

The impact of geography and climate on the population structure and local adaptation in a wild bee

Farida Samad-zada | Evan P. Kelemen  | Sandra M. Rehan 

Department of Biology, York University,
Toronto, Canada

Correspondence

Sandra M. Rehan, Department of Biology,
York University, Toronto M3J 1P3,
Canada.
Email: sandra.rehan@gmail.com

Funding information

National Science Foundation; Natural
Sciences and Engineering Research
Council of Canada

Abstract

Deciphering processes that contribute to genetic differentiation and divergent selection of natural populations is useful for evaluating the adaptive potential and resilience of organisms faced with various anthropogenic stressors. Insect pollinator species, including wild bees, provide critical ecosystem services but are highly susceptible to biodiversity declines. Here, we use population genomics to infer the genetic structure and test for evidence of local adaptation in an economically important native pollinator, the small carpenter bee (*Ceratina calcarata*). Using genome-wide SNP data ($n = 8302$), collected from specimens across the species' entire distribution, we evaluated population differentiation and genetic diversity and identified putative signatures of selection in the context of geographic and environmental variation. Results of the analyses of principal component and Bayesian clustering were concordant with the presence of two to three genetic clusters, associated with landscape features and inferred phylogeography of the species. All populations examined in our study demonstrated a heterozygote deficit, along with significant levels of inbreeding. We identified 250 robust outlier SNPs, corresponding to 85 annotated genes with known functional relevance to thermoregulation, photoperiod, and responses to various abiotic and biotic stressors. Taken together, these data provide evidence for local adaptation in a wild bee and highlight genetic responses of native pollinators to landscape and climate features.

KEYWORDS

gene–environment associations, pollinators, population genomics, signatures of selection

1 | INTRODUCTION

Evolutionary processes in wild populations are determined by a variety of factors, at different spatial and temporal scales (Chevin et al., 2010). Understanding the adaptive potential of a species requires considering both neutral and selective forces that have historically shaped the population of interest (Holderegger et al., 2006).

Characterizing population structure, diversity, and gene flow is crucial for conservation as this helps identify biogeographical features that have shaped population dynamics over time (Hedrick, 2001; Manel et al., 2003). Another important goal of conservation genomics is elucidating the complex interactions between ecological processes and selection pressures (Otto, 2018). There is increasing interest in understanding how historical demography shaped by

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Evolutionary Applications* published by John Wiley & Sons Ltd.

environmental and geographic variation affects population structure of species of conservation concern. Determining how natural populations have adapted to their local environments can help mitigate the threat of biodiversity declines by informing conservation management strategies (Hohenlohe et al., 2021).

Global biodiversity loss has a direct negative impact on crucial ecosystem services, a prime example of which is the decline of insect pollinator species (Potts et al., 2010). According to different estimates, crop reliance on animal-mediated pollination ranges from 78% to 94% (Ollerton et al., 2011) with bees (Hymenoptera: Anthophila) providing the majority of pollination services (Klein et al., 2007). There exist more than 20,000 species of bees (Michener, 2007) of which at least 4000 are found in North America. Recent research has shown that wild bees provide a significant proportion of pollination services both in agricultural (Kennedy et al., 2013) and urban settings (Lowenstein et al., 2015). At the community level, many plants rely on uncommon bees for pollination, highlighting the importance of rare species in plant-pollinator networks (Simpson et al., 2022).

Despite the fact that wild bee populations are particularly vulnerable to habitat loss (Kline & Joshi, 2020; Koh et al., 2016; Potts et al., 2010) as evidenced by their continued decline on the global scale (Aldercotte et al., 2022; Jacobson et al., 2018; Zattara & Aizen, 2021), bee genetic research has concentrated on a few taxa (primarily *Apis* and *Bombus* spp.), while the majority of solitary species continue to be underrepresented in conservation and genetics literature (Cameron & Sadd, 2020; Kelemen & Rehan, 2021). As sequencing becomes more accessible and is routinely used to characterize nonmodel species, newly emerging techniques and analyses make it possible to answer previously intractable conservation questions. This allows researchers to not only infer population connectivity and levels of inbreeding, but also to perform analyses that identify signatures of selection, estimate historic population sizes, and describe migration patterns (Allendorf et al., 2010; Fraser & Whiting, 2020).

In recent years, there has been an increase in the number of studies aiming to characterize local adaptation in bees. Genome-wide screening has identified regions under selection across bee genomes (reviewed in Kelemen & Rehan, 2021); but research has predominantly focused on traits unrelated to environmental conditions (honeybees: (Cridland et al., 2018; Grozinger et al., 2007; Harpur et al., 2014; Mikheyev et al., 2015)), sweat bees: (Kocher et al., 2013), bumble bees: (Harpur et al., 2017) and carpenter bees (Rehan et al., 2016). Only a few number of studies have investigated patterns of selection with respect to geographic and climatic variation. For instance, a study on *Bombus vosnesenskii* and *B. vancouverensis*, has found an association between temperature and genes related to neural and neuromuscular function, as well as ion transport (Jackson et al., 2020). This study also found an association between precipitation and genes related to cuticle formation, homeostasis, tracheal, and respiratory system development (Jackson et al., 2020). There is also putatively adaptive genetic variation associated with latitude in the stingless bee *Melipona subnitida* (Jaffé et al., 2019) and the honeybee *Apis mellifera* (Hadley & Betts, 2012; Henriques et al., 2018).

Iberian Peninsula populations of *A. mellifera* show latitudinal gradients associated with clock genes, suggesting that circadian rhythms are involved in local adaptation (Henriques et al., 2018). This species also shows distinct adaptive genetic variation along elevational gradients (Jaffé et al., 2019; Wallberg et al., 2017). Furthermore, populations of *A. mellifera* in East Africa not only exhibit panmixia but also show divergence at two haplotype blocks, one of which contains all octopamine receptor genes (Wallberg et al., 2017), which could indicate that adaptation to altitude is connected to processes associated with learning and memory.

Identification of genes under selection can provide information on molecular processes driving phenotypic divergences, as well as shed light on evolutionary trajectories responsible for local adaptations (Fraser & Whiting, 2020). As analysis of genome-wide markers becomes a standard approach in conservation genomics, more studies aim to identify signatures of selection associated with various questions of evolutionary importance. Examples include searching for candidate genes underlying phenotypic divergence (Pimsler et al., 2017), parallel evolution (Deagle et al., 2012), adaptation to climatic conditions (Schmidt et al., 2021), and anthropogenic disturbances (Theodorou et al., 2018). Moreover, diverging loci can be used to infer patterns of genetic differentiation in recently isolated populations, in cases where conventional neutral markers do not provide enough resolution (Funk et al., 2012; Russello et al., 2012). While local adaptation patterns are largely species-specific (Nooten & Rehan, 2020), including more species in population genomics studies can be useful for uncovering patterns in global bee declines and identifying potential conservation management strategies, by increasing our understanding of how variation in environmental conditions and historical demography affects the contemporary genetic structure of native pollinators.

Despite the recognized value of investigating evolutionary responses across a diverse array of species, population genomics studies on nonsocial bees remain scarce. Given that social bees comprise less than 10% of all bee species (Danforth et al., 2019), understanding the biology of solitary and subsocial species is crucial for evaluating the adaptive potential of bee communities in the face of changing environmental conditions. Here, we investigate whether a small carpenter bee, *Ceratina calcarata*, exhibits evidence of local adaptation across its entire geographic distribution, spanning multiple climatic and elevational gradients. *Ceratina calcarata* is a common stem-nesting bee endemic to eastern North America (Nooten et al., 2020; Rehan & Sheffield, 2011) that occurs in high abundances across various landscapes from southern Ontario in Canada to Georgia in the United States; the species is subsocial, exhibiting prolonged maternal care (Rehan & Richards, 2010). *Ceratina calcarata* is a generalist crop pollinator (Kennedy et al., 2013), among the top 20 most economically important wild bee species (Kleijn et al., 2015), and a key contributor to plant-pollinator networks (Russo et al., 2013). Given its ubiquity and the availability of genomic (Rehan et al., 2016) and transcriptomic (Rehan et al., 2014) resources, *C. calcarata* can serve as an important model species for investigating the effects of environmental variation on adaptive responses in wild bees.

Here, we employed genotyping by sequencing to collect genome-wide data for *C. calcarata* samples across the species' entire geographic distribution. The resulting single nucleotide polymorphism (SNP) dataset was used to address several questions. First, we inferred the population structure of *C. calcarata* and compared genetic diversity estimates across different sampling locations to help characterize the ecological biogeography of the species. Additionally, we tested for evidence of adaptive divergence by identifying putative signatures of selection associated with genetic structure and environmental variation. Combined, these objectives provide insights into the population genomics of an ecologically important pollinator species, and, more broadly, demonstrate how geography and climate shape patterns of local adaptation in wild bees.

2 | METHODS

2.1 | Specimen collection, DNA extraction, and sequencing

Adult female *C. calcarata* ($n=129$) were collected from 33 sites from 14 regions spanning the range of the species (Figure 1; Table S1). Regions were separated by a minimum of 225 km with a maximum distance of 2787 km and a mean distance of $1079 \text{ km} \pm 565 \text{ SD}$. Bees were collected via sweep nets and pan traps in May–August 2015 to 2019 and preserved using 95% ethanol or pinning. We extracted DNA from specimens in two ways. For samples preserved in 95% ethanol, DNA was extracted from abdomens using DNeasy Blood & Tissue Kits (Qiagen). For pinned specimens, we extracted DNA from whole bodies employing a nondestructive technique (whole body

incubation) using Quick-DNA Miniprep Plus extraction kits (Zymo Research) and following a modified protocol (Freitas et al., 2021). DNA quality was assessed using agarose gel electrophoresis and a spectrophotometer (Thermo Fisher Scientific). For each individual, DNA concentration was normalized to $20 \text{ ng}/\mu\text{L}$. RAD-seq library preparation was carried out by Floragenex, Inc, according to the original RAD protocol described by Baird et al. (2008). DNA from each individual was digested with the restriction enzyme *Pst*I. Sequencing was done using two lanes of Illumina HiSeq 3000 at the Center for Genome Research and Biocomputing at Oregon State University for single pair end 100bp reads and produced a total of 521 million reads.

