecology*, *31*(19), 4919–4931. https://doi.org/10.1111/mec.16643

De la Torre, A. R., Wilhite, B., Neale, D. B., & Slotte, T. (2019). Environmental genome-wide association reveals climate adaptation is shaped by subtle to moderate allele frequency shifts in loblolly pine. *Genome Biology and Evolution*, *11*(10), 2976–2989. https://doi.org/10.1093/gbe/evz220

Dillon, M. E. (2006). Into thin air: Physiology and evolution of alpine insects. *Integrative and Comparative Biology*, *46*(1), 49–61. https://doi.org/10.1093/icb/icj007

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

Fisher, K., Watrous, K. M., Williams, N. M., Richardson, L. L., & Woodard, S. H. (2022). A contemporary survey of bumble bee diversity across the state of California. *Ecology and Evolution*, *12*(3), 1–12. https://doi.org/10.1002/ece3.8505

Fraser, D. F., Copley, C. R., Elle, E., & Cannings, R. A. (2012). Changes in the status and distribution of the yellow-faced bumble bee (*Bombus vosnesenskii*) in British Columbia. *Journal of the Entomological Society of British Columbia*, *109*(250), 31–37.

Fuentes-Pardo, A. P., & Ruzzante, D. E. (2017). Whole-genome sequencing approaches for conservation biology: Advantages, limitations and practical recommendations. *Molecular Ecology*, *26*(20), 5369–5406. https://doi.org/10.1111/mec.14264

Gain, C., & François, O. (2021). LEA 3: Factor models in population genetics and ecological genomics with R. *Molecular Ecology Resources*, *21*(8), 2738–2748. https://doi.org/10.1111/1755-0998.13366

Gallegos, C., Hodgins, K. A., & Monro, K. (2023). Climate adaptation and vulnerability of foundation species in a global change hotspot. *Molecular Ecology*, *00*, 1–15. https://doi.org/10.1111/mec.16848

Garrison, E., & Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. *ArXiv*, 1–9. Retrieved from http://arxiv.org/abs/1207.3907

Gillespie, J. M., & Hodge, J. J. L. (2013). CASK regulates CaMKII autophosphorylation in neuronal growth, calcium signaling, and learning. *Frontiers in Molecular Neuroscience*, *6*, 1–14. https://doi.org/10.3389/fnmol.2013.00027

Gower, G., Tuke, J., Rohrlach, A. B., Soubrier, J., Llamas, B., Bean, N., & Cooper, A. (2018). Population size history from short genomic scaffolds: How short is too short? *BioRxiv*, 1–11. https://doi.org/10.1101/382036

Greenleaf, S. S., & Kremen, C. (2006). Wild bee species increase tomato production and respond differently to surrounding land use in Northern California. *Biological Conservation*, *133*(1), 81–87. https://doi.org/10.1016/j.biocon.2006.05.025

Gruber, B., Unmack, P. J., Berry, O. F., & Georges, A. (2018). dartr: An r package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Molecular Ecology Resources*, *18*(3), 691–699. https://doi.org/10.1111/1755-0998.12745

Hartke, J., Waldvogel, A., Sprenger, P. P., Schmitt, T., Menzel, F., Pfenninger, M., & Feldmeyer, B. (2021). Little parallelism in genomic signatures of local adaptation in two sympatric, cryptic sister species. *Journal of Evolutionary Biology*, *34*(6), 937–952. https://doi.org/10.1111/jeb.13742

Heinrich, B., & Kammer, A. E. (1973). Activation of the fibrillar muscles in the bumblebee during warm up, stabilization of thoracic temperature and flight. *Journal of Experimental Biology*, *58*(3), 677–688. https://doi.org/10.1242/jeb.58.3.677

Heraghty, S. D., Rahman, S. R., Jackson, J. M., & Lozier, J. D. (2022). Whole genome sequencing reveals the structure of environment-associated divergence in a broadly distributed montane bumble bee, *Bombus vancouverensis*. *Insect Systematics and Diversity*, *6*(5), 1–17.

Heraghty, S. D., Sutton, J. M., Pimsler, M. L., Fierst, J. L., Strange, J. P., & Lozier, J. D. (2020). De novo genome assemblies for three north american bumble bee species: *Bombus bifarius*, *Bombus vancouverensis*, and *Bombus vosnesenskii*. *Genes|Genomes|Genetics*, *10*(8), 2585–2592. https://doi.org/10.1534/g3.120.401437

Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., … Whitlock, M. C. (2016). Finding the genomic basis of local adaptation: Pitfalls, practical solutions, and future directions. *American Naturalist*, *188*(4), 379–397. https://doi.org/10.1086/688018

Hoffmann, A. A., Weeks, A. R., & Sgro, C. M. (2021). Opportunities and challenges in assessing

climate change vulnerability through genomics. *Cell*, *184*(6), 1420–1425. https://doi.org/10.1016/j.cell.2021.02.006

Huml, J. V, Ellis, J. S., Rustage, S., Billington, R., Knight, M. E., & Brown, M. J. F. (2023). The tragedy of the common? A comparative population genomic study of two bumblebee species. *Insect Conservation and Diversity*, *00*, 1–20. https://doi.org/10.1111/icad.12626

Iannucci, A., Benazzo, A., Natali, C., Arida, E. A., Zein, M. S. A., Jessop, T. S., … Ciofi, C. (2021). Population structure, genomic diversity and demographic history of Komodo dragons inferred from whole-genome sequencing. *Molecular Ecology*, *30*(23), 6309–6324. https://doi.org/10.1111/mec.16121

IPCC, 2021: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*[Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, In press, doi:10.1017/9781009157896.

