Geography, climate and changes in host plants distribution explain patterns of genomic variation within the cactus moth

Daniel Poveda-Martínez[1], Victor Noguerales[2], Laura Varone[3], Stephen Hight[4], Guillermo Logarzo[3], Brent Emerson[5], and Esteban Hasson[6]

[1]Instituto de Ecología, Genética y Evolución de Buenos Aires
[2]Island Ecology and Evolution Research Group, Instituto de Productos Naturales y Agrobiología (IPNA-CSIC)
[3]Fundación para el Estudio de Especies Invasivas
[4]USDA-ARS, Insect Behavior and Biocontrol Research Unit (IBBRU), Tallahassee, Florida, USA.
[5]IPNA-CSIC
[6]CONICET-UNIVERSIDAD DE BUENOS AIRES

September 1, 2022

Abstract

Landscape heterogeneity and the reconfiguration of host plant distributions as a consequence of Quaternary climate oscillations are suggested to play a determinant role in shaping the evolutionary history of herbivorous insects. The cactus moth, Cactoblastis cactorum, is a southern South American phytophagous insect specialized in the use of cacti as feeding and breeding resources. It can be found across broad latitudinal and longitudinal gradients feeding on diverse native Opuntia species as well as the exotic and cultivated species Opuntia ficus-indica. Using high-throughput sequence data for the nuclear genome and mitochondrial DNA sequencing, we investigated patterns of genomic variation of C. cactorum across its native distribution. We integrated a demographic modeling approach for inferring gene flow and divergence times between C. cactorum populations, within a landscape genomic framework, to test alternative spatially-explicit hypotheses of past and current population connectivity based on climatically suitable areas for the focal species and distributions of host plants. Regions currently exhibiting high genomic diversity were evaluated for congruence with areas where suitable climatic conditions remained stable from the last glacial maximum to the present. Results revealed significant population structure across the range of C. cactorum, that can be explained by the spatial configuration of persistently suitable environmental conditions and host plant ranges during interglacial and glacial periods. Moreover, genomic data supported a hypothesis of long-term habitat stability in the northern regions of the distribution that served as a refuge for C. cactorum, enabling the accumulation and maintenance of high levels of genetic diversity over time.

Geography, climate and changes in host plants distribution explain patterns of genomic variation within the cactus moth

Running head: Genomic variation of the cactus moth

Daniel Poveda-Martínez[1,2*], Víctor Noguerales[3], Laura Varone[1], Stephen D. Hight[4], Guillermo Logarzo[1], Brent C. Emerson[3], Esteban Hasson[2].

[1]Fundación para el Estudio de Especies Invasivas (FuEDEI), Hurlingham, Argentina.

[2]Instituto de Ecología Genética y Evolución de Buenos Aires (IEGEBA-CONICET), Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires, Buenos Aires, Argentina.

[3]Island Ecology and Evolution Research Group, Instituto de Productos Naturales y Agrobiología (IPNA-CSIC), San Cristóbal de La Laguna, Canary Islands, Spain.

[4]USDA-ARS, Insect Behavior and Biocontrol Research Unit (IBBRU), Tallahassee, Florida, USA.

*Corresponding author:*danielpovedam@gmail.com*(D.P.M).

Abstract

Landscape heterogeneity and the reconfiguration of host plant distributions as a consequence of Quaternary climate oscillations are suggested to play a determinant role in shaping the evolutionary history of herbivorous insects. The cactus moth, *Cactoblastis cactorum* , is a southern South American phytophagous insect specialized in the use of cacti as feeding and breeding resources. It can be found across broad latitudinal and longitudinal gradients feeding on diverse native *Opuntia* species as well as the exotic and cultivated species *Opuntia ficus-indica* . Using high-throughput sequence data for the nuclear genome and mitochondrial DNA sequencing, we investigated patterns of genomic variation of *C. cactorum* across its native distribution. We integrated a demographic modeling approach for inferring gene flow and divergence times between *C. cactorum* populations, within a landscape genomic framework, to test alternative spatially-explicit hypotheses of past and current population connectivity based on climatically suitable areas for the focal species and distributions of host plants. Regions currently exhibiting high genomic diversity were evaluated for congruence with areas where suitable climatic conditions remained stable from the last glacial maximum to the present. Results revealed significant population structure across the range of *C. cactorum* , that can be explained by the spatial configuration of persistently suitable environmental conditions and host plant ranges during interglacial and glacial periods. Moreover, genomic data supported a hypothesis of long-term habitat stability in the northern regions of the distribution that served as a refuge for *C. cactorum* , enabling the accumulation and maintenance of high levels of genetic diversity over time.