2.2 | Read mapping, variant calling, and filtering

We demultiplexed FASTQ files and removed multiplex identifying barcode sequences using *process_radtags* in STACKS v2.3 (Catchen et al., 2011; Rochette et al., 2019). We excluded reads if they contained an ambiguous *Pst*I cut site, incorrect barcode, or average Phred quality score < 10 within a sliding window (default), with 468.5 million reads (89.86%) retained following demultiplexing. Reads were filtered out because of ambiguous barcodes (9.41%), low quality scores (0.09%), or ambiguous RAD tags (0.63%). Due to a low number of reads (< 7000), two individuals were removed from further analysis. Generated FASTQ files were aligned to the *Ceratina calcarata* genome (Rehan et al., 2016; BioProject: PRJNA791561) using *bwa-mem* v0.7.17 (Li & Durbin, 2009). SAM outputs were converted to BAM and sorted with SAMTOOLS v1.6 (Li et al., 2009). SNP calling was made using *gstacks* in STACKS v2.3

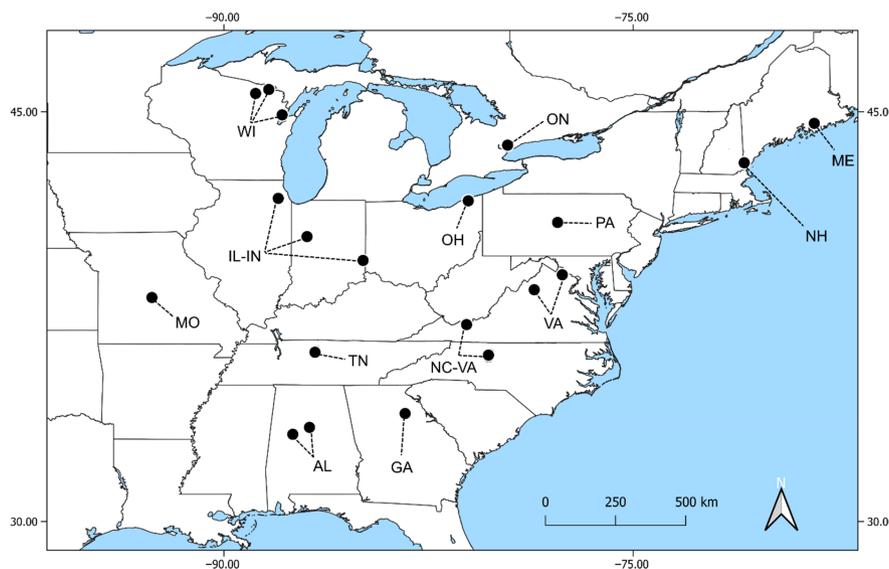


FIGURE 1 Map showing sample locations and populations of *Ceratina calcarata* examined in this study. Abbreviations of the populations, along with sample sizes are provided below. AL, Alabama ($n=5$), GA, Georgia ($n=9$), IL-IN, Illinois/Indiana ($n=10$), ME, Maine ($n=8$), MO, Missouri ($n=10$), NC-VA, North Carolina/Virginia ($n=10$), NH, New Hampshire ($n=8$), OH, Ohio ($n=9$), ON, Ontario ($n=10$), PA, Pennsylvania ($n=10$), TN, Tennessee ($n=5$), VA, Virginia ($n=8$), and WI, Wisconsin ($n=10$). Map produced using QGIS.org (2020), QGIS Geographic Information System, Open Source Geospatial Foundation Project (<http://qgis.org>). Shape files retrieved from <https://www.naturalearthdata.com/>.

(Rochette et al., 2019) under the Marukilow model, with an alpha threshold for discovering SNPs and calling genotypes set at 0.05. We used the *populations* module in STACKS to call SNPs, present in more than 80% of individuals across all populations ($-R\ 80$), minimum minor allele frequency of 0.05 ($--min-maf\ 0.05$), maximum observed heterozygosity of 0.5 ($--max-obs-het\ 0.5$), and retaining one SNP per locus ($--write-single-snp$) to decrease the effects of linkage disequilibrium. Following the initial *populations* run, we identified and removed low coverage individuals (depth < 6x) and/or individuals with high missing data percentage (>50%), calculating both statistics using *vcftools* v0.1.16 (Danecek et al., 2011). After removing 15 samples due to low coverage and/or high missing data percentage, we reran *populations* on the remaining samples, using the same parameters as above. We used *vcftools* v0.1.16 to filter the resulting VCF file, keeping SNPs with minimum depth of 10x and maximum depth of 100x ($--min-meanDP\ 10$ and $--max-meanDP\ 100$) and missing data below 10% ($--max-missing\ 0.9$). To ensure that no potential sibling pairs were present in the dataset, we calculated relatedness using the kinship coefficient method (Manichaikul et al., 2010), as implemented in *vcftools* v0.1.16 ($--relatedness2$). The resulting dataset was used to identify outlier SNPs.

2.3 | Detection of outlier SNPs

Genome scans are used in population genomics studies to identify potentially divergent loci; however, this approach is characterized by a high level of false positives (Narum & Hess, 2011). To mitigate this issue, a common strategy is to use several outlier detection methods (Alves-Pereira et al., 2020). Here, we combined four different approaches. First, we employed the F_{ST} -based method implemented in Bayescan v2.0 (Foll & Gaggiotti, 2008), running the program for 100,000 iterations with a 50,000 burn-in period and prior-odds set to 100. SNPs with a $q < 0.05$ were considered outliers.

Second, we detected outliers using the principal component analysis (PCA) method, as implemented in *pcadapt* v4.3.3 (Luu et al., 2017). Using Cuttlet's rule, we inferred the optimal number of principal components (PC), and then employed the *ld.clumping* procedure to thin the SNP dataset, with the window radius of 200 and the squared correlation threshold of 0.2. We corrected the resulting p -values for multiple comparisons using the method of Benjamini and Yekutieli (2001), using the R package *q-value*, assuming all SNPs with a $q < 0.05$ to be outliers.

Third, to identify signatures of local adaptation potentially associated with environmental variables, we used the gene-environment association (GEA) approach for outlier loci detection. We selected the following geographical/environmental variables for this analysis: latitude, longitude, elevation, mean annual temperature, and mean annual precipitation between the years 1970–2000 (<https://www.worldclim.org/data/worldclim21.html>). For each sample in the dataset, we extracted environmental variables

based on sample coordinates, using the R package *raster* v3.4–5. To control for multicollinearity, we conducted a collinearity test between potential environmental predictors using the *cor* command within R package *stats* v4.04 and removed one variable from each pair, where collinearity estimate (r) was higher than 0.8. Based on this test, temperature was highly correlated with latitude, and hence, only one of these variables (temperature) was kept as a predictor for GEA tests. First, we used latent factor mixed modelling (LFMM), as implemented in the R package *LEA* (Frichot & François, 2015). This method uses individual level data to find associations between allele frequencies and environmental variables of interest, and accounts for background population structure using latent factors. To select the optimal number of latent factors for this analysis, we estimated admixture coefficients using Sparse nonnegative matrix factorization (SNMF), selecting a number of clusters (K) with the lowest cross entropy. We then imputed missing data using the *impute* function in the *LEA* package, using the selected value of K . LFMM was run with 10 repetitions, 100,000 iterations and a burn-in step of 50,000. Finally, we combined the resulting z -scores across all runs, and recalculated p -values, using the genomic inflation factor (λ). Following λ recalibration and a subsequent correction for multiple comparisons, SNPs with a q -value below 0.05 were considered outliers.

Finally, we used another GEA approach implemented in BayPass (Gautier, 2015), which identifies loci that are associated with population-specific covariates. To account for the shared ancestry and population structure within the dataset, we used the core model within BayPass to generate a covariance matrix of allele frequencies. The resulting matrix was then used as an input file to run the auxiliary covariate model, which tests for associations between SNPs and population covariates. We ran both models for 20 pilot runs, 10,000 iterations, and a burn-in period of 50,000. SNPs with a Bayes factor above 20 dB were considered outliers.

2.4 | Annotation of outlier SNPs

We extracted shared SNPs among the above four methods, using the R packages *dplyr* and *tidyft* within *tidyverse* v1.3.1 (Wickham et al., 2019) and *gplots* v3.1.1 (Warnes et al., 2020): SNPs that were identified as outliers by two or more methods were considered robust. We used *bedtools* v2.26.0 (Quinlan & Hall, 2010) and *BEDOPS* v2.4.38 (Neph et al., 2012) *bedmap --echo-map-id-uniq* command to extract gene IDs associated with genomic positions of the robust outlier SNPs, using the latest annotation of the *C. calcarata* genome (<http://www.rehanlab.com/genomes.html> PO1409_Ceratina_calcarata.annotation.gff.gz). We investigated the putative role of these genes, using the R package *TopGO* (Alexa & Rahnenfuhrer, 2022), by testing the associated gene ontology (GO) terms (terms generated by BLAST2GO (Gotz et al., 2008)), for enrichment using the *elim* algorithm and Fisher's exact test to test for significance. Only biological process (BP) terms were used in this analysis and associations with a p -value below 0.05 were considered significant.

2.5 | Population genetics analyses

To create a putatively neutral SNP dataset for downstream population genetics analyses, we removed all SNPs that were identified as outliers by at least one of the above methods. Furthermore, we removed SNPs that significantly ($p < 0.01$) deviated from Hardy-Weinberg equilibrium (HWE) in at least 50% of the populations. The resulting SNP dataset was used for all population genetics analyses below, unless otherwise specified.

First, we calculated observed (H_O) and expected (H_S) heterozygosities, observed and effective numbers of alleles, and inbreeding coefficients (G_{IS}) for each sampling location in the dataset using GenoDive v3.04 (Meirmans, 2020). We estimated pairwise differentiation between sampling locations by calculating θ (Weir & Cockerham, 1984), as implemented in GenoDive v3.04, using 10,000 permutations to test for significance.

To evaluate the population structure of *C. calcarata*, we used the Bayesian clustering approach, as implemented in STRUCTURE v3.4 (Pritchard et al., 2000). Run length was set to 100,000 Markov chain Monte Carlo (MCMC) replicates after a burn-in period of 100,000 using correlated allele frequencies under an admixture model using the LOCPRIOR option. The number of clusters (K) was varied from 1 to 14, with 10 iterations of each. To infer the optimal K value, we used the ΔK method (Evanno et al., 2005), summarized using STRUCTURE HARVESTER (Earl & vonHoldt, 2012). We also plotted the log probability of the data to assess where $\ln \Pr(X|K)$ plateaued (Pritchard et al., 2000). Lastly, we visualized the output, using CLUMPAK (Kopelman et al., 2015). To further examine the population structure not revealed by neutral loci, we conducted a separate STRUCTURE analysis using only robust outlier SNPs and the same parameters as above. While usage of SNPs that are not in HWE violates one of the key assumptions of STRUCTURE, some studies suggest that population structure identified by outlier loci can help identify adaptive divergence patterns not found by analysis of neutral SNPs alone, but

which might be important for informing conservation management decisions (Funk et al., 2012; Milano et al., 2014; Schmidt et al., 2021).