Jackson, H. M., Johnson, S. A., Morandin, L. A., Richardson, L. L., Guzman, L. M., & M'Gonigle, L. K. (2022). Climate change winners and losers among North American bumblebees. *Biology Letters*, *18*(6). https://doi.org/10.1098/rsbl.2021.0551

Jackson, J. M., Pimsler, M. L., Oyen, K. J., Koch-Uhuad, J. B., Herndon, J. D., Strange, J. P., … Lozier, J. D. (2018). Distance, elevation and environment as drivers of diversity and divergence in bumble bees across latitude and altitude. *Molecular Ecology*, *27*(14), 2926–2942. https://doi.org/10.1111/mec.14735

Jackson, J. M., Pimsler, M. L., Oyen, K. J., Strange, J. P., Dillon, M. E., & Lozier, J. D. (2020). Local adaptation across a complex bioclimatic landscape in two montane bumble bee species. *Molecular Ecology*, *29*(5), 920– 939. https://doi.org/10.1111/mec.15376

Jha, S. (2015). Contemporary human-altered landscapes and oceanic barriers reduce bumble bee gene flow. *Molecular Ecology*, *24*(5), 993–1006. https://doi.org/10.1111/mec.13090

Jha, Shalene, & Kremen, C. (2013a). Resource diversity and landscape-level homogeneity drive native bee foraging. *Proceedings of the National Academy of Sciences*, *110*(2), 555–558. https://doi.org/10.1073/pnas.1208682110

Jha, Shalene, & Kremen, C. (2013b). Urban land use limits regional bumble bee gene flow. *Molecular Ecology*, *22*(9), 2483–2495. https://doi.org/10.1111/mec.12275

Kania, A., Han, P. L., Kim, Y. T., & Bellen, H. (1993). Neuromusculin, a *Drosophila* gene expressed in peripheral neuronal precursors and muscles, encodes a cell adhesion molecule. *Neuron*, *11*(4), 673–687. https://doi.org/10.1016/0896-6273(93)90078-6

Kerr, J. T., Pindar, A., Galpern, P., Packer, L., Potts, S. G., Roberts, S. M., … Gall, L. F. (2012). Climate change impacts on bumblebees converge across continents. *Science*, *349*(6244), 177–180. https://doi.org/10.1126/science.aaa7031

Knaus, B. J., & Grünwald, N. J. (2017). 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

Koch, J. B., Looney, C., Sheppard, W. S., & Strange, J. P. (2017). Patterns of population genetic structure and diversity across bumble bee communities in the Pacific Northwest. *Conservation Genetics*, *18*(3), 507–520. https://doi.org/10.1007/s10592-017-0944-8

Koch, J., Strange, J., & Williams, P. (2012). Bumble bees of the western United States. *USDA*

Forest Service Research Notes, 143. https://doi.org/10.1603/0022-0493-99.2.443

Lal, M. M., Southgate, P. C., Jerry, D. R., Bosserelle, C., & Zenger, K. R. (2016). A parallel population genomic and hydrodynamic approach to fishery management of highly-dispersive marine invertebrates: The case of the Fijian Black- Lip Pearl Oyster *Pinctada margaritifera*. *PLoS ONE*, *11*(8), 1–26. https://doi.org/10.1371/journal.pone.0161390

Lenormand, T. (2002). Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, *14*(4), 183–189. https://doi.org/10.1111/j.1365-2184.1976.tb01248.x

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

Li, H., & Durbin, R. (2011). Inference of human population history from individual whole-genome sequences. *Nature*, *475*, 493–496. https://doi.org/10.1038/nature10231

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, Y., Zhang, X. X., Mao, R. L., Yang, J., Miao, C. Y., Li, Z., & Qiu, Y. X. (2017). Ten years of landscape genomics: Challenges and opportunities. *Frontiers in Plant Science*, *8*(December), 1–7. https://doi.org/10.3389/fpls.2017.02136

Liu, H., Jia, Y., Sun, X., Tian, D., Hurst, L. D., & Yang, S. (2017). Direct determination of the mutation rate in the bumblebee reveals evidence for weak recombination-associated mutation and an approximate rate constancy in insects. *Molecular Biology and Evolution*, *34*(1), 119–130. https://doi.org/10.1093/molbev/msw226

Lozier, J. D., Parsons, Z. M., Rachoki, L., Jackson, J. M., Pimsler, M. L., Oyen, K. J., … Dillon,

M. E. (2021). Divergence in body mass, wing loading, and population structure reveals species-specific and potentially adaptive trait variation across elevations in montane bumble bees. *Insect Systematics and Diversity*, *5*(5), 1–15. https://doi.org/10.1093/isd/ixab012

Lozier, J. D., Strange, J. P., & Heraghty, S. D. (2023). Whole genome demographic models indicate divergent effective population size histories shape contemporary genetic diversity gradients in a montane bumble bee. *Ecology and Evolution*, *13*(2), e9778. https://doi.org/10.1002/ece3.9778

Lozier, J. D., Strange, J. P., Stewart, I. J., & Cameron, S. A. (2011). Patterns of range-wide genetic variation in six North American bumble bee (Apidae: Bombus) species. *Molecular Ecology*, *20*(23), 4870–4888. https://doi.org/10.1111/j.1365-294X.2011.05314.x

Lozier, J. D., & Zayed, A. (2016). Bee conservation in the age of genomics. *Conservation Genetics*, *18*, 713–729. https://doi.org/10.1007/s10592-016-0893-7

Lüning, S., & Vahrenholt, F. (2017). Paleoclimatological context and reference level of the 2°C and 1.5°C paris agreement long-term temperature limits. *Frontiers in Earth Science*, *5*, 1–7. https://doi.org/10.3389/feart.2017.00104

Luo, L., Tang, Z. zheng, Schoville, S. D., & Zhu, J. (2021). A comprehensive analysis comparing linear and generalized linear models in detecting adaptive SNPs. *Molecular Ecology Resources*, *21*(3), 733–744. https://doi.org/10.1111/1755-0998.13298

Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: Combining landscape ecology and population genetics. *Trends in Ecology and Evolution*, *18*(4), 189–197. https://doi.org/10.1016/S0169-5347(03)00008-9

Mather, N., Traves, S. M., & Ho, S. Y. W. (2020). A practical introduction to sequentially Markovian coalescent methods for estimating demographic history from genomic data.