Key words : Cactus pest, *Cactoblastis cactorum* , ddRAD, landscape genomics, paleoclimate, *Opuntia* .

Introduction

The study of herbivorous insect pests and the spatio-ecological factors affecting their historical and contemporary population genetic structure is of interest for both evolutionary understanding and pest management. When such studies involve species distributed across environmental gradients, with multiple host species, they can also provide a window on the processes driving diversification and shaping gene flow dynamics in heterogeneous landscapes (Borer, Arrigo, Buerki, Naisbit & Alvarez, 2012; Laukkanen, Mutikainen, Muola & Leimu, 2014).

The use of alternative host plants is thought to be a strong determinant in the evolutionary history of phytophagous insects (Forbes et al., 2017; Funk, Nosil, & Etges, 2006), a factor which together with adaptation to differing environmental conditions can lead to ecological specialization and diversification (isolation by environment, IBE; Wang & Bradburd, 2014). Geography (isolation by distance, IBD; Slatkin, 1993) and the spatial distribution of suitable habitats (isolation by resistance, IBR; McRae, 2006; McRae & Beier, 2007) may also be determinants of population genetic differentiation across the landscape (Driscoe et al., 2019; Peterman, Connette, Semlitsch & Eggert, 2014; Vidal, Quinn, Stireman III, Tinghitella & Murphy, 2019). Additionally, changes in host plant distributions driven by climate oscillations throughout the Quaternary period may have an important role in shaping contemporary patterns of genetic variation in species with narrow feeding requirements (Noguerales, Cordero & Ortego, 2018).

Southern South America has a complex and ancient geological history in which orogenic processes, such as the uplift of the Andes, together with marine transgressions and biotic landscape changes caused by Quaternary climatic oscillations, have been hypothesized to shape intra and interspecific diversification in the region (Agrain, Domínguez, Carrara, Griotti & Roig-Junent, 2021; Hewitt, 2000; Ortiz-Jaureguizar & Cladera, 2006; Rodriguero, Lanteri, Guzman, Carus Guedes & Confalonieri, 2016). These historical events have synergistically contributed to promote periods of both population isolation and subsequent secondary contact (Lavinia, Barreira, Campagna, Tubaro & Lijtmaer, 2019; Rocha et al., 2020). Evidence

for evolutionary consequences of landscape changes promoted by Quaternary climatic dynamics is being increasingly documented in southern South America. The distributional shifts of open vegetation biomes, such as the Chacoan and Pampean domains, have been hypothesized to determine the diversification history of many faunal taxa intimately linked to these habitats (Rodriguero et al., 2016; Rosetti, Krohling & Remis, 2022; Turchetto-Zolet, Pinheiro, Salgueiro & Palma-Silva, 2013).

Among the southern South America invertebrate fauna, the cactus moth, *Cactoblastis cactorum* (Berg) (Lepidoptera: Pyralidae), provides a suitable opportunity to investigate the aforementioned factors and their potential role in diversification within the region. *Cactoblastis cactorum* is a phytophagous moth native to southern South America. It can be found across wide latitudinal and longitudinal gradients, from the arid and semi-arid lands of northwestern and northeastern Chaco, the grasslands of the Pampa, through to the northern area of the Patagonian steppe (McFadyen, 1985; Morrone, 2014; Varone, Logarzo, Briano, Hight & Carpenter, 2014) (Figure 1). Across its broad distribution, *C. cactorum* exploits a broad spectrum of host species within the genus *Opuntia* (Cactaceae, Opuntioideae), including several native species such as *O. quimilo* K. Schum., *O. megapotamica* Arechav., *O. bonaerensis* Speg., *O. elata* Link & Otto ex Salm-Dyck, *O. rioplatense* Font, *O. anacantha* Speg. and *O. penicilligera* Speg. and the exotic and commercially cultivated *O. ficus-indica* (L.) Miller (Varone et al., 2012, 2014).