Additionally, we examined population structure using TESS3 (Caye et al., 2016), which incorporates spatial information into the calculation of ancestry coefficients. We ran the program as implemented in the associated R package, using the default parameters and varying the number of clusters from 1 to 10. To choose the optimal number of clusters, we examined the cross-validation scores, to identify where the plot exhibited a plateau. To further evaluate population differentiation, we conducted a PCA using SNPRelate v1.14.0 (Zheng et al., 2012) on both neutral ($n=6517$) and robust outlier ($n=250$) SNPs and visualized the results using R packages *plotly* v4.9.0 (Sievert, 2020) and *ggplot2* v3.3.5 (Wickham et al., 2019).

3 | RESULTS

3.1 | Population genetics analyses

After full filtering, 112 individuals and 8302 SNPs were retained in the dataset, with an average depth of 55.16x and mean missing data percentage of 6.03%. Following the removal of unique outliers, no SNPs were out of HWE: therefore, 6517 putatively neutral SNPs were retained for population genetics analyses. Across all sampling locations, observed heterozygosity was lower than expected heterozygosity (Table 1) and ranged from 0.074 (VA) to 0.164 (ON). The overall inbreeding coefficient across populations was significantly different than 0 at 0.336 (95% CI: 0.329–0.342), ranging from 0.037 (ON) to 0.701 (VA) (Table 1). Based on calculations of pairwise genetic differentiation between sampling locations, WI population was most highly differentiated from all other sampling locations (Table 2).

The STRUCTURE analysis based on the putatively neutral SNPs revealed evidence for two clusters that best explained the genetic variation within the dataset (Figures 2a, S3). The first cluster was

Population	N	N_A	N_E	H_O	H_S	G_{IS}
AL	5	1.538	1.275	0.123	0.205	0.400 ($p < 0.01$)
GA	9	1.6	1.243	0.161	0.169	0.047 ($p < 0.01$)
IL-IN	10	1.883	1.333	0.153	0.24	0.362 ($p < 0.01$)
ME	8	1.387	1.195	0.075	0.141	0.465 ($p < 0.01$)
MO	10	1.83	1.276	0.149	0.202	0.263 ($p < 0.01$)
NC-VA	10	1.766	1.273	0.152	0.197	0.230 ($p < 0.01$)
NH	8	1.527	1.229	0.115	0.161	0.285 ($p < 0.01$)
OH	9	1.607	1.245	0.159	0.17	0.064 ($p < 0.01$)
ON	10	1.63	1.246	0.164	0.17	0.037 ($p < 0.01$)
PA	10	1.595	1.227	0.125	0.161	0.225 ($p < 0.01$)
TN	5	1.686	1.403	0.129	0.283	0.544 ($p < 0.01$)
VA	8	1.692	1.352	0.074	0.249	0.701 ($p < 0.01$)
WI	10	1.647	1.31	0.11	0.207	0.470 ($p < 0.01$)

TABLE 1 Population diversity metrics calculated for all sampling locations (populations) based on 6517 putatively neutral SNPs, as estimated using GenoDive v3.04 (Meirmans, 2020).

Abbreviations: G_{IS} , inbreeding coefficient; H_O , observed heterozygosity; H_S , expected heterozygosity; N , sample size; N_A , number of alleles; N_E , number of effective alleles.

TABLE 2 Estimates of pairwise θ (Weir & Cockerham, 1984) between all pairs of populations (above the diagonal), with their associated p -values (below the diagonal), as estimated using GenoDive v3.04 (Meirmans, 2020).

	AL	GA	IL-IN	ME	MO	NC-VA	NH	OH	ON	PA	TN	VA	WI
AL	-	0.025	-0.014	0.035	-0.005	0.001	0.02	0.032	0.031	0.025	0.058	-0.012	0.373
GA	0.002	-	0.03	0.039	0.01	0.002	0.014	0.02	0.016	0.012	0.178	0.121	0.458
IL-IN	0.505	0.128	-	0.031	-0.013	0.001	0.02	0.026	0.029	0.025	0.016	0.003	0.284
ME	0.01	0	0.218	-	0.013	0.024	0.018	0.04	0.026	0.028	0.175	0.12	0.473
MO	0.384	0	0.218	0.008	-	-0.008	0.001	0.01	0.006	0.005	0.091	0.056	0.378
NC-VA	0.145	0.16	0.211	0.001	0.475	-	0.007	0.012	0.012	0.006	0.116	0.074	0.399
NH	0.039	0	0.327	0	0.432	0.019	-	0.013	0.006	0.006	0.171	0.115	0.461
OH	0.001	0	0.204	0	0	0	0	-	0.012	0.007	0.177	0.123	0.459
ON	0	0	0	0	0	0	0.001	0	-	0.005	0.184	0.127	0.462
PA	0.023	0	0.225	0	0	0.001	0.009	0.001	0.001	-	0.179	0.121	0.465
TN	0.182	0.008	0.311	0.034	0.043	0.028	0.028	0.008	0.002	0.023	-	-0.049	0.106
VA	0.495	0	0.315	0.045	0.081	0.06	0.004	0	0	0.002	0.666	-	0.197
WI	0.003	0.001	0.004	0	0	0	0	0	0	0	0.086	0.014	-

mostly formed by WI samples, and included some individuals from TN, VA, IL-IN, MO, and NC-VA, while the second cluster included all the remaining populations. Further, STRUCTURE analysis based on the robust outlier SNPs showed highest support for $K=3$ (Figure S3), separating the southern populations (AL, GA, NC-VA, and MO) from the original cluster. While TESS3 analysis was inconclusive between $K=2$ and $K=3$, the clustering results revealed by TESS3 largely agreed with those found by STRUCTURE (Figure 2b). The PCA on the neutral dataset supported the results of the clustering programs, with the first two principal components separating WI from other populations in the dataset; PCA on the robust outlier SNPs revealed further divergence of the southern populations (Figure 3).

3.2 | Outlier detection and annotation of outlier SNPs

BayeScan detected 90 SNPs, while *pcadapt*, LFMM, and BayPass detected 884, 887, and 251 outlier SNPs, respectively (Table S2, Figure S1). All 90 outliers identified by BayeScan were under diversifying selection ($\alpha > 0$; Figure S2). LFMM identified 429 outliers associated with longitude, 402 with temperature, 183 with precipitation, and 150 with elevation: of these, 250 outliers were associated with two or more variables. Out of the 251 BayPass outliers, 50 were associated with longitude, 144 with temperature, 37 with precipitation, and 53 with elevation: 33 outliers were associated with two or more environmental variables. Combined, we detected 1784 unique outliers, of which 23 SNP was detected in common by all four methods, 32 SNPs were detected in at least three out of four methods, and 250 were detected by two or more methods, and thus considered robust (Table S2, Figure S1). Of the 250 robust outliers, 70 were associated with longitude by at least one of the GEA analyses, 171 were associated with temperature, 91 with precipitation, and 39 with elevation (Tables S2, S3). When comparing robust outliers by covariables, LFMM and BayPass shared 39 outliers associated with longitude, 134 outliers associated with temperature, 31 outliers associated with precipitation, and 17 outliers associated with elevation.

Annotation against the *C. calcarata* genome revealed that 250 robust outliers were mapped against 126 unique (143 total) gene identifiers, of which 97 (Table S3) were from genes of known function. Of the 134 temperature-associated SNPs detected by both LFMM and BayPass, annotations were available for 80 SNPs, including eye-specific diacylglycerol kinase (*rdgA*), a gene crucial for photoreceptor maintenance (Masai et al., 1997); multidrug resistance protein homolog 49 (*Mdr49*), which contributes to insecticide resistance; synaptic vesicle membrane protein VAT-1 homolog-like (*Vat1l*)—a gene expressed in brains of a variety of taxa from zebrafish to humans (Loeb-Hennard et al., 2004). Several outlier SNPs exhibited divergence patterns that could be linked to geographic and climatic factors. For example, one SNP (Scaffold_3_12741967; hereafter, the number after the underscore refers to the SNP position within the scaffold) that was detected as an outlier associated with longitude,

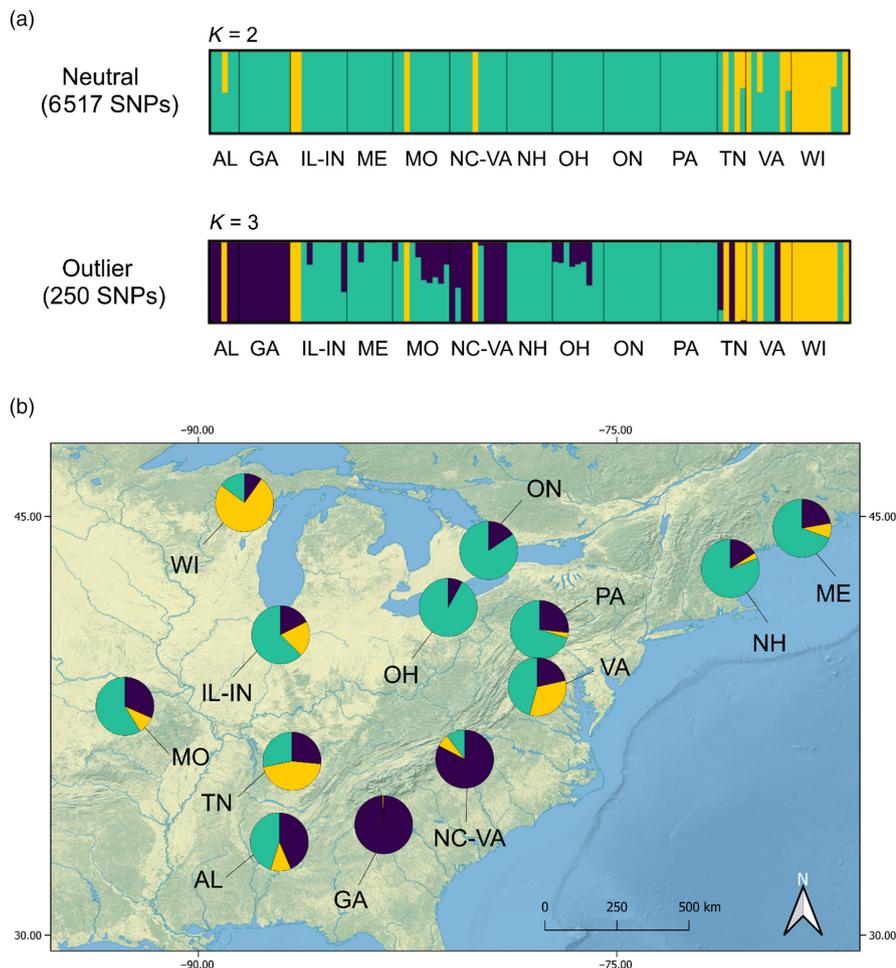


FIGURE 2 (a) Results of the Bayesian clustering method, as implemented in STRUCTURE v3.4 (Pritchard et al., 2000). Output results represent the optimal K value ($K=2$ or $K=3$), as determined by the ΔK method (Evanno et al., 2005), using STRUCTURE HARVESTER (Earl & vonHoldt, 2012). Visualized using CLUMPAK (Kopelman et al., 2015). (b) Results of the $K=3$ clustering generated by TESS3, based on 6517 putatively neutral SNPs. Map produced using QGIS.org (2020), QGIS Geographic Information System, Open Source Geospatial Foundation Project (<http://qgis.org>). Shape files retrieved from <https://www.naturalearthdata.com/>.