Ecology and Evolution, *10*(1), 579–589. https://doi.org/10.1002/ece3.5888

Mccaw, B. A., Stevenson, T. J., & Lancaster, L. T. (2020). Integrative and Comparative Biology Epigenetic Responses to Temperature and Climate, *60*(6), 1469–1480. https://doi.org/10.1093/icb/icaa049

Melero, Y., Evans, L. C., Kuussaari, M., Schmucki, R., Roy, D. B., Oliver, T. H., & Stefanescu, C. (2022). Local adaptation to climate anomalies relates to species phylogeny. *Communications Biology*, *5*(143), 1–9. https://doi.org/10.1038/s42003-022-03088-3

Mola, J. M., Miller, M. R., O'Rourke, S. M., & Williams, N. M. (2020a). Forests do not limit bumble bee foraging movements in a montane meadow complex. *Ecological Entomology*, *45*(5), 955–965. https://doi.org/10.1111/een.12868

Mola, J. M., Miller, M. R., O'Rourke, S. M., & Williams, N. M. (2020b). Wildfire reveals transient changes to individual traits and population responses of a native bumble bee *Bombus vosnesenskii*. *Journal of Animal Ecology*, *89*(8), 1799–1810. https://doi.org/10.1111/1365-2656.13244

Morin, P. A., Archer, F. I., Avila, C. D., Balacco, J. R., Bukhman, Y. V., Chow, W., … Jarvis, E. D. (2021). Reference genome and demographic history of the most endangered marine mammal, the vaquita. *Molecular Ecology Resources*, *21*(4), 1008–1020. https://doi.org/10.1111/1755-0998.13284

Nadachowska-Brzyska, K., Burri, R., Smeds, L., & Ellegren, H. (2016). PSMC analysis of effective population sizes in molecular ecology and its application to black-and-white Ficedula flycatchers. *Molecular Ecology*, *25*(5), 1058–1072. https://doi.org/10.1111/mec.13540

Patil, A. B., Shinde, S. S., Raghavendra, S., Satish, B. N., Kushalappa, C. G., & Vijay, N.

(2021). The genome sequence of *Mesua ferrea* and comparative demographic histories of forest trees. *Gene*, *769*(August 2020), 145214. https://doi.org/10.1016/j.gene.2020.145214

Patil, A. B., & Vijay, N. (2021). Repetitive genomic regions and the inference of demographic history. *Heredity*, *127*(2), 151–166. https://doi.org/10.1038/s41437-021-00443-8

Phillips, S. J., Anderson, R. P., Dudík, M., Schapire, R. E., & Blair, M. E. (2017). Opening the black box: an open-source release of Maxent. *Ecography*, *40*(7), 887–893. https://doi.org/10.1111/ecog.03049

Pimsler, M. L., Oyen, K. J., Herndon, J. D., Jackson, J. M., Strange, J. P., Dillon, M. E., & Lozier, J. D. (2020). Biogeographic parallels in thermal tolerance and gene expression variation under temperature stress in a widespread bumble bee. *Scientific Reports*, *10*(1), 1–11. https://doi.org/10.1038/s41598-020-73391-8

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rahbek, C., Borregaard, M. K., Antonelli, A., Colwell, R. K., Holt, B. G., Nogues-Bravo, D., … Fjeldså, J. (2019). Building mountain biodiversity: Geological and evolutionary processes. *Science*, *365*(6458), 1114–1119. https://doi.org/10.1126/science.aax0151

Räsänen, K., & Hendry, A. P. (2008). Disentangling interactions between adaptive divergence and gene flow when ecology drives diversification. *Ecology Letters*, *11*(6), 624–636. https://doi.org/10.1111/j.1461-0248.2008.01176.x

Reed, D. H., & Frankham, R. (2003). Correlation between fitness and genetic diversity. *Conservation Biology*, *17*(1), 230–237. https://doi.org/10.1046/j.1523-1739.2003.01236.x

Revelle, W. (2022) psych: Procedures for personality and psychological research, Northwestern

University, Evanston, Illinois, USA, https://CRAN.R-project.org/package=psych Version = 2.2.9.

Savolainen, O., Lascoux, M., & Merilä, J. (2013). Ecological genomics of local adaptation. *Nature Reviews Genetics*, *14*(11), 807–820. https://doi.org/10.1038/nrg3522

Sears, M. W., Riddell, E. A., Rusch, T. W., & Angilletta, M. J. (2019). The world still is not flat: Lessons learned from organismal interactions with environmental heterogeneity in terrestrial environments. *Integrative and Comparative Biology*, *59*(4), 1049–1058. https://doi.org/10.1093/icb/icz130

Serpe, M., Ralston, A., Blair, S. S., & O'Connor, M. B. (2005). Matching catalytic activity to developmental function: Tolloid-related processes Sog in order to help specify the posterior crossvein in the *Drosophila* wing. *Development*, *132*(11), 2645–2656. https://doi.org/10.1242/dev.01838

Skovrind, M., Louis, M., Westbury, M. V., Garilao, C., Kaschner, K., Castruita, J. A. S., … Lorenzen, E. D. (2021). Circumpolar phylogeography and demographic history of beluga whales reflect past climatic fluctuations. *Molecular Ecology*, *30*(11), 2543–2559. https://doi.org/10.1111/mec.15915

Slatkin, M. (1993). Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, *47*(1), 264–279.

Stephen, W. P. (1957). Bumble bees of western america (Hymenoptera: Apoidea), (November), 163.

Stolle, E., Wilfert, L., Schmid-Hempel, R., Schmid-Hempel, P., Kube, M., Reinhardt, R., & Moritz, R. F. A. (2011). A second generation genetic map of the bumblebee *Bombus terrestris* (Linnaeus, 1758) reveals slow genome and chromosome evolution in the Apidae.

BMC Genomics, *12*. https://doi.org/10.1186/1471-2164-12-48

Storey, J.D., Bass A.J., Dabney, A., Robinson, D. (2022). *qvalue: Q-value estimation for false discovery rate control*. R package version 2.30.0, http://github.com/jdstorey/qvalue.