The affinity of *C. cactorum* for species of *Opuntia* (prickly pear cactus) has encouraged its use as a biological control agent since 1925, particularly during the successful programs against alien *Opuntia* species in Australia and South Africa (Julien & Griffiths, 1998). Despite the substantial success in controlling invasive cacti in Australia, the deliberate introduction of *C. cactorum* as a biological control agent into the Caribbean (Simmonds & Bennett 1966) resulted in threats to both cactus diversity as well as the prickly pear fruit industry in the Caribbean, and more recently in North America (Hight & Carpenter, 2009). In its native range, *C. cactorum* also represents a serious threat to the regional prickly pear fruit production industry based on the Mexican species *O. ficus-indica* . The use of this species as a feeding and breeding resource represents a relatively recent expansion of host use by *C. cactorum* (Ervin, 2012). Almost one third of *O. ficus-indica* crops are infested with *C. cactorum* in some regions of Argentina (Varone et al., 2014). *Opuntia ficus-indica* is considered a traditional crop in arid and semi-arid regions of Argentina, being part of family subsistence economies and also having a prominent role in the regional economy (Ochoa, Targa, Abdala & Leguizamon, 2007). While the prickly pear industry is mainly concentrated in central and northern Argentina, patches of *O. ficus-indica* can also be found in other parts of the country (Varone et al., 2014). This distribution has led to the suggestion that *O. ficus-indica* may promote the movement of *C. cactorum* among distant populations, through human activity

Previous studies have reported that populations of *C. cactorum* are not only genetically differentiated (Marsico et al., 2011), but that patterns of coloration of mature larvae and host plant use also vary across regions (Brooks, Ervin, Varone & Logarzo, 2012; McFadyen, 1985). However, only limited conclusions can be drawn from these studies, due to the following: (i) only a single mitochondrial genetic marker was used; (ii) sampling did not cover the entire species range; and (iii) the relative abundance of host species were not taken into account (Brooks et al., 2012; Marsico et al., 2011; McFadyen, 1985). More recently, an extensive survey evaluating host plant preferences across the broad range of *C. cactorum* concluded that patterns of host use reflect variation in host availability rather than female preference (Varone et al., 2014). In accordance with this finding, a recent study of *C. cactorum* and two closely related specialist species, *Cactoblastis bucyrus* Dyar and *Cactoblastis doddi* Heinrich, demonstrated that geography is more important than host species as a driver of genetic differentiation in *C. cactorum* , contrasting with the two specialist species for which intra-specific genetic differentiation is mainly determined by host species rather than geography (Poveda-Martinez et al., 2022). These findings are consistent with the idea that the use of alternative host species in generalist herbivores may not promote divergence, due to relatively low fitness variation across different host species (Vidal & Murphy, 2018; Vidal et al., 2019). However, additional work is needed for a complete understanding of the effects of spatiotemporal changes in landscape factors on the evolutionary history of *C. cactorum* in its native range (Varone et al., 2014; Poveda-Martinez et al., 2022).

In this study, we use high-throughput nuclear sequence data and Sanger sequence data for mitochondrial DNA to investigate the evolutionary history of the cactus moth, *C. cactorum* . We sample populations of*C. cactorum* representatively across its broad geographic range and diversity of native and non-native host species. We combine inferences from phylogenomics and demographic modeling to uncover the tempo and mode of diversification among genetic groups, together with patterns of gene flow, population genetic structure and diversity. By integrating ecological niche modeling into a landscape genomic framework, we are able to test a comprehensive suite of competing scenarios of population isolation, specifically evaluating the hypotheses that (i) contemporary spatial patterns of population differentiation in *C. cactorum* can be explained by the geographic reconfiguring of suitable areas for the focal species and its host species, caused by Quaternary climate dynamics, and (ii) regions where*C. cactorum* exhibits high genomic diversity, together with no evidence for genetic admixture, are those where suitable climatic conditions have remained stable from the last glacial maximum (LGM, ca. 21 ka) to the present-day. Finally, we also evaluated if the broadened distribution of *O. ficus-indica* resulting from its introduction and cultivation is facilitating contemporary gene flow among geographically distant populations, thus promoting genetic admixture within *C. cactorum.*

Methods

Population sampling and DNA extraction

Between 2018 and 2019, we collected 140 individuals of *C. cactorum* from 28 sampling sites spanning a broad range of its native distribution in southern South America (Figure 1; Table S1). Individuals were collected directly from host plants, including six native taxa,*O. megapotamica, O. rioplatense, O. elata, O. penicilligera, O. quimilo,* and *O. anacantha* , and the exotic *O. ficus-indica*(Varone et al., 2014). Given that adult moths have nocturnal activity and a short lifespan, only eggsticks and I-IV instar larvae were collected. Collected eggsticks were maintained in the lab under constant temperature (25degC) and humidity (70%). After eggs hatched, larvae were reared on *O. ficus-indica* until the IV instar stage. To avoid the inclusion of siblings, only a single individual per sampled plant and eggstick was used for genetic analysis. Larvae were individually transferred to vials containing 100% ethanol and stored at –20degC until DNA extraction. Before DNA extraction, individual larvae were dissected to remove any food residue, and ˜30 mg of clean tissue was used for DNA extraction using the Qiagen DNeasy Blood & Tissue Kit following manufacturer's instructions (Qiagen, Inc.).