temperature, and elevation (Table S3) annotated to ATP-dependent 6-phosphofructokinase (*Pfk*), which is involved in glycometabolism (Leite et al., 1988); at this SNP, New Hampshire and Maine had the same genotype in contrast to all other populations. Furthermore, three outlier SNPs (Scaffold 1_18492518; Scaffold 1_18495934; Scaffold 1_18401015) found in association with longitude, temperature, and precipitation were located on Scaffold 1 and annotated to steroid receptor seven-up (*svp*) that has a role in eye development (Tsachaki & Sprecher, 2012), and *king-tubby*—a gene known for its function in retinal degeneration (Carbone et al., 2016). At two of these three SNPs, southern populations (Alabama, Georgia, North Carolina-Virginia, and Missouri) were differentiated from the rest of the populations. Lastly, across the 126 candidate genes, TopGO analysis identified 43 significantly enriched GO terms (Table S4), including those related to cellular response to UV, cellular response to light stimulus, and disruption by virus of host envelope lipopolysaccharide during virus entry.

4 | DISCUSSION

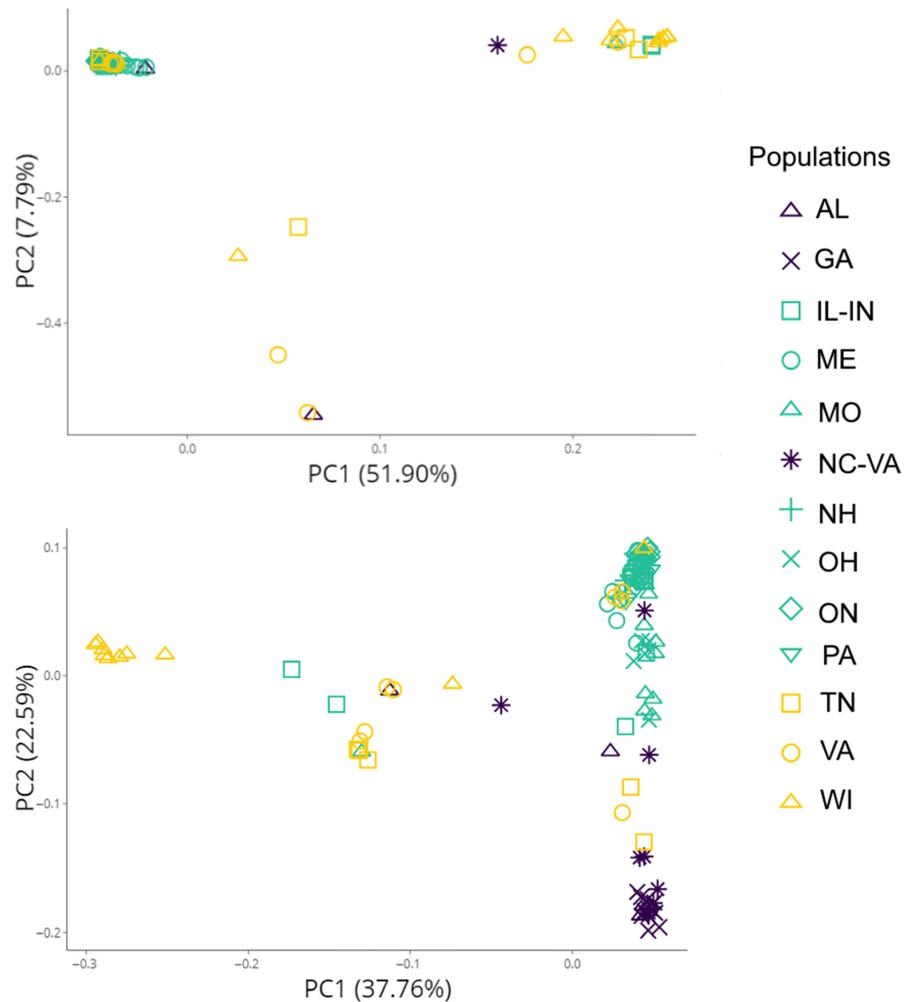
The loss of native pollinator abundance and diversity has wide-ranging impacts on natural ecosystems (Potts et al., 2010), agriculture, and economy (Garibaldi et al., 2013; Potts et al., 2010).

However, little is known about the genetics of wild bees and molecular mechanisms that enable their ability to adapt to environmental conditions. Here, we investigate the population genomics of an important native pollinator across its entire geographic distribution. The results of our data analysis demonstrate that *C. calcarata* population is separated into multiple genetic clusters, reflecting the geographical features of the species' range across eastern North America. Identification of putative signatures of selection revealed outlier loci involved in molecular processes potentially contributing to local adaptation, including thermoregulation, light sensitivity, stress response, and immunity.

4.1 | Patterns of genetic differentiation and inbreeding

The results of several population structure analyses conducted on a genome-wide set of SNPs were consistent with the presence of two to three genetic clusters within *C. calcarata*. The majority of the samples (Maine, New Hampshire, Ontario, Ohio, and Pennsylvania) belonged to the Northeastern genetic cluster; the Wisconsin samples formed a Western cluster, while Georgia and North Carolina-Southern Virginia samples comprised the Southern cluster (Figures 2, 3). Alabama, Illinois-Indiana, Tennessee, Missouri,

FIGURE 3 Principal component analysis (PCA) for 112 *Ceratina calcarata* samples across 13 populations, produced using 6517 putatively neutral SNPs (top panel) or 250 robust outlier SNPs (bottom panel). PCA analysis was conducted using SNPRelate v1.14.0 (Zheng et al., 2012).



and Virginia samples appeared to be highly admixed, and hence are not included as part of the three clusters.

We found strong evidence for differentiation of the Wisconsin population, as it was highly diverged from the rest of the samples (Table 2), forming a separate genetic cluster based on the analysis of both neutral and putatively adaptive SNPs (Figures 2, 3). The longitudinal differentiation was also reflected in the admixture shown by samples collected in the midwestern part of the study area (i.e., Illinois-Indiana, Missouri, Tennessee, and Virginia). This clustering suggests a West to East divergence of *C. calcarata*: a pattern also observed in other species that could be indicative of phylogeographical breaks (Soltis et al., 2006). For instance, molecular studies on plants propose that climatic fluctuations during the Pleistocene reduced the available range for many species, forcing them to occupy glacial refugia and subsequently diverge after the last glacial maximum ~20,000 years ago (Hewitt, 1996; Ony et al., 2021). Previous research has indicated a role for the Appalachian Mountains in postglaciation expansions: historically, the mountain range could have acted as a barrier to dispersal (Soltis et al., 2006). This discontinuity has been investigated for many animal and plant species (Austin, 2002; Church et al., 2003; Vickruck & Richards, 2017; Zamudio & Savage, 2003). For instance, in a generalist pollinator *Xylocopa virginica*, Iowa and Arkansas populations are highly differentiated from the core genetic

cluster located in the Eastern part of the continent (Vickruck & Richards, 2017). Several studies have investigated postglaciation divergence in North America and effects of climate on range expansions focusing on plants (Hoban et al., 2010; McLachlan et al., 2005). For example, in eastern redbud *Cercis canadensis* population, differentiation occurs along the Appalachian Mountains (Ony et al., 2021), a genetic split comparable to that exhibited by *X. virginica*. It is probable that other pollinator species have followed similar recolonization routes. A phylogeographical analysis of the five eastern North American *Ceratina* species showed that they have experienced a population expansion approximately 20,000 years ago coinciding with glacial retreat (Shell & Rehan, 2016). Following angiosperms that were able to occupy new suitable habitats as the ice gradually retreated (Hewitt, 1996), insect populations started a similar recolonization process (Shell & Rehan, 2016). The postglaciation refugia hypothesis has been examined on other continents: for example, a pollen-specialist bee *Colletes gigas* endemic to China was shown to exhibit fluctuations in effective population size during the last glacial maximum (Su et al., 2022).

Another signature of a recolonization process following glaciation is reduced genetic diversity along the dispersal routes and higher differentiation of the peripheral clusters. In accord with this prediction, *C. calcarata* exhibits high levels of inbreeding across

TABLE 3 A subset of robust outlier SNPs identified in this study and their associated annotations. For the complete list of SNPs ($n=250$) and annotations, see [Table S3](#).