Strange, J. P. (2015). *Bombus huntii*, *Bombus impatiens*, and *Bombus vosnesenskii* (Hymenoptera: Apidae) pollinate greenhouse-grown tomatoes in Western North America. *Journal of Economic Entomology*, *108*(3), 873–879. https://doi.org/10.1093/jee/tov078

Sun, C., Huang, J., Wang, Y., Zhao, X., Su, L., Thomas, G. W. C., … Mueller, R. L. (2020). Genus-wide characterization of bumblebee genomes provides insights into their evolution and variation in ecological and behavioral traits. *Molecular Biology and Evolution*, *38*(2), 486–501. https://doi.org/10.1093/molbev/msaa240

Taylor, R. S., Manseau, M., Klütsch, C. F. C., Polfus, J. L., Steedman, A., Hervieux, D., … Wilson, P. J. (2021). Population dynamics of caribou shaped by glacial cycles before the last glacial maximum. *Molecular Ecology*, *30*(23), 6121–6143. https://doi.org/10.1111/mec.16166

Thorp, R. W., Horning Jr., D. S., & Dunning, L. L. (1983). Bumble bees and cuckoo bumble bees of California (Hymenoptera: Apidae). Bulletin of the California Insect Survey (Vol. 23).

Velthuis, H. W., & Doorn, A. Van. (2006). A century of advances in bumblebee domestication and the economic and environmental aspects of its commercialization for pollination. *Apidologie*, *37*, 421–451.

Von Deimling, T. S., Ganopolski, A., Held, H., & Rahmstorf, S. (2006). How cold was the last Glacial maximum? *Geophysical Research Letters*, *33*(14), 1–5. https://doi.org/10.1029/2006GL026484

Wang, A., & Singh, A. (2019). Ecological specialization in populations adapted to constant versus heterogeneous environments, 1–9. https://doi.org/10.1111/evo.13725

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, *38*(6), 1358–1370. https://doi.org/10.1111/j.1558-5646.1984.tb05657.x

Williams, P. H., Cameron, S. A., Hines, H. M., Cederberg, B., & Rasmont, P. (2008). A simplified subgeneric classification of the bumblebees (genus *Bombus*). *Apidologie*, *39*(1), 46–74. https://doi.org/10.1051/apido:2007052

Williams, P. H., Françoso, E., Martinet, B., Orr, M. C., Ren, Z., Júnior, J. S., … Vandame, R. (2022). When did bumblebees reach South America? Unexpectedly old montane species may be explained by Mexican stopover (Hymenoptera: Apidae). *Systematics and Biodiversity*, *20*(1), 1–24. https://doi.org/10.1080/14772000.2022.2092229

Williams, P. H., Lobo, J. M., & Meseguer, A. S. (2018). Bumblebees take the high road: climatically integrative biogeography shows that escape from Tibet, not Tibetan uplift, is associated with divergences of present-day Mendacibombus. *Ecography*, *41*(3), 461–477. https://doi.org/10.1111/ecog.03074

Yang, F., Crossley, M. S., Schrader, L., & Dubovskiy, I. M. (2022). Polygenic adaptation contributes to the invasive success of the Colorado potato beetle. *Molecular Ecology*, *31*, 5568–5580. https://doi.org/10.1111/mec.16666

Yeaman, S. (2022). Evolution of polygenic traits under global vs local adaptation. *Genetics*, *220*(1), 1–15. https://doi.org/10.1093/genetics/iyab134

Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., & Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*, *28*(24), 3326–3328. https://doi.org/10.1093/bioinformatics/bts606

Data Accessibility

Raw sequencing data is available on NCBI SRA (Bioproject PRJNA909611; accession numbers SAMN32298940- SAMN32298956, SAMN32230877 - SAMN32230970 SAMN32298940). Variant data for SNPs as well as scripts used in data filtering and analysis are available on FigShare.

Author Contributions

S.D.H. processes samples, performed statistical analysis, wrote the manuscript, and edited the manuscript. J.M.J. collected samples and helped to process samples. J.D.L. designed the study, obtained funding, conducted field work, helped with statistical analysis, and edited manuscript.

Tables

Site	Number of samples	Latitude	Longitude	Elevation (m)	π
CA01.2015	8	36.458	-118.879	313	0.2745
CA04.2012	11	41.356	-122.208	2261	0.2743
CA06.2015	11	37.012	-119.732	131	0.27478
CA11.2013	12	39.316	-120.326	2164	0.2745
CA11.2015	7	39.237	-121.451	49	0.2747
CA12.2015	10	40.508	-122.322	138	0.2746
CA13.2015	8	36.619	-118.811	2268	0.2751
CA29.2015	9	37.291	-119.102	2796	0.2746
OR02.2015	5	45.238	-123.063	53	0.2741
OR05.2016	10	45.333	-121.670	1698	0.2743
OR08.2012	2	42.076	-122.717	2055	0.2747*
OR09.2012	8	42.073	-122.754	2134	0.2747*
OR11.2012	10	42.118	-123.085	518	0.2746

Table 1: Summary of sampling sites and average nucleotide diversity (π) per population. * OR08.2012 and OR09.2012 were pooled together as a single population when calculating nucleotide diversity.

A)

Model	Inertia	R^2	*p*	Proportion of explainable variance	Proportion of total variance
Full (Env + Lat)	13985	0.037	0.001	1	0.0486
Environment pRDA	11140	0.009	0.09	0.7966	0.0387
Latitude pRDA	2781		0.503	0.1989	0.0097
Confounded	64			0.0046	0.0002
Unexplained	273535				0.9514
Total	287520				1

B)

Model	Inertia	R^2	*p*	Variance Explained	Total Variance
Full (Env + Lat)	13985	0.037	0.001	0.0486	287520
Environment RDA	11205	0.009	0.001	0.0386	290301
Geography RDA	2845		0.001	0.0095	298660

Table 2: A) Results of the pRDA analysis reporting the results of the full RDA model (BIO1 + BIO3 + BIO12 + Elevation + Latitude), the results of the Environment pRDA (BIO1 + BIO3 + BIO12 + Elevation | Latitude) and the results of the Latitude pRDA (Latitude | BIO1 + BIO3 + BIO12 + Elevation). It also shows that amount of confounded variation (variation that could be associated with either pRDA model), and the amount of unexplained variation. B) Results of different the full RDA model (BIO1 + BIO3 + BIO12 + Elevation + Latitude) (note this is the same as in table 1A), the results of the Environment RDA (BIO1 + BIO3 + BIO12 + Elevation) and the result of the Latitude RDA (Latitude).