Genomic libraries preparation, filtering and assembling

Individuals were processed into two genomic libraries following the double digest restriction-site Associated DNA (ddRAD) procedure described in Peterson, Weber, Kay, Fisher and Hoekstra (2012). Genomic DNA was fragmented using *NspI* and *MboI* restriction enzymes, and then purified and ligated to barcoded adapters. Samples were pooled within multiplexing batches and bead purified. Libraries were sequenced in paired-end 125 bp mode on a HiSeq2500 Illumina instrument.

Demultiplexed raw reads were checked for quality using fastqc v0.11.15 (Andrews et al., 2010), and outputs were compiled and summarized on MultiQC v1.10.1 (Ewels, Magnusson, Lundin & Kaller, 2016). Reads were assembled with IPYRAD v.0.9.59 (Eaton & Overcast, 2020) using the strictest filter for Illumina adapters. A reference based assembly method was implemented for mapping the filtered reads against the*C. cactorum* draft genome (NCBI accession number: JADGIL000000000). As this genome was assembled from an adult male (clean DNA can be obtained from males, as they do not feed), it serves as a filter for exogenous reads derived from larval food (Poveda-Martinez et al., 2022). In subsequent IPYRAD assembly steps, we allowed 20% of SNPs per RAD locus, and shared heterozygous sites occurring across a maximum of 25% of samples. A minimum of 90% was set for the minimum number of samples scored per locus. We used vcftools v1.15 (Danecek et al., 2011) to remove SNPs with minimum allele frequency lower than 3%, missing data per site across individuals exceeding 25%, and to keep SNPs with read depths between 6X to 100X. Individuals with more than 20% missing data were excluded from further analyses. In order to prune SNPs in linkage disequilibrium, PLINK v1.9 (Purcell et al., 2007) was used with a window size of 50 bp, window shift of 5

and VIF threshold of 2. SNPs under selection were identified using Bayescan v.2.1 (Foll & Gaggiotti, 2008) and excluded from the dataset. A fragment of the mitochondrial gene encoding for the Cytochrome oxidase subunit I (*COI*) (mtDNA) was amplified using primers C1-J-2183 (Simon et al., 1994) and PatII (Caterino & Sperling, 1999). Amplifications, PCR-product purification and sequencing were performed as detailed in Poveda-Martinez et al., (2022).

Assessing genomic population structure

Population genetic structure was inferred using genome-wide SNP data (nDNA hereafter) and the sparse non-negative matrix factorization approach (sNMF) (Frichot, Mathieu, Trouillon, Bouchard & Francois, 2014), as implemented in R using the LEA package (Frichot & Francois, 2015). The number of genetic clusters that best described our data was assessed with 100 repetitions for each possible K value (from K=1 to K=10) and using the cross-entropy criterion. In addition to sNMF, genetic structure was approximated by using the major axes of genomic variation obtained from a Principal Component Analysis (PCA), as implemented in the hierfstat R package (Goudet, 2005). We used mtDNA data to construct a haplotype network using the Median Joining algorithm (Bandelt, Forster, & Rohl, 1999) as implemented in PopArt v.1.7.1 (Leigh & Bryant, 2015).

Phylogenetic analyses

Phylogenetic relationships were reconstructed among the main genetic groups as inferred in sNMF using nDNA data and the coalescent-based method implemented in SNAPP (Bryant, Bouckaert, Felsenstein, Rosenberg, & RoyChoudhury, 2012). Due to computational burden, SNAPP analyses were run including only five individuals per genetic cluster, selecting those with an ancestry coefficient higher than 0.85, according to sNMF. One SNP per locus was randomly selected resulting in a new matrix of 2,086 SNPs shared across tips. The default model parameters were used in SNAPP for U and V equal to one, and the analysis was run for 5,000,000 MCMC generations, sampling every 1000 generations in Beast v.2.5.2 (Bouckaert et al., 2014). The complete set of trees was visualized in Densitree v.2.2.5 (Bouckaert & Heled, 2014), removing the first 10% of the trees as burn-in. A maximum credibility tree was generated using TreeAnnotator v.1.7.5 (Drummond, Suchard, Xie, & Rambaut, 2012).