	Scaffold	SNP	Gene ID	Gene description
Thermoregulation and energy metabolism	Scaffold_3	12,741,967	Ccalc.v3.03962	Similar to Pfk: ATP-dependent 6-phosphofructokinase (<i>Drosophila melanogaster</i>)
	Scaffold_7	3,528,284	Ccalc.v3.06845	Similar to Tret1: Facilitated trehalose transporter Tret1 (<i>Apis mellifera ligustica</i>)
	Scaffold_14	53,709	Ccalc.v3.13777	Similar to Atp2c1: Calcium-transporting ATPase type 2C member 1 (<i>Rattus norvegicus</i>)
	Scaffold_3	12,172,435	Ccalc.v3.03863	Similar to katnal1: Katanin p60 ATPase-containing subunit A-like 1 (<i>Danio rerio</i>)
	Scaffold_7	1,457,239	Ccalc.v3.06639	Similar to CG15743: Putative inositol monophosphatase 3 (<i>Drosophila melanogaster</i>)
Photoperiod	Scaffold_1	18,241,238	Ccalc.v3.02814	Similar to rdgA: Eye-specific diacylglycerol kinase (<i>Drosophila melanogaster</i>)
	Scaffold_1	18,492,518	Ccalc.v3.02835	Similar to svp: Steroid receptor seven-up, isoforms B/C (<i>Drosophila melanogaster</i>)
	Scaffold_1	18,401,015	Ccalc.v3.02826	Similar to king-tubby: Protein king tubby (<i>Drosophila persimilis</i>)
	Scaffold_3	816,030	Ccalc.v3.02948	Similar to UVOP: Opsin, ultraviolet-sensitive (<i>Apis mellifera</i>)
Response to stress	Scaffold_8	8,436,752	Ccalc.v3.10243	Similar to Mdr49: Multidrug resistance protein homolog 49 (<i>Drosophila melanogaster</i>)
Neural development	Scaffold_5	173,692	Ccalc.v3.08566	Similar to Vat1l: Synaptic vesicle membrane protein VAT-1 homolog-like (<i>Mus musculus</i>)
	Scaffold_1	18,444,442	Ccalc.v3.02831	Similar to OTOF: Otoferlin (<i>Homo sapiens</i>)

its entire geographic range ([Table 1](#)). According to the central-marginal hypothesis, populations located toward the periphery of the species' range exhibit lower levels of genetic diversity (Eckert et al., 2008). For example, increased differentiation and decreased genetic diversity in marginal populations have been observed in another *Ceratina* species, *C. australensis*, where levels of inbreeding were elevated in the recently established southern populations, but not in the more genetically diverse and older Queensland population (Harpur & Rehan, 2021). Similarly, we observe that the level of inbreeding in the Western (Wisconsin) cluster is higher ($G_{IS}=0.47$) than the average level of inbreeding in the Northeastern (mean $G_{IS}=0.30\pm 0.18$) or the Southern genetic cluster (mean $G_{IS}=0.14\pm 0.09$). However, given the variation in sample size between different locations in this study, this comparison should be interpreted with caution. Additionally, we did not find any evidence for substructure within the Northeastern cluster; despite this region spanning a large geographic distance, there was little genetic differentiation between sampling locations. This pattern could be attributed either to high gene flow between these locations or indicate that the founding population was characterized by low levels of genetic diversity. Bee foraging distance is correlated with body size (Gathmann & Tscharrntke, 2002): given the small body size of *C. calcarata*, it is unlikely that the absence of population structure in this region is due to recent gene flow and more probably attributed to the founder effect after glaciation. Finally, results of the genetic clustering analyses using robust outlier SNPs reveal additional patterns of differentiation providing support for the separation of the

Southern populations. This divergence might indicate that *C. calcarata* populations are undergoing local adaptation, reflective of current or historic climatic processes.

4.2 | Signatures of selection

Studies of local adaptation help to disentangle interactions between various evolutionary forces and environmental heterogeneity, contributing to our understanding of speciation and natural selection at a broader scale. These findings can be used by conservation biologists to identify current and future threats to economically important or critically endangered populations. Here, we used a genome scan approach to identify signatures of differential selection in a wild pollinator, focusing on its entire geographic distribution. We identified putatively divergent SNPs associated with population structure, as well as geographic and climatic variation. The results revealed several genes with potential relevance to bee physiology and response to external stressors, providing clues into gene-environment interactions in wild bees more generally. While this study represents a starting point in understanding the patterns of divergence in *C. calcarata*, we recognize that reduced representation sequencing methods (such as RAD-seq) capture only a fraction of the genome and likely underestimate the true number of divergent loci (Lowry et al., 2017). Thus, a more thorough understanding of the local adaptation in this species would require a follow-up whole-genome analysis.

In our study, we identified a number of putatively selected genes, potentially linked to thermal regulation (Tables 3, S3). For instance, one annotated gene (*pfk*) was related to glycolysis regulation, and more specifically to the phosphofructokinase/fructose-1,6-bisphosphatase-mediated substrate cycle (Leite et al., 1988). In bumblebees, this substrate cycle is critical for flight muscle activity, and is dependent on environmental temperature (Leite et al., 1988). Interestingly, at the SNP located on *pfk* gene, New Hampshire and Maine samples show similar genotypes. Within our dataset, Maine and New Hampshire are characterized by colder temperatures: this is in line with the prediction that selection on genes involved in thermoregulation will result in divergent genotypes in populations found in different climatic conditions. A study on bumblebees (Liu et al., 2020) found that *Pfk1* was among the 19 genes that were up-regulated in populations inhabiting high altitude conditions, further highlighting the role of this gene in cold tolerance. Another outlier SNP in our study annotated to facilitated trehalose transporter (*Tret1-like*), which is involved in the uptake of trehalose—sugar circulating in insect hemolymph—which plays a role in energy metabolism and stress recovery (Shukla et al., 2015). Additionally, one of the robust outlier SNPs annotated to inositol monophosphatase 3 (CG15743): in eukaryotic cells, inositol monophosphatase is an enzyme involved in myo-inositol synthesis (Michell, 2008). Myo-inositol has been shown to increase in overwintering *D. melanogaster*, suggesting the enzyme's role in cold resistance (Vesala et al., 2012). Previous studies found that genes involved in sugar metabolism and energy biosynthesis are differentially expressed in bees subject to extreme temperature conditions (Liu et al., 2020; Xu et al., 2017); hence, it is expected that these genes will show signatures of divergent selection in populations distributed across a climatic gradient. Also, of interest are outlier SNPs annotated to ATPase genes—katanin p60 ATPase-containing subunit A-like 1 (*katnal1*) and calcium-transporting ATPase type 2C member 1 (*Atp2c1*), which have been associated with climate adaptations (Eydivandi et al., 2021; Yu et al., 2021). While most insects are ectotherms, many, including bees, exhibit thermoregulatory behavior, making them facultatively endothermic (Heinrich, 1975; Stone & Willmer, 1989). This ability makes it possible for bees to function under different climatic conditions, inhabiting large geographic regions (Heinrich, 1975). Thermoregulation in bees is critical for flight, efficient foraging, and young rearing (Heinrich, 1975; Kovac & Schmaranzer, 1996; Somanathan et al., 2019); therefore, selection is predicted to act on genes that are tightly linked to functions associated with thermal responses. Two temperature-associated outlier SNPs annotated to genes with a role in neural development (*Vat1l*, *OTOF*); furthermore, one of the significantly enriched GO terms (GO:0007616) was associated with long-term memory. These findings are consistent with previous research that revealed genes involved in neural and neuromuscular processes to be under selection in other bee species, across temperature (Jackson et al., 2020), and elevation gradients (Wallberg et al., 2017).

In our study, we also identified several genes (*svp*, *king-tubby*, and *rdgA*), which play a role in the development of photoreceptor cells

(Tables 3, S3). In a study on *D. melanogaster* senescence, *king-tubby* was among the top candidate genes that promote visual decline (Carbone et al., 2016). Selection on genes that regulate eye development could be driven by variation in daylight amount among various locations in our study. For instance, our results show Alabama, Georgia, North Carolina, and Missouri populations show similar genotypes at genes associated with eye development—this could be explained by a difference in photoperiod length experienced by populations at Southern latitudes. Together with temperature, photoperiod is known to impact many critical bee life-history traits, such as the timing of brood onset (Bennett et al., 2018; Nürnbergger et al., 2018) and diapause regulation (Pitts-Singer, 2020).

In addition to genes linked to thermal regulation and light sensitivity, we also found outlier SNPs putatively associated with adaptation to various biotic and abiotic threats. Of the 43 significantly enriched GO terms, three (GO:0034644: cellular response to UV; GO:0071482: cellular response to light stimulus; GO:0009416: response to light stimulus) were linked to biological processes associated with light responses. Exposure to ultraviolet (UV) light constitutes a type of oxidative stress; in honeybees, nurse bees subject to UV radiation show an elevated expression of heat shock proteins in abdomens and heads, which is a marker typically indicative of thermal stress (Kim et al., 2019). Likewise, in a solitary bee *Osmia bicornis*, exposure to UV light was found to cause body deformities and increased mortality (Wasielewski et al., 2015). Another outlier SNP annotated to multidrug resistance protein homolog 49 (*Mdr49*), which could be a signature of selection related to increased pesticide usage. Research has shown that interaction with certain compounds commonly found in bee habitats (e.g., fumagillin), can inhibit multidrug resistance transporters, increasing honeybee sensitivity to ivermectin and acetamiprid (Levy & Marshall, 2013). Finally, one of the significantly enriched GO terms (GO:0098995: disruption by virus of host envelope lipopolysaccharide during virus entry) was related to immune function. Both managed and wild bees are highly susceptible to pathogenic infections (Mallinger et al., 2017), making immune genes a highly probable target of selection pressure. This finding is congruent with previous studies that detected immune-related genes showing signatures of selection in other bee species, such as *A. mellifera iberiensis* (Chávez-Galarza et al., 2013) and *B. terrestris* (Kent et al., 2018).

5 | CONCLUSION

Combined, our data show how geographic and climatic features shape genetic structure and diversity of natural populations and provide important insights into mechanisms by which native pollinator insects adapt to environmental conditions, using *C. calcarata* as an example. We demonstrate that despite its widespread occurrence, this wild pollinator exhibits decreased heterozygosity and elevated levels of inbreeding. Furthermore, our results provide evidence that patterns of neutral divergence in *C. calcarata* are associated with longitude, forming two main genetic clusters; at the

same time, evaluation of adaptive divergence reveals a third genetic cluster in the Southern part of the continent. Analysis of signatures of selection shows that *C. calcarata* is undergoing geography- and climate-associated adaptation, with molecular processes linked to thermoregulation, abiotic and biotic stress, and light sensitivity being under putatively divergent selection. Finally, our work highlights the utility of a small carpenter bee as a potential model species for evaluating the interactions between genetics and environment. As anthropogenic disturbances and global warming continue to modify critical pollinator habitats, wild bees are going to be disproportionately affected. Continuous monitoring of the genetic diversity, structure and adaptive processes of vulnerable populations is going to be vital in determining the resilience potential of bee communities as a whole.

ACKNOWLEDGMENTS

We thank members of the Rehan laboratory for constructive feedback on earlier versions of this manuscript. Data analysis was enabled by support provided by the high-performance computing (HPC) clusters from the Digital Research Alliance of Canada. This work was funded by NSF Biological Collections Postdoctoral Fellowship DBI-1906494 to EPK and NSERC Discovery Grant, Supplement, and EWR Steacie Memorial Fellowship to SMR.

DATA AVAILABILITY STATEMENT

Sequence data are publicly available in NCBI, BioProject: PRJNA791561. Reference genome, annotation, and gene ontology files used in this study are available at <http://www.rehanlab.com/genomes.html>.