SCF	POS	LOC	Name	Fly homolog	Environmental variable
NW_022882922.1	3208914	LOC117230477	fat-like cadhherin-related tumor suppressor homolog	kug	BIO3,BIO12
NW_022882923.1	104464	LOC117232989-LOC117234423	uncharacterized - uncharacterized	#N/A - #N/A	BIO3
NW_022882923.1	713637	LOC117232142	uncharacterized	#N/A	BIO12
NW_022882923.1	4077534	LOC117234457	tropomodulin	tmod	BIOI3
NW_022882923.1	7217833	LOC117232441	protein spliit ends-like	sba	BIO12
NW_022882925.1	6861603	LOC117242995-LOC117243288	trehalase - uncharacterized	Treh - #N/A	BIO12
NW_022882926.1	2932739	LOC117243561	lysine-specific histone demethylase 1A	Su(var)3-3	BIO3,BIO12
NW_022882927.1	1136770	LOC117230390	collagen alpha chain CG42342	CG42342	BIO12
NW_022882927.1	5589210	LOC117230819	conserved oligomeric Golgi complex subunit 1	Cog1	BIO1
NW_022882927.1	7339228	LOC117230712	max dimerization protein 1-like	#N/A	BIO3,BIO12
NW_022882928.1	1315015	LOC117230941-LOC117230955	uncharacterized - protein adenylyltransferase Fic	Sox14 - Fic	BIO12
NW_022882928.1	1771666	LOC117231000	nephrin-like	#N/A	BIO12
NW_022882928.1	2109466	LOC117231079	nephrin-like	#N/A	BIO12
NW_022882930.1	642634	LOC117231202	hemicentin-1	nrm	BIO12
NW_022882930.1	1400678	LOC117231243	collageen alphda-1 (XVIII) chain	Mp	BIO12
NW_022882930.1	1701967	LOC117231281	tRNA dimethylallytransferase	CG31381	BIO12
NW_022882930.1	4543233	LOC117231348	potassium voltage-gated channel protein eag	eag	BIO3,BIO12
NW_022882930.1	5271102	LOC117231280-LOC117231225	fringe glycosyltransferase - uncharacterized	fng - CG32432	BIO12
NW_022882932.1	911460	LOC117231809	uncharacterized	#N/A	BIO12
NW_022882933.1	211778	LOC117231908-LOC117231891	innexin inx2-like - innexin inx7-like	Inx2 - Inx7	BIO12
NW_022882934.1	701863	LOC117231927	calcium/calmoodulin-dependent ptoetin kinase type II alpha chain	CaMKII	BIO3
NW_022882934.1	701882	LOC117231927	calcium/calmoodulin-dependent ptoetin kinase type II alpha chain	CaMKII	BIO3

SCF	POS	LOC	Name	Fly homolog	Environmental variable
NW_022882934.1	738415	LOC117231927	calcium/calmoodulin-dependent ptoetin kinase type II alpha chain	CaMKII	BIO12
NW_022882938.1	2188879	LOC117232181	uncharacterized	#N/A	BIO12
NW_022882938.1	2192345	LOC117232181	uncharacterized	#N/A	BIO12
NW_022882938.1	3280509	LOC117232145-LOC117232141	pre-mRNA 3'-end-processing factor FIP1 - pumilio homolog 2	#N/A - pum	BIO12
NW_022882939.1	813717	LOC117232387-LOC117232337	zinc finger MIZ domain-containing protein 1 - uncharacterized	tna - #N/A	BIO12
NW_022882942.1	154820	LOC117232587	uncharacterized	#N/A	BIO1
NW_022882945.1	670729	LOC117233040	small conductance calcium-activated potassium channel protein	SK	BIO12
NW_022882945.1	1365986	LOC117233045	class E basic helix-loop-helix proteiini 23	Oli	BIO12
NW_022882946.1	5568048	LOC117233408	lysophoospholipid acyltransferase 1	oys	BIO12
NW_022882946.1	6135136	LOC117233164	serine/threoonine-proteiin phosphatase 4 regulatory subunit 1-like	#N/A	BIO12
NW_022882947.1	1392144	LOC117233588	integrin alpha-PS2	if	BIO3
NW_022882947.1	1928801	LOC117233584	putatiive polypeptide N-acetylgalactosaimnyltransferase 9	Pgant9	BIO12
NW_022882947.1	2031308	LOC117233584	putatiive polypeptide N-acetylgalactosaimnyltransferase 9	Pgant9	BIO12
NW_022882947.1	2169738	LOC117233552-LOC117233534	multiple coagulation factor deficiency protein 2 homolog - homeobox protein prophet of Pit-1	#N/A - CG32532	BIO12
NW_022882949.1	1126918	LOC117233891	dorsal-ventral patterning protein Sog	sog	BIO12
NW_022882951.1	567652	LOC117234079-LOC117234069	uncharacterized - uncharacterized	#N/A - CG5756	BIO3,BIO12
NW_022882951.1	1054364	LOC117234181	uncharacterized	#N/A	BIO12
NW_022882953.1	2335188	LOC117234317	uncharacterized	#N/A	BIO3,BIO12
NW_022882953.1	3616066	LOC117234263	potassium voltage-gated channel subfamily KQT member 1	KCNQ	BIO1
NW_022882957.1	72400	LOC117234678-LOC117234677	BET1 homolog - KN motif and ankyrin repeat domain-containing protein 2	Bet1 - Kank	BIO12
NW_022882957.1	515680	LOC117234759-	protein lethal(2)essential for life-like - high affinity cAMP-	l(2)efl -	BIO3