Demographic history

A coalescent-based simulation approach was implemented in Fastsimcoal2 based on the site frequency spectrum (SFS) of nDNA data (Excoffier, Dupanloup, Huerta-Sanchez, Sousa & Foll, 2013) to further investigate the demographic history of the three main lineages identified in clustering analysis and species tree inference (Center, East and South, see results section). Initially, to further evaluate the consistency of phylogenetic relationships among *C. cactorum* populations inferred in SNAPP, models representing the three possible topologies (A, B and C) were considered in each one of the subsequent models of gene flow (Figure S1). For each topology, divergence times and contemporary effective population sizes were estimated, as well as alternative scenarios of gene flow. Models 1-3 considered no gene flow among lineages; models 4-6 contemplated scenarios of symmetric gene flow among lineages, while models 7-9 considered scenarios of asymmetric gene flow among lineages. For these analyses, 20 individuals were selected with the highest ancestry coefficient (> 0.85) according to sNMF (e.g., Ortego, Cespedes, Millan & Green, 2021). The folded joint SFS was calculated considering a single SNP per locus using the Python script written by Isaac Overcast and available at GitHub (*https://github.com/isaacovercast/easySFS*). To remove missing data and minimize errors with allele frequency estimates, each genetic cluster was downsampled to 16 individuals yielding a total of 2,565 variable sites. Assuming that invariable sites were not considered in the SFS calculation, we used the "removeZeroSFS" option in Fastsimcoal2 and fixed the effective population size of one of the demes (South lineage, NE_{SOUTH}) to enable the estimation of other parameters with Fastsimcoal2 (Excoffier et al., 2013). NE_{SOUTH} was calculated according to the equation $NE= \pi/4\mu$ (Lynch & Conery, 2003). Nucleotide diversity (π=0.004) was estimated using variant and invariant sites with DNAsp v6 (Rozas et al., 2017). The mutation rate (μ) was considered to be 2.9×10^{-9} substitutions per site per generation, previously estimated for the butterfly *Heliconius melpomene* (Lepidoptera: Nymphalidae) (Keightley et al., 2015). Each model

was run with 50 replicates, considering 100,000–250,000 simulations for the calculation of the composite likelihood, 10–40 expectation-conditional maximization (ECM) cycles, and a stopping criterion of 0.001. Once maximum likelihood was estimated per run, the best demographic model was selected according to Akaike's information criterion (AIC). AIC values were rescaled in terms of AIC differences (Δi) according to the formula: Δi = AICi - AICmin. A model with a ΔAIC value of 0 and the highest AIC weight (ωi) served as the best model. A parametric bootstrapping approach was used to construct 95% confidence intervals of the estimated parameters running 100 bootstrap replicates using initialized values from the best model (Excoffier et al., 2013).

Environmental niche modeling

An environmental niche modeling approach was implemented to predict contemporary and historical geographic distributions of suitable areas for the cactus moth in its native range. Environmental niche models (ENMs) were used to determine the impact of suitable habitat distribution on *C. cactorum* diversification as is detailed below in the landscape genomic analysis section. ENMs were built using MaxEnt v.3.4.1 (Phillips, Anderson, & Schapire, 2006) and the bioclimatic data available in CHELSA v.1.2 database (*https://chelsa-climate.org/bioclim/*). Model parameters in MaxEnt were selected and optimized using the kuenm R package (Cobos, Peterson, Barve & Osorio-Olvera, 2019). For both temporal bioclimatic conditions, present-day and Last Glacial Maximum (LGM, ca. 21 kya), we evaluated 15 environmental variables retrieved at 30 arc-sec (˜1 km) of resolution. Four bioclimatic variables (Bio 8, 9, 18 and 19) were discarded for having artificial breaks (Oliveira et al., 2020). Highly correlated variables (R >0.9) according to variance inflation factor criterion were excluded for downstream analyses resulting in a final dataset of 6 bioclimatic variables (Bio 2, 3, 5, 13, 14 and 15). Suitability maps during the LGM were obtained by projecting the present-day ENM onto LGM bioclimatic conditions derived from the Community Climate System Model (CCSM4; Braconnot et al. 2007) resulting in two suitability maps based on bioclimatic variables data (Set 1, $\text{Current}_{\text{ENV}}$ and LGM_{ENV} from hereafter).