ORCID

Evan P. Kelemen  <https://orcid.org/0000-0003-3600-9878>

Sandra M. Rehan  <https://orcid.org/0000-0002-6441-5155>

REFERENCES

- Aldercotte, A. H., Simpson, D. T., & Winfree, R. (2022). Crop visitation by wild bees declines over an 8-year time series: A dramatic trend, or just dramatic between-year variation? *Insect Conservation and Diversity*, 15(5), 522–533. <https://doi.org/10.1111/icad.12589>
- Alexa, A., & Rahnenfuhrer, J. (2022). topGO: Enrichment analysis for gene ontology. R Package Version 2.48.0.
- Allendorf, F. W., Hohenlohe, P. A., & Luikart, G. (2010). Genomics and the future of conservation genetics. *Nature Reviews Genetics*, 11(10), 697–709. <https://doi.org/10.1038/nrg2844>
- Alves-Pereira, A., Clement, C. R., Picanço-Rodrigues, D., Veasey, E. A., Dequigiovanni, G., Ramos, S. L. F., Pinheiro, J. B., de Souza, A. P., & Zucchi, M. I. (2020). A population genomics appraisal suggests independent dispersals for bitter and sweet manioc in Brazilian Amazonia. *Evolutionary Applications*, 13(2), 342–361. <https://doi.org/10.1111/eva.12873>
- Austin, J. (2002). Cryptic lineages in a small frog: The post-glacial history of the spring peeper, *Pseudacris crucifer* (Anura: Hylidae). *Molecular Phylogenetics and Evolution*, 25(2), 316–329. [https://doi.org/10.1016/S1055-7903\(02\)00260-9](https://doi.org/10.1016/S1055-7903(02)00260-9)
- Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., Selker, E. U., Cresko, W. A., & Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One*, 3(10), e3376. <https://doi.org/10.1371/journal.pone.0003376>
- Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*, 29(4), 1165–1188.
- Bennett, M. M., Rinehart, J. P., Yocum, G. D., Doetkott, C., & Greenlee, K. J. (2018). Cues for cavity nesters: Investigating relevant zeitgebers for emerging leafcutting bees, *Megachile rotundata*. *Journal of Experimental Biology*, 221(10), jeb175406. <https://doi.org/10.1242/jeb.175406>
- Cameron, S. A., & Sadd, B. M. (2020). Global trends in bumble bee health. *Annual Review of Entomology*, 65(1), 209–232. <https://doi.org/10.1146/annurev-ento-011118-111847>
- Carbone, M. A., Yamamoto, A., Huang, W., Lyman, R. A., Meadors, T. B., Yamamoto, R., Anholt, R. R., & Mackay, T. F. (2016). Genetic architecture of natural variation in visual senescence in *Drosophila*. *Proceedings of the National Academy of Sciences*, 113(43), E6620–E6629. <https://doi.org/10.1073/pnas.1613833113>
- Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: Building and genotyping loci *De novo* from short-read sequences. *Genes*(Genomes)Genetics, 1(3), 171–182. <https://doi.org/10.1534/g3.111.000240>
- Caye, K., Deist, T. M., Martins, H., Michel, O., & François, O. (2016). TESS3: Fast inference of spatial population structure and genome scans for selection. *Molecular Ecology Resources*, 16, 540–548. <https://doi.org/10.1111/1755-0998.12471>
- Chávez-Galarza, J., Henriques, D., Johnston, J. S., Azevedo, J. C., Patton, J. C., Muñoz, I., de la Rúa, P., & Pinto, M. A. (2013). Signatures of selection in the Iberian honey bee (*Apis mellifera iberiensis*) revealed by a genome scan analysis of single nucleotide polymorphisms. *Molecular Ecology*, 22, 5890–5907.
- Chevin, L.-M., Lande, R., & Mace, G. M. (2010). Adaptation, plasticity, and extinction in a changing environment: Towards a predictive theory. *PLoS Biology*, 8(4), e1000357. <https://doi.org/10.1371/journal.pbio.1000357>
- Church, S. A., Kraus, J. M., Mitchell, J. C., Church, D. R., & Taylor, D. R. (2003). Evidence for multiple Pleistocene refugia in the postglacial expansion of the eastern Tiger salamander, *Ambystoma tigrinum*. *Evolution*, 57, 372–383.
- Cridland, J. M., Ramirez, S. R., Dean, C. A., Sciligo, A., & Tsutsui, N. D. (2018). Genome sequencing of museum specimens reveals rapid changes in the genetic composition of honey bees in California. *Genome Biology and Evolution*, 10(2), 458–472. <https://doi.org/10.1093/gbe/evy007>
- 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., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>
- Danforth, B. N., Minckley, R. L., Neff, J. L., & Fawcett. (2019). *The solitary bees: Biology, evolution, conservation*. Princeton University Press. <https://doi.org/10.2307/j.ctvd1c929>
- Deagle, B. E., Jones, F. C., Chan, Y. F., Absher, D. M., Kingsley, D. M., & Reimchen, T. E. (2012). Population genomics of parallel phenotypic evolution in stickleback across stream-lake ecological transitions. *Proceedings of the Royal Society B: Biological Sciences*, 279(1732), 1277–1286. <https://doi.org/10.1098/rspb.2011.1552>
- Earl, D. A., & von Holdt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4(2), 359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- Eckert, C. G., Samis, K. E., & Loughheed, 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>

- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. *Molecular Ecology*, 14(8), 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Eydivandi, S., Roudbar, M. A., Karimi, M. O., & Sahana, G. (2021). Genomic scans for selective sweeps through haplotype homozygosity and allelic fixation in 14 indigenous sheep breeds from Middle East and South Asia. *Scientific Reports*, 11(1), 2834. <https://doi.org/10.1038/s41598-021-82625-2>
- 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>
- Fraser, B. A., & Whiting, J. R. (2020). What can be learned by scanning the genome for molecular convergence in wild populations? *Annals of the New York Academy of Sciences*, 1476(1), 23–42. <https://doi.org/10.1111/nyas.14177>
- Freitas, F. V., Branstetter, M. G., Griswold, T., & Almeida, E. A. B. (2021). Partitioned gene-tree analyses and gene-based topology testing help resolve incongruence in a phylogenomic study of host-specialist bees (Apidae: Eucerinae). *Molecular Biology and Evolution*, 38(3), 1090–1100. <https://doi.org/10.1093/molbev/msaa277>
- Frichot, E., & François, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, 6, 925–929.
- Funk, W. C., McKay, J. K., Hohenlohe, P. A., & Allendorf, F. W. (2012). Harnessing genomics for delineating conservation units. *Trends in Ecology & Evolution*, 27(9), 489–496. <https://doi.org/10.1016/j.tree.2012.05.012>
- Garibaldi, L. A., Steffan-Dewenter, I., Winfree, R., Aizen, M. A., Bommarco, R., Cunningham, S. A., Kremen, C., Carvalheiro, L. G., Harder, L. D., Afik, O., Bartomeus, I., Benjamin, F., Boreux, V., Cariveau, D., Chacoff, N. P., Dudenhöffer, J. H., Freitas, B. M., Ghazoul, J., Greenleaf, S., ... Klein, A. M. (2013). Wild pollinators enhance fruit set of crops regardless of honey bee abundance. *Science*, 339(6127), 1608–1611. <https://doi.org/10.1126/science.1230200>
- Gathmann, A., & Tscharrnke, T. (2002). Foraging ranges of solitary bees. *Journal of Animal Ecology*, 71(5), 757–764. <https://doi.org/10.1046/j.1365-2656.2002.00641.x>
- Gautier, M. (2015). Genome-wide scan for adaptive divergence and association with population-specific covariates. *Genetics*, 201, 1555–1579.
- Gotz, S., Garcia-Gomez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., Robles, M., Talon, M., Dopazo, J., & Conesa, A. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Research*, 36(10), 3420–3435. <https://doi.org/10.1093/nar/gkn176>
- Grozinger, C. M., Fan, Y., Hoover, S. E. R., & Winston, M. L. (2007). Genome-wide analysis reveals differences in brain gene expression patterns associated with caste and reproductive status in honey bees (*Apis mellifera*). *Molecular Ecology*, 16(22), 4837–4848. <https://doi.org/10.1111/j.1365-294X.2007.03545.x>
- Hadley, A. S., & Betts, M. G. (2012). The effects of landscape fragmentation on pollination dynamics: Absence of evidence not evidence of absence. *Biological Reviews*, 87(3), 526–544. <https://doi.org/10.1111/j.1469-185X.2011.00205.x>
- Harpur, B. A., Dey, A., Albert, J. R., Patel, S., Hines, H. M., Hasselmann, M., Packer, L., & Zayed, A. (2017). Queens and workers contribute differently to adaptive evolution in bumble bees and honey bees. *Genome Biology and Evolution*, 9(9), 2395–2402. <https://doi.org/10.1093/gbe/evx182>
- Harpur, B. A., Kent, C. F., Molodtsova, D., Lebon, J. M. D., Alqarni, A. S., Owayss, A. A., & Zayed, A. (2014). Population genomics of the honey bee reveals strong signatures of positive selection on worker traits. *Proceedings of the National Academy of Sciences*, 111(7), 2614–2619. <https://doi.org/10.1073/pnas.1315506111>
- Harpur, B. A., & Rehan, S. M. (2021). Connecting social polymorphism to single nucleotide polymorphism: Population genomics of the small carpenter bee, *Ceratina australensis*. *Biological Journal of the Linnean Society*, 132(4), 945–954. <https://doi.org/10.1093/biolinnean/blab003>
- Hedrick, P. W. (2001). Conservation genetics: Where are we now? *Trends in Ecology & Evolution*, 16(11), 629–636. [https://doi.org/10.1016/S0169-5347\(01\)02282-0](https://doi.org/10.1016/S0169-5347(01)02282-0)
- Heinrich, B. (1975). Thermoregulation in bumblebees. *Journal of Comparative Physiology*, 96, 155–166.
- Henriques, D., Wallberg, A., Chávez-Galarza, J., Johnston, J. S., Webster, M. T., & Pinto, M. A. (2018). Whole genome SNP-associated signatures of local adaptation in honeybees of the Iberian Peninsula. *Scientific Reports*, 8(1), 11145. <https://doi.org/10.1038/s41598-018-29469-5>
- Hewitt, G. M. (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, 58, 247–276.
- Hoban, S. M., Borkowski, D. S., Brosi, S. L., McCleary, T. S., Thompson, L. M., McLachlan, J. S., & Romero-Severson, J. (2010). Range-wide distribution of genetic diversity in the north American tree *Juglans cinerea*: A product of range shifts, not ecological marginality or recent population decline. *Molecular Ecology*, 19(22), 4876–4891. <https://doi.org/10.1111/j.1365-294X.2010.04834.x>
- Hohenlohe, P. A., Funk, W. C., & Rajora, O. P. (2021). Population genomics for wildlife conservation and management. *Molecular Ecology*, 30(1), 62–82. <https://doi.org/10.1111/mec.15720>
- Holderregger, R., Kamm, U., & Gugerli, F. (2006). Adaptive vs. neutral genetic diversity: Implications for landscape genetics. *Landscape Ecology*, 21, 797–807.
- 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>
- Jacobson, M. M., Tucker, E. M., Mathiasson, M. E., & Rehan, S. M. (2018). Decline of bumble bees in northeastern North America, with special focus on *Bombus terricola*. *Biological Conservation*, 217, 437–445. <https://doi.org/10.1016/j.biocon.2017.11.026>
- Jaffé, R., Veiga, J. C., Pope, N. S., Lanes, É. C. M., Carvalho, C. S., Alves, R., Andrade, S. C. S., Arias, M. C., Bonatti, V., Carvalho, A. T., de Castro, M. S., Contrera, F. A. L., Franco, T. M., Freitas, B. M., Giannini, T. C., Hrcir, M., Martins, C. F., Oliveira, G., Saraiva, A. M., ... Imperatriz-Fonseca, V. L. (2019). Landscape genomics to the rescue of a tropical bee threatened by habitat loss and climate change. *Evolutionary Applications*, 12(6), 1164–1177. <https://doi.org/10.1111/eva.12794>
- Kelemen, E. P., & Rehan, S. M. (2021). Conservation insights from wild bee genetic studies: Geographic differences, susceptibility to inbreeding, and signs of local adaptation. *Evolutionary Applications*, 14(6), 1485–1496. <https://doi.org/10.1111/eva.13221>
- Kennedy, C. M., Lonsdorf, E., Neel, M. C., Williams, N. M., Ricketts, T. H., Winfree, R., Bommarco, R., Brittain, C., Burley, A. L., Cariveau, D., Carvalheiro, L. G., Chacoff, N. P., Cunningham, S. A., Danforth, B. N., Dudenhöffer, J. H., Elle, E., Gaines, H. R., Garibaldi, L. A., Gratton, C., ... Kremen, C. (2013). A global quantitative synthesis of local and landscape effects on wild bee pollinators in agroecosystems. *Ecology Letters*, 16(5), 584–599. <https://doi.org/10.1111/ele.12082>
- Kent, C. F., Dey, A., Patel, H., Tsvetkov, N., Tiwari, T., MacPhail, V. J., Gobeil, Y., Harpur, B. A., Gurtowski, J., Schatz, M. C., Colla, S. R., & Zayed, A. (2018). Conservation genomics of the declining north American bumblebee *Bombus terricola* reveals inbreeding and selection on immune genes. *Frontiers in Genetics*, 9, 316. <https://doi.org/10.3389/fgene.2018.00316>
- Kim, S., Kim, K., Lee, J. H., Han, S. H., & Lee, S. H. (2019). Differential expression of acetylcholinesterase 1 in response to various stress