SCF	POS	LOC	Name	Fly homolog	Environmental variable
		LOC117234755	specific and IBMX-insensitive 3',5'-cyclic phosphodiesterase 8-like	Pde8	
NW_022882958.1	204189	LOC117234852	uncharacterized	#N/A	BIO12
NW_022882958.1	1005699	LOC117234833-LOC117234847	SAGA-associated factor 29 - uncharacterized	Sgf29 - #N/A	BIO12
NW_022882959.1	1208400	LOC117234939	microtubule-assoociated protein futsch-like	futsch	BIO12
NW_022882960.1	3778493	LOC117235117	protein cortex-like	cort	BIO12
NW_022882961.1	265994	LOC117235314-LOC117235313	uncharacterized - alpha-glucosidase-like	#N/A - #N/A	BIO12
NW_022882962.1	236847	LOC117235573	long-chaiin fatty acid transport protein 4	Fatp2	BIO12
NW_022882962.1	236867	LOC117235573	long-chaiin fatty acid transport protein 4	Fatp2	BIO12
NW_022882962.1	4277581	LOC117235535-LOC117235540	receptor-type guanylate cyclase Gyc76C-like - somatostatin receptor type 2-like	CG42637 - Astc-R2	BIO12
NW_022882964.1	1850900	LOC117235795	uncharacterized	#N/A	BIO3
NW_022882964.1	1850938	LOC117235795	uncharacterized	#N/A	BIO12
NW_022882972.1	399855	LOC117235909	ras-responsive element-biinding protein 1	peb	BIO12
NW_022882977.1	723073	LOC117236241	band 3 anion transport protein	Ae2	BIO3,BIO12
NW_022882985.1	1602088	LOC117236599	midasin	CG13185	BIO3
NW_022882988.1	1948816	LOC117236799	transient receptor potential-gamma protein	Trpgamma	BIO12
NW_022882989.1	355203	LOC117237000-LOC117237016	E3 ubiquitin-protein ligase RBBP6-like - uncharacterized	#N/A - #N/A	BIO12
NW_022882990.1	345704	LOC117237067	uncharacterized	#N/A	BIO12
NW_022882990.1	345853	LOC117237067	uncharacterized	#N/A	BIO12
NW_022882990.1	952989	LOC117237065	histamine H2 receptor-like	#N/A	BIO12
NW_022882990.1	992602	LOC117237065	histamine H2 receptor-like	#N/A	BIO12
NW_022883009.1	908150	LOC117237571	sperm flagellar protein 2-likee	#N/A	BIO12
NW_022883019.1	226656	LOC117237628-LOC117237625	uncharacterized - protein slit	#N/A - sli	BIO12
NW_022883022.1	213524	LOC117237685-	LHFPL tetraspan subfamily member 6 protein-like -	#N/A -	BIO12

SCF	POS	LOC	Name	Fly homolog	Environmental variable
		LOC117237701	uncharacterized	#N/A	
NW_022883036.1	516302	LOC117237948-LOC117237935	serine/threonine-protein kinase 17B-like - probable serine/threonine-protein kinase MARK-A	Drak - #N/A	BIO12
NW_022883058.1	145023	LOC117238319	homeotic protein spalt-major-like	salm	BIO3,BIO12
NW_022883144.1	503689	LOC117238692-LOC117238739	vesicular glutamate transporter 1 - uncharacterized	Vglut - #N/A	BIO3,BIO12
NW_022883264.1	1402456	LOC117239068-LOC117239065	uncharacterized - uncharacterized	#N/A - #N/A	BIO12
NW_022883289.1	248365	LOC117239335	uncharacterized	#N/A	BIO12
NW_022883289.1	248458	LOC117239335	uncharacterized	#N/A	BIO12
NW_022883317.1	14976	LOC117239422	uncharacterized	#N/A	BIO12
NW_022883371.1	3253233	LOC117239509	phosphatase and actin regulator 4B	CG32264	BIO12
NW_022883452.1	410515	LOC117239853	protein madd-4-like	nolo	BIO12
NW_022883531.1	1137597	LOC117240238	odorant receptor 13a-like	#N/A	BIO12
NW_022884330.1	20490	LOC117241676	bone morphogenetiic protein 1-like	tok	BIO1
NW_022884342.1	679964	LOC117242267	homeotic protein antennapedia-like	Antp	BIO12
NW_022884342.1	1078031	LOC117242285	uncharacterized	#N/A	BIO3
NW_022884343.1	3681294	LOC117242443	cAMP-specific 3',5'-cyclic phosphodiesterase-like	dnc	BIO3
NW_022884344.1	1195029	LOC117242689-LOC117242710	uncharacterized - uncharacterized	#N/A - #N/A	BIO12
NW_022884348.1	2233766	LOC117242787	uncharacterized	#N/A	BIO12

Table 3: SNPs that were identified as environmentally associated outliers noting the scaffold the SNP is located on (SCF), the position of the of the SNP (POS), the NCBI identifier (LOC), the NCBI gene name (Name), the homologous gene in *D. melanogaster* if available (Fly homolog), and the environmental variable(s) associated with the SNP.