Additionally, we constructed ENMs for five of the six native host species of *C. cactorum* considering the aforementioned approach for variable selection and model building. ENM for *O. penicilligera* could not be implemented due to limited occurrence data. The resulting ENMs obtained for current and LGM conditions from each host species were used as variables to build an additional ENM for *C. cactorum* (Set 2, $\text{Current}_{\text{HOST}}$ and LGM_{HOST} from hereafter). A third ENM was constructed for *C. cactorum* considering the input of both the bioclimatic variables (Set 1) and the host plant ENMs (Set 2) (Set 3, $\text{Current}_{\text{ENV-HOST}}$ and $\text{LGM}_{\text{ENV-HOST}}$ from hereafter). Occurrence data for the host species and *C. cactorum*were obtained from field surveys made from 2007 to 2019 and complemented, after detailed curation, with distribution information available at GBIF (*www.gbif.org*). Redundant occurrences (e.g. points occurring within 1,500 km^2) were excluded using spThin R package (Aiello-Lammens, Boria, Radosavljevic, Vilela & Anderson, 2015). After thinning occurrence data, 81 records for *C. cactorum* locations and their associated occurance of 205 host species remained [*O. quimilo* (50), *O. megapotamica* (47), *O. rioplatense* (39),*O. elata* (38), and *O. anacantha* (31)] and were used to conduct ENMs. Model performance for each scenario was evaluated independently based on statistical significance (Partial ROC), omission rates (OR), and the Akaike information criterion corrected for small sample sizes (AICc) using the kuenm R package (Cobos et al. 2019).

Landscape genomic analyses

A landscape genomic approach was implemented to study potential factors that could explain patterns of genetic differentiation within *C. cactorum* . As a measure of genetic differentiation, pairwise F_{ST} estimates were derived from genome-wide SNP data using the Weir and Cockerham (1984) method with the StAMPP R package (Pembleton, Cogan & Forster, 2013), using 9,999 bootstrap replicates. We evaluated several plausible scenarios of population connectivity based on historical and contemporary spatial and ecological data.

Isolation by resistance scenarios (IBR) were tested by calculating resistance surfaces based on the suitability maps obtained from ENMs for present-day and LGM conditions considering different subsets of factors:

(i) only climatic variables ($\text{Current}_{\text{ENV}}$ and LGM_{ENV}); (ii) climate-based habitat suitability maps for host species ($\text{Current}_{\text{HOST}}$ and LGM_{HOST}); and (iii) combining climatic variables and climate-based habitat suitability maps for host species ($\text{Current}_{\text{HOST-ENV}}$ and $\text{LGM}_{\text{HOST-ENV}}$). Resistance distances for all pairs of populations were calculated using an eight-neighbor cell connection scheme in Circuitscape v.5 (Hall et al., 2021) through Julia v.1.5.2 (*https://julialang.org/*). We also calculated resistance distances based on a "flat landscape" where all cells have an equal resistance value (=1) representing a null model of isolation by resistance (IBR_{NULL}). Resulting between-population resistance matrices were used as input for downstream statistical analyses.

Climatic dissimilarities were estimated between populations to evaluate an isolation by environment scenario (IBE_{CLI}). Climatic data was extracted from the 15 CHELSA bioclimatic variables for each of the 28 sampling sites, as well as from 500 random points covering our study area to avoid potential biases resulting from only considering conditions at focal sites. Due to collinearity among the bioclimatic variables, we ran a PCA using the ade4 R package and summarized the environmental variation in the three first axes accounting for ca. 82% (PC1=43.79%; PC2=23.96%; PC3=14.31%) of total variation. The environmental dissimilarity matrix was obtained by calculating the Euclidean distances for PC scores between pairs of sampling sites (Ortego et al., 2021).

An isolation by distance scenario was also evaluated using weighted topographic distances (IBD_{WTD}), which incorporate an additional overland distance covered by an organism due to changes in elevation imposed by the topography. The IBD_{WTD} distance matrix was calculated on a digital elevation model (DEM) at ˜1 km of resolution retrieved from WorldClim v.2.1 dataset (Fick & Hijmans, 2017) using the TopoWeightedDist function implemented in the topoDistance R package (Wang, 2020). We assumed a linear function to weight aspect changes (hFunction parameter) and an exponential function to weight the slope (vFunction parameter) (e.g. Noguerales et al., 2021).

Relationships between explanatory distance matrices and genetic differentiation between populations (F_{ST}) were evaluated using univariate and multiple matrix regressions with randomization using the function MMRR (Wang, 2013) as implemented in R. An initial full model was constructed considering all significant explanatory terms identified previously in univariate analysis, and a final best-fit model was selected using a backward-stepwise procedure by progressively removing non significant variables until all retained terms within the model were significant (Ortego, Gugger & Sork, 2014). The result was the minimal most adequate model for explaining the variability in the response variable, where only the significant explanatory terms were retained.