- factors in honey bee workers. *Scientific Reports*, 9(1), 10342. <https://doi.org/10.1038/s41598-019-46842-0>
- Kleijn, D., Winfree, R., Bartomeus, I., Carvalho, L. G., Henry, M., Isaacs, R., Klein, A. M., Kremen, C., M'Gonigle, L. K., Rader, R., Ricketts, T. H., Williams, N. M., Lee Adamson, N., Ascher, J. S., Báldi, A., Batáry, P., Benjamin, F., Biesmeijer, J. C., Blitzer, E. J., ... Potts, S. G. (2015). Delivery of crop pollination services is an insufficient argument for wild pollinator conservation. *Nature Communications*, 6(1), 7414. <https://doi.org/10.1038/ncomms8414>
- Klein, A.-M., Vaissière, B. E., Cane, J. H., Steffan-Dewenter, I., Cunningham, S. A., Kremen, C., & Tscharntke, T. (2007). Importance of pollinators in changing landscapes for world crops. *Proceedings of the Royal Society B: Biological Sciences*, 274(1608), 303–313. <https://doi.org/10.1098/rspb.2006.3721>
- Kline, O., & Joshi, N. K. (2020). Mitigating the effects of habitat loss on solitary bees in agricultural ecosystems. *Agriculture*, 10(4), 115. <https://doi.org/10.3390/agriculture10040115>
- Kocher, S. D., Li, C., Yang, W., Tan, H., Yi, S. V., Yang, X., Hoekstra, H. E., Zhang, G., Pierce, N. E., & Yu, D. W. (2013). The draft genome of a socially polymorphic halictid bee, *Lasioglossum albipes*. *Genome Biology*, 14(12), R142. <https://doi.org/10.1186/gb-2013-14-12-r142>
- Koh, I., Lonsdorf, E. V., Williams, N. M., Brittain, C., Isaacs, R., Gibbs, J., & Ricketts, T. H. (2016). Modeling the status, trends, and impacts of wild bee abundance in the United States. *Proceedings of the National Academy of Sciences*, 113(1), 140–145. <https://doi.org/10.1073/pnas.1517685113>
- Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, 15(5), 1179–1191. <https://doi.org/10.1111/1755-0998.12387>
- Kovac, H., & Schmaranzer, S. (1996). Thermoregulation of honeybees (*Apis mellifera*) foraging in spring and summer at different plants. *Journal of Insect Physiology*, 42(11–12), 1071–1076. [https://doi.org/10.1016/S0022-1910\(96\)00061-3](https://doi.org/10.1016/S0022-1910(96)00061-3)
- Leite, A., Neto, J. A., Leyton, J. F., Crivellaro, O., & El-Dorri, H. A. (1988). Phosphofructokinase from bumblebee flight muscle. Molecular and catalytic properties and role of the enzyme in regulation of the fructose 6-phosphate/fructose 1, 6-bisphosphate cycle. *Journal of Biological Chemistry*, 263, 17527–17533.
- Levy, S. B., & Marshall, B. M. (2013). Honeybees and tetracycline resistance. *MBio*, 4(1), e00045-13. <https://doi.org/10.1128/mBio.00045-13>
- 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., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Liu, Y., Jin, H., Naeem, M., & An, J. (2020). Comparative transcriptome analysis reveals regulatory genes involved in cold tolerance and hypoxic adaptation of high-altitude Tibetan bumblebees. *Apidologie*, 51(6), 1166–1181. <https://doi.org/10.1007/s13592-020-00795-w>
- Loeb-Hennard, C., Cousin, X., Pregel, I., & Kremmer, E. (2004). Cloning and expression pattern of vat-1 homolog gene in zebrafish. *Gene Expression Patterns*, 5(1), 91–96. <https://doi.org/10.1016/j.modgep.2004.06.002>
- Lowenstein, D. M., Matteson, K. C., & Minor, E. S. (2015). Diversity of wild bees supports pollination services in an urbanized landscape. *Oecologia*, 179(3), 811–821. <https://doi.org/10.1007/s00442-015-3389-0>
- Lowry, D. B., Hoban, S., Kelley, J. L., Lotterhos, K. E., Reed, L. K., Antolin, M. F., & Storfer, A. (2017). Breaking RAD: An evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. *Molecular Ecology Resources*, 17(2), 142–152. <https://doi.org/10.1111/1755-0998.12635>
- Luu, K., Bazin, E., & Blum, M. G. B. (2017). *Pcadapt*: An R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, 17(1), 67–77. <https://doi.org/10.1111/1755-0998.12592>
- Mallinger, R. E., Gaines-Day, H. R., & Gratton, C. (2017). Do managed bees have negative effects on wild bees?: A systematic review of the literature. *PLoS One*, 12(12), e0189268. <https://doi.org/10.1371/journal.pone.0189268>
- Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: Combining landscape ecology and population genetics. *Trends in Ecology & Evolution*, 18(4), 189–197. [https://doi.org/10.1016/S0169-5347\(03\)00008-9](https://doi.org/10.1016/S0169-5347(03)00008-9)
- Manichaikul, A., Mychalek, J. C., Rich, S. S., Daly, K., Sale, M., & Chen, W.-M. (2010). Robust relationship inference in genome-wide association studies. *Bioinformatics*, 26(22), 2867–2873. <https://doi.org/10.1093/bioinformatics/btq559>
- Masai, I., Suzuki, E., Yoon, C. S., Kohyama, A., & Hotta, Y. (1997). Immunolocalization of drosophila eye-specific diacylglycerol kinase, rdgA, which is essential for the maintenance of the photoreceptor. *Journal of Neurobiology*, 32(7), 695–706.
- McLachlan, J. S., Clark, J. S., & Manos, P. S. (2005). Molecular indicators of tree migration capacity under rapid climate change. *Ecology*, 86(8), 2088–2098. <https://doi.org/10.1890/04-1036>
- Meirmans, P. G. (2020). Genodive version 3.0: Easy-to-use software for the analysis of genetic data of diploids and polyploids. *Molecular Ecology Resources*, 20, 1126–1131. <https://doi.org/10.1111/1755-0998.13145>
- Michell, R. H. (2008). Inositol derivatives: Evolution and functions. *Nature Reviews Molecular Cell Biology*, 9(2), 151–161. <https://doi.org/10.1038/nrm2334>
- Michener, C. D. (2007). *The bees of the world* (2nd ed.). John Hopkins University Press.
- Mikheyev, A. S., Tin, M. M. Y., Arora, J., & Seeley, T. D. (2015). Museum samples reveal rapid evolution by wild honey bees exposed to a novel parasite. *Nature Communications*, 6(1), 7991. <https://doi.org/10.1038/ncomms8991>
- Milano, I., Babbucci, M., Cariani, A., Atanassova, M., Bekkevold, D., Carvalho, G. R., ... Bargelloni, L. (2014). Outlier SNP markers reveal fine-scale genetic structuring across European hake populations (*Merluccius merluccius*). *Molecular Ecology*, 23(1), 118–135. <https://doi.org/10.1111/mec.12568>
- Narum, S. R., & Hess, J. E. (2011). Comparison of FST outlier tests for SNP loci under selection. *Molecular Ecology Resources*, 11, 184–194.
- Neph, S., Kuehn, M. S., Reynolds, A. P., Haugen, E., Thurman, R. E., Johnson, A. K., Rynes, E., Maurano, M. T., Vierstra, J., Thomas, S., Sandstrom, R., Humbert, R., & Stamatoyannopoulos, J. A. (2012). BEDOPS: High-performance genomic feature operations. *Bioinformatics*, 28(14), 1919–1920. <https://doi.org/10.1093/bioinformatics/bts277>
- Nooten, S. S., Odanaka, K. A., & Rehan, S. M. (2020). Effects of farmland and seasonal phenology on wild bees in blueberry orchards. *Northeastern Naturalist*, 27(4), 841–860. <https://doi.org/10.1656/045.027.0420>
- Nooten, S. S., & Rehan, S. M. (2020). Historical changes in bumble bee body size and range shift of declining species. *Biodiversity and Conservation*, 29, 451–467.
- Nürnberg, F., Härtel, S., & Steffan-Dewenter, I. (2018). The influence of temperature and photoperiod on the timing of brood onset in hibernating honey bee colonies. *PeerJ*, 6, e4801. <https://doi.org/10.7717/peerj.4801>
- Ollerton, J., Winfree, R., & Tarrant, S. (2011). How many flowering plants are pollinated by animals? *Oikos*, 120(3), 321–326. <https://doi.org/10.1111/j.1600-0706.2010.18644.x>
- Ony, M., Klingeman, W. E., Zobel, J., Trigiano, R. N., Ginzel, M., Nowicki, M., Boggess, S. L., Everhart, S., & Hadziabdic, D. (2021). Genetic