856

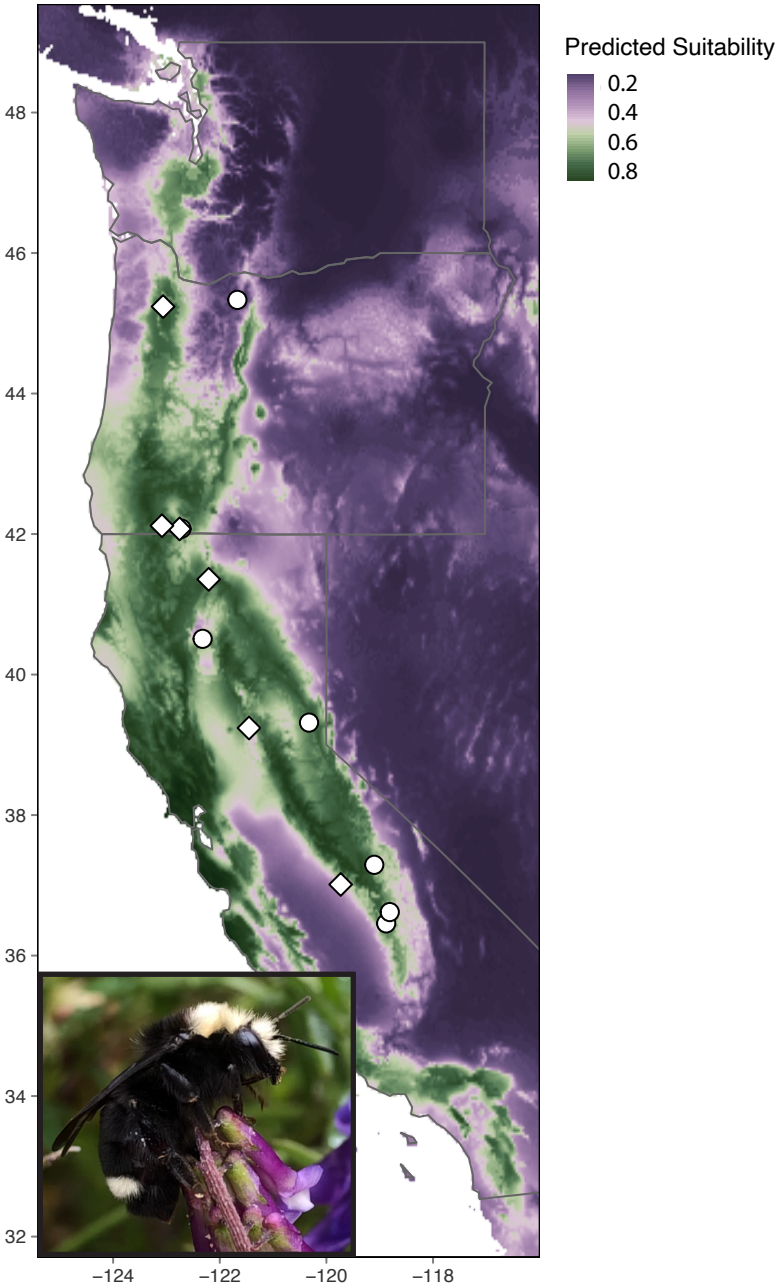
Figures

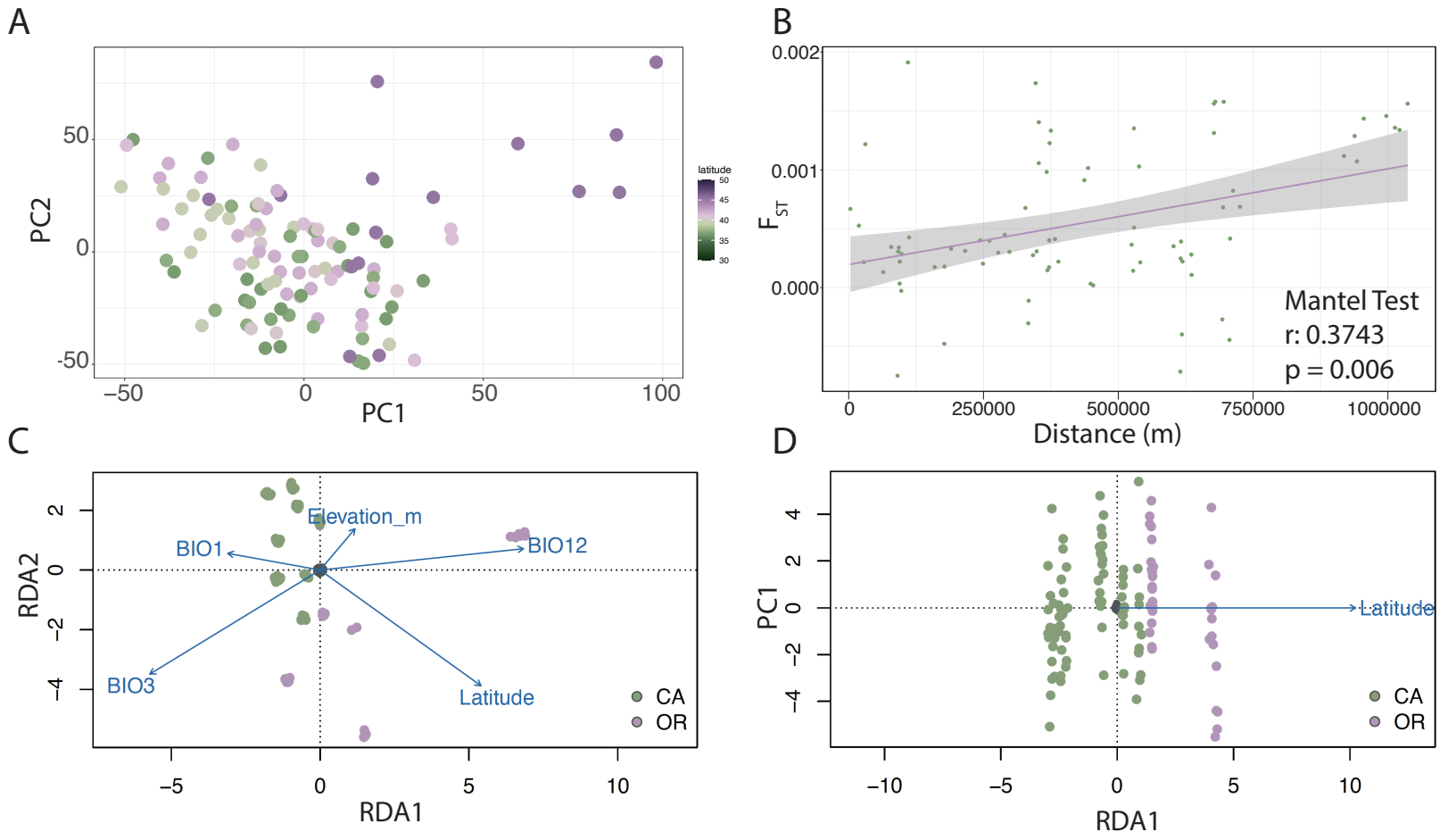
Figure 1: Map showing Maxent (Phillips, Anderson, Dudík, Schapire, & Blair, 2017) range of the *B vosnesenskii* using select environmental variables (BIO1, BIO3, and BIO12) with presence absence data from (Cameron et al., 2011). White diamonds indicate sites that were used for PSMC analysis.

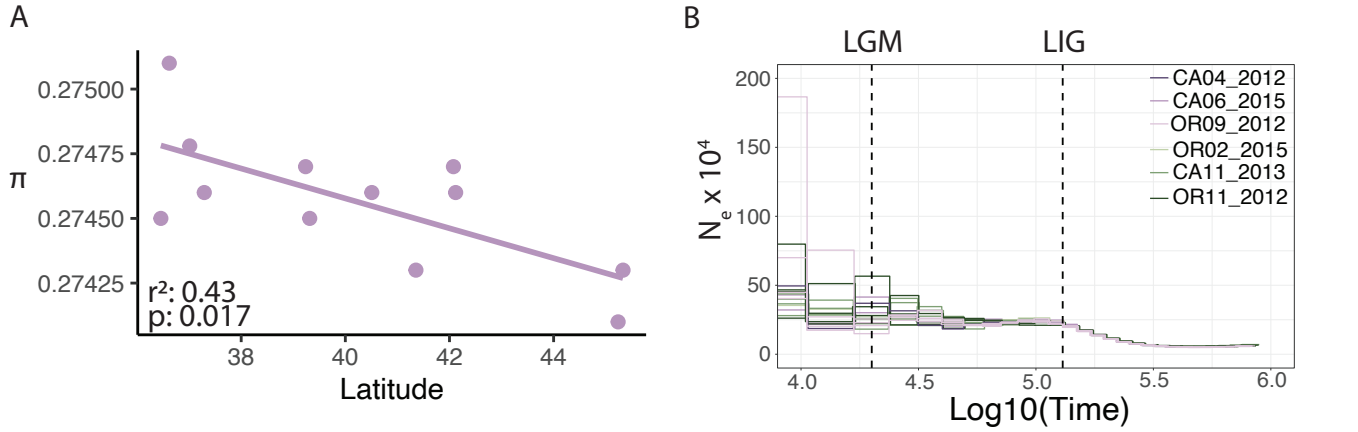
Figure 2: A) Pearson Principal Component analysis with samples colored by latitude of origin. B) Results of the isolation by distance analysis. C) The full RDA model containing all selected environmental variables as well as latitude (a proxy of distance) with samples colored by state of origin. Note this model is essentially identical to the RDA with only environmental variables, so we only present the results of the full model. D) The geography RDA model with samples colored by state of origin.

Figure 3: A) Demographic inference from PSMC models, marked with the approximate dates of the last glacial maximum (LGM - 20,000 years ago) and last interglacial period (LIG – 130,000 years ago) B) Linear regression between population π and latitude.

Figure 4: Manhattan plot of LFMM2 *q*-values for each of the four selected environmental variables (BIO1, BIO3, BIO12, and elevation). The black line represents a *q*-value threshold of 0.05, which is the threshold for significance.


Predicted Suitability
0.2
0.4
0.6
0.8
48
46
44
42
40
38
36
34
32
−124
−122
−120
−118

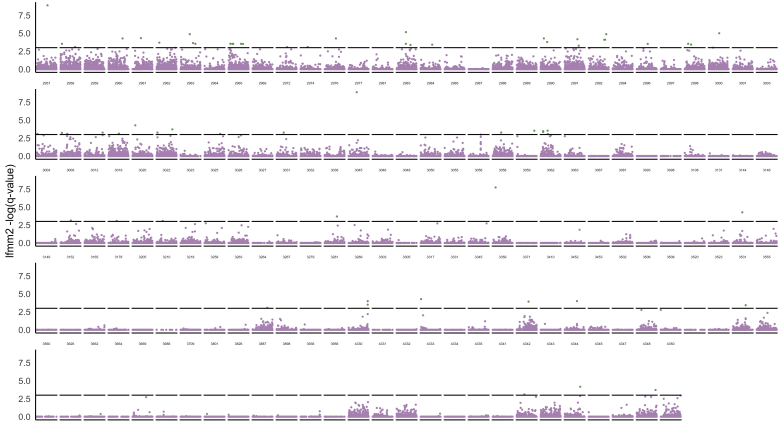
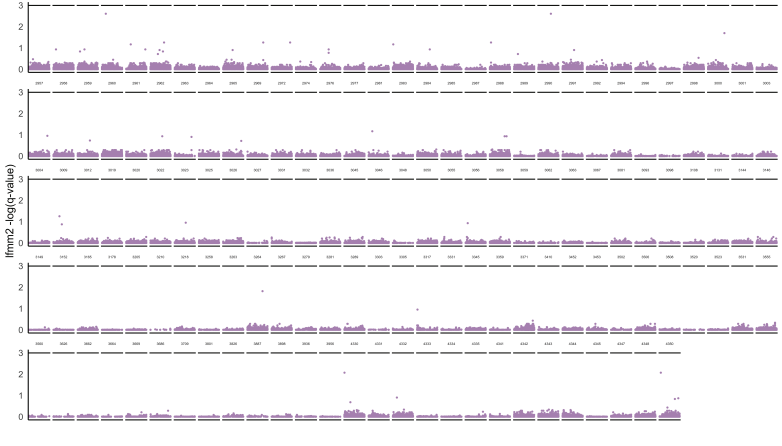

A
PC2
PC1
latitude
50
45
40
35
30
B
F_{ST}
Distance (m)
Mantel Test
r: 0.3743
p = 0.006
C
RDA2
RDA1
Elevation_m
BIO1
BIO12
BIO3
Latitude
CA
OR
D
PC1
RDA1
Latitude
CA
OR


A
0.27500
0.27475
0.27450
0.27425
π
r²: 0.43
p: 0.017
38
40
42
44
Latitude
B
LGM
LIG
200
150
100
50
0
N_e x 10^4
CA04_2012
CA06_2015
OR09_2012
OR02_2015
CA11_2013
OR11_2012
4.0
4.5
5.0
5.5
6.0
Log10(Time)

BIO1: Mean Annual Temperature - 4

BIO3: Isothermality - 16

BIO12: Annual Precipitation - 66



Elevation - 0