Population genetic diversity and climate/habitat stability

Given that suitable environments tend to support larger populations (Carnaval, Hickerson, Haddad, Rodrigues & Moritz, 2009; Soley-Guardia, Carnaval & Anderson, 2019), we tested whether the regions exhibiting high genetic diversity were those where suitable climatic conditions remained stable through time, from the LGM (˜21 Kya) to the present. To test this hypothesis, environmental stability maps were constructed by averaging current and LGM suitability maps. For each sampling site, values of climate/habitat stability were extracted for each of the three scenarios considering (i) only climatic variables ($\text{Stability}_{\text{ENV}}$), (ii) climate-based habitat suitability maps for host species ($\text{Stability}_{\text{HOST}}$), and (iii) combining climatic variables and climate-based habitat suitability maps for host species ($\text{Stability}_{\text{ENV-HOST}}$). Raster calculations were conducted using the R raster package (Hijmans & van Etten, 2016).

Nuclear and mitochondrial diversity were characterized for each of the 28 sampling sites. Genetic diversity estimates were normalized to three individuals per sampling location to avoid potential bias resulting from uneven population sample size (range: 3-11 individuals). For nDNA, expected heterozygosities (H_{E}) and nucleotide diversity (π) were calculated using the diveRsity R package (Keenan, McGinnity, Cross, Crozier & Prodöhl, 2013), and DNAsp v6 (Rozas et al., 2017), respectively. We also estimated π and haplotype diversity (H_{d}) for mtDNA data using DNAsp v6 (Rozas et al., 2017). Correlations between stability values and estimates of nuclear (H_{E}, π) and mtDNA (H_{D}, π) population genetic diversity were tested using

linear regressions. Longitude and latitude were also included as explanatory factors to account for potential geographical clines of genetic diversity (Guo, 2012).

Results

Genomic data

We successfully genotyped 138 individuals of *C. cactorum* using ddRAD sequencing which were representative of populations feeding on both exotic *O. ficus-indica* (71 individuals), and native host plants *O. megapotamica* (34), *O. rioplatense* (11), *O. elata* (10), *O. bonaerensis* (4), *O. penicilligera* (3),*O. quimilo* (3), and *O. anacantha* (2). After filtering steps, the average number of paired-end reads retained per individual was 1,859,116, of which an average of 1,253,277 reads mapped to the*C. cactorum* reference genome. After discarding loci in linkage disequilibrium and under selection, we recovered a total of 3,506 unlinked sequence loci with an average coverage of 30X which contained 17,084 biallelic SNPs (Table S2). A total of 143 individuals from the same 28 locations were successfully sequenced for a fragment of 790 bp of the *COI* gene. The analysis of mtDNA data revealed 66 haplotypes, 171 variable sites, 97 of which were parsimony-informative sample-wide.

Assessing population genomic structure

Clustering analysis with sNMF indicated K=6 to be the most likely number of genetic clusters: North, Central, Western, Southeast, Southwest and East. However, these analyses revealed a certain degree of admixture among clusters; some Northern and Central individuals exhibited a relatively high degree of admixture with both the Western and Eastern clusters (Figure 2A-C). Interestingly, Northern, Central and Western populations appeared to form a unique cluster (Central lineage from hereafter) when assuming K=3, whereas both Southern (South lineage from hereafter) and Eastern (East lineage from hereafter) populations remained as separated panmictic groups. With K=4, the Southern population appeared subdivided into two geographic units; Southeastern and Southwestern. Finally, with K=5, Northern populations appeared separated from the Central population (Figure S2). Principal component analysis (PCA) reveals the Eastern lineage to appear separated from the Central lineage along PC1, whereas the South lineage can be distinguished from the Central and Eastern lineages along PC2 (Figure 2D). Additional PCA structuring was also evident, mainly in the Southern lineage in agreement with the clustering pattern observed with K=4 in sNMF. Moreover, the PCA revealed an apparent admixture signal between the Central and Eastern lineages.

The median-joining network obtained with mtDNA sequence data revealed strong spatial genetic structure consistent with the clusters identified with nDNA data (Figure 2E). The Central haplogroup was one of the most frequent haplotypes and was separated by seven mutational steps from the Eastern haplogroup, whereas four mutational steps separated the Central and Northern haplogroups. Southeastern and Southwestern haplogroups appeared close to the Central haplogroup and shared a common haplotype. Despite spatial genetic structure, a few Central haplotypes were distributed in other regions (three central haplotypes in the East, two in the Southeast and one in the North). A frequent haplotype in the Northern haplogroup was also detected in the central region and a haplotype from the East was observed in the north region.

Species tree reconstruction

Phylogenetic relationships inferred in SNAPP were consistent with the hierarchical spatial genetic structure observed with sNMF (Figure 2B-C; Figure S2). The most ancestral split corresponded to the separation of South and East lineages from the Central lineage. Southwestern and Southeastern clusters, as well as, West, Central and North clusters diverged subsequently from their respective lineages, as observed in clustering analyses. The maximum clade credibility tree indicated a high posterior probability (> 0.9) supporting all nodes in the species tree (Figure 2B).

Demographic inference using coalescent-based simulations

The scenario considering full asymmetric interpopulation gene flow (model 8) was the most supported among the nine demographic models tested in Fastsimcoal2 (Table S3). Under this model, divergence among the

three main lineages of *C. cactorum* was estimated to have occurred during the Late Pleistocene (Figure 3; Table 1). Specifically, this model describes an initial split (T_{DIV1}) of the Central lineage ˜75 kya (considering two generations per year in native host species, as reported by Varone et al., 2014) and a more recent split (T_{DIV2}) separating the South and East lineages occurring ˜19 kya. Demographic simulations estimated an effective population size close to 250,000 for the Central lineage (NE_{CENTER}), which was six times higher than the value obtained for the East lineage (NE_{EAST}=41,551). Estimates of gene flow rates varied among lineages with the highest rate of contemporary gene flow estimated to occur from the Central to the East lineage (M_{CE}=1.50 x 10^{-5}) and the lowest from the Central lineage to the South lineages (M_{CS}=2.71 x 10^{-8}) (Table 1).

Ecological niche modeling through time

Ecological niche models (ENMs) for both the focal species (*C. cactorum*) and its host species presented relatively high AUC scores (Table S4 and S5), indicating an overall good model performance. Suitability maps of host species were concordant with their respective current distributions (Figure S3). The three models for *C. cactorum* (Set 1: $Current_{ENV}$, Set 2: $Current_{HOST}$, Set 3: $Current_{ENV\text{-}HOST}$) provided similar predicted distributions, with no major differences (Figure S4). However, the most supported model for *C. cactorum*according to AIC values was that considering as input variables the predicted distribution of the host plants (Set 2), followed by the model including these and the six bioclimatic variables (Set 3). Additional information on the results of ENMs for both *C. cactorum* and host species are detailed in Tables S4 and S5. The present-day distribution inferred by the most supported scenario ($Current_{HOST}$) was in line with the contemporary range of *C. cactorum* . Under this model, higher suitability areas were predicted in both Chaco and north of Pampa biogeographic regions whereas areas of lower suitability were predicted in the southern portions of the Pampas and the Monte biogeographic regions (Figure 1B). Other biogeographic provinces (Yungas, Puna and Prepuna) had extremely low values of suitability for*C. cactorum* . Projections of ENM to LGM predicted high environmental suitability along the Chaco biogeographic province to the most southern portion of Pampas, with areas of low suitability located in eastern Argentina, a confluence zone of La Plata river basin (Figure 1C). Comparing within each scenario across time, from the LGM to the present, areas of high environmental stability were predicted in North and Central Argentina (the arid and semi-arid regions of Chaco), with low environmental stability predicted in southern Argentina.

Landscape genomic analyses

Estimates of genetic differentiation (F_{ST}) among sampling sites ranged from 0.023 to 0.448 for nDNA data (Table S6). Univariate matrix regressions indicated that nuclear genetic differentiation was significantly correlated with all distance matrices except with $Current_{ENV}$ and $Current_{ENV\text{-}HOST}$ (Table S7). Yet, only $LGM_{ENV\text{-}HOST}$ was significantly retained in the best-fit model after the backward-stepwise procedure (Table 2).

Population genetic diversity and climate/habitat stability

We found higher nuclear and mitochondrial diversity in the Northern, Central and Eastern population, than in Southern populations (Figure 4; Table S8). Linear regression analyses revealed that nuclear genetic diversity was significantly correlated with latitude for nDNA but not with longitude (Table 3). Conversely, we found that mtDNA genetic diversity was not correlated with latitude nor with longitude. Likewise, analyses showed that all nuclear genetic diversity estimates (H_E, π) were positively correlated with the three environmental stability estimates ($Stability_{ENV}$, $Stability_{HOST}$ and $Stability_{ENV\text{-}HOST}$) from the LGM (˜21 Kya) to the present, but mitochondrial genetic diversity estimators (H_D, π) were not correlated (Table 3; Figure 4).

Discussion

Nuclear SNPs and mtDNA data revealed significant population structure across the range of *C. cactorum* . Landscape genomic analyses provided support for the hypothesis that a combination of environmental conditions and