- diversity in north American *Cercis Canadensis* reveals an ancient population bottleneck that originated after the last glacial maximum. *Scientific Reports*, 11(1), 21803. <https://doi.org/10.1038/s41598-021-01020-z>
- Otto, S. P. (2018). Adaptation, speciation and extinction in the Anthropocene. *Proceedings of the Royal Society B: Biological Sciences*, 285(1891), 20182047. <https://doi.org/10.1098/rspb.2018.2047>
- Pimsler, M. L., Jackson, J. M., & Lozier, J. D. (2017). Population genomics reveals a candidate gene involved in bumble bee pigmentation. *Ecology and Evolution*, 7(10), 3406–3413. <https://doi.org/10.1002/ece3.2935>
- Pitts-Singer, T. L. (2020). Photoperiod effect on *Megachile rotundata* (Hymenoptera: Megachilidae) female regarding diapause status of progeny: The importance of data scrutiny. *Environmental Entomology*, 49(2), 516–527. <https://doi.org/10.1093/ee/nvaa004>
- Potts, S. G., Biesmeijer, J. C., Kremen, C., Neumann, P., Schweiger, O., & Kunin, W. E. (2010). Global pollinator declines: Trends, impacts and drivers. *Trends in Ecology & Evolution*, 25(6), 345–353. <https://doi.org/10.1016/j.tree.2010.01.007>
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155, 945–959.
- Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, 26(6), 841–842. <https://doi.org/10.1093/bioinformatics/btq033>
- Rehan, S. M., Berens, A. J., & Toth, A. L. (2014). At the brink of eusociality: Transcriptomic correlates of worker behaviour in a small carpenter bee. *BMC Evolutionary Biology*, 14(1), 260. <https://doi.org/10.1186/s12862-014-0260-6>
- Rehan, S. M., Glastad, K. M., Lawson, S. P., & Hunt, B. G. (2016). The genome and methylome of a subsocial small carpenter bee, *Ceratina calcarata*. *Genome Biology and Evolution*, 8(5), 1401–1410. <https://doi.org/10.1093/gbe/evw079>
- Rehan, S. M., & Richards, M. H. (2010). Nesting biology and subsociality in *Ceratina calcarata* (Hymenoptera: Apidae). *The Canadian Entomologist*, 142(1), 65–74. <https://doi.org/10.4039/n09-056>
- Rehan, S. M., & Sheffield, C. S. (2011). Morphological and molecular delineation of a new species in the *Ceratina* dupla species-group (Hymenoptera: Apidae: Xylocopinae) of eastern North America. *Zootaxa*, 2873(1), 35. <https://doi.org/10.11646/zootaxa.2873.1.3>
- Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, 28, 4737–4754. <https://doi.org/10.1111/mec.15253>
- Russello, M. A., Kirk, S. L., Frazer, K. K., & Askey, P. J. (2012). Detection of outlier loci and their utility for fisheries management: Outlier loci for fisheries management. *Evolutionary Applications*, 5(1), 39–52. <https://doi.org/10.1111/j.1752-4571.2011.00206.x>
- Russo, L., DeBarros, N., Yang, S., Yang, K., & Mortensen, D. (2013). Supporting crop pollinators with floral resources: Networkbased phenological matching. *Ecology and Evolution*, 3(9), 3125–3140.
- Schmidt, D. A., Waterhouse, M. D., Sjodin, B. M. F., & Russello, M. A. (2021). Genome-wide analysis reveals associations between climate and regional patterns of adaptive divergence and dispersal in American pikas. *Heredity*, 127(5), 443–454. <https://doi.org/10.1038/s41437-021-00472-3>
- Shell, W. A., & Rehan, S. M. (2016). Recent and rapid diversification of the small carpenter bees in eastern North America. *Biological Journal of the Linnean Society*, 117(3), 633–645. <https://doi.org/10.1111/bj.12692>
- Shukla, E., Thorat, L. J., Nath, B. B., & Gaikwad, S. M. (2015). Insect trehalase: Physiological significance and potential applications. *Glycobiology*, 25(4), 357–367. <https://doi.org/10.1093/glycob/cwu125>
- Sievert, C. (2020). *Interactive web-based data visualization with R, plotly, and shiny*. Chapman and Hall/CRC.
- Simpson, D. T., Weinman, L. R., Genung, M. A., Roswell, M., MacLeod, M., & Winfree, R. (2022). Many bee species, including rare species, are important for function of entire plant–pollinator networks. *Proceedings of the Royal Society B: Biological Sciences*, 289(1972), 20212689. <https://doi.org/10.1098/rspb.2021.2689>
- Soltis, D. E., Morris, A. B., McLachlan, J. S., Manos, P. S., & Soltis, P. S. (2006). Comparative phylogeography of unglaciated eastern North America: Phylogeography of unglaciated eastern North America. *Molecular Ecology*, 15(14), 4261–4293. <https://doi.org/10.1111/j.1365-294X.2006.03061.x>
- Somanathan, H., Saryan, P., & Balamurali, G. S. (2019). Foraging strategies and physiological adaptations in large carpenter bees. *Journal of Comparative Physiology A*, 205(3), 387–398. <https://doi.org/10.1007/s00359-019-01323-7>
- Stone, G. N., & Willmer, P. G. (1989). Warm-up rates and body temperatures in bees: The importance of body size, thermal regime and phylogeny. *Journal of Experimental Biology*, 147, 303–328.
- Su, T., He, B., Zhao, F., Jiang, K., Lin, G., & Huang, Z. (2022). Population genomics and phylogeography of *Colletes gigas*, a wild bee specialized on winter flowering plants. *Ecology and Evolution*, 12(4), e8863. <https://doi.org/10.1002/ece3.8863>
- Theodorou, P., Radzevičiūtė, R., Kahnt, B., Soro, A., Grosse, I., & Paxton, R. J. (2018). Genome-wide single nucleotide polymorphism scan suggests adaptation to urbanization in an important pollinator, the red-tailed bumblebee (*Bombus lapidarius* L.). *Proceedings of the Royal Society B: Biological Sciences*, 285(1877), 20172806. <https://doi.org/10.1098/rspb.2017.2806>
- Tsachaki, M., & Sprecher, S. G. (2012). Genetic and developmental mechanisms underlying the formation of the *Drosophila* compound eye: Formation of the *Drosophila* compound eye. *Developmental Dynamics*, 241(1), 40–56. <https://doi.org/10.1002/dvdy.22738>
- Vesala, L., Salminen, T. S., Košťál, V., Zahradníčková, H., & Hoikkala, A. (2012). Myo-inositol as a main metabolite in overwintering flies: Seasonal metabolomic profiles and cold stress tolerance in a northern drosophilid fly. *Journal of Experimental Biology*, 215(16), 2891–2897. <https://doi.org/10.1242/jeb.069948>
- Vickruck, J. L., & Richards, M. H. (2017). Nesting habits influence population genetic structure of a bee living in anthropogenic disturbance. *Molecular Ecology*, 26(10), 2674–2686. <https://doi.org/10.1111/mec.14064>
- Wallberg, A., Schöning, C., Webster, M. T., & Hasselmann, M. (2017). Two extended haplotype blocks are associated with adaptation to high altitude habitats in east African honey bees. *PLoS Genetics*, 13(5), e1006792. <https://doi.org/10.1371/journal.pgen.1006792>
- Warnes, G. R., Bolker, B., Bonebakker, L., Gentleman, R., Huber, W., Liaw, A., Lumley, T., Maechler, M., Magnusson, A., Moeller, S., Schwartz, M., & Venables, B. (2020). *gplots: Various R Programming Tools for Plotting Data*. R package version 3.1.1. <https://CRAN.R-project.org/package=gplots>
- Wasielowski, O., Wojciechowicz, T., Giejdasz, K., & Krishnan, N. (2015). Enhanced UV-B radiation during pupal stage reduce body mass and fat content, while increasing deformities, mortality and cell death in female adults of solitary bee *Osmia bicornis*. *Insect Science*, 22, 512–520. <https://doi.org/10.1111/1744-7917.12124>
- Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 38(6), 1358–1370.
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., & Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, 4(43), 1686. <https://doi.org/10.21105/joss.01686>
- Xu, K., Niu, Q., Zhao, H., Du, Y., & Jiang, Y. (2017). Transcriptomic analysis to uncover genes affecting cold resistance in the Chinese honey bee (*Apis cerana cerana*). *PLoS One*, 12(6), e0179922. <https://doi.org/10.1371/journal.pone.0179922>
- Yu, F., Peng, W., Tang, B., Zhang, Y., Wang, Y., Gan, Y., & Ke, C. (2021). A genome-wide association study of heat tolerance in Pacific abalone

based on genome resequencing. *Aquaculture*, 536, 736436. <https://doi.org/10.1016/j.aquaculture.2021.736436>

Zamudio, K. R., & Savage, W. K. (2003). Historical isolation, range expansion, and secondary contact of two highly divergent mitochondrial lineages in spotted salamanders (*Ambystoma maculatum*). *Evolution*, 57(7), 1631–1652.

Zattara, E. E., & Aizen, M. A. (2021). Worldwide occurrence records suggest a global decline in bee species richness. *One Earth*, 4(1), 114–123. <https://doi.org/10.1016/j.oneear.2020.12.005>

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>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Samad-zada, F., Kelemen, E. P., & Rehan, S. M. (2023). The impact of geography and climate on the population structure and local adaptation in a wild bee. *Evolutionary Applications*, 16, 1154–1168. <https://doi.org/10.1111/eva.13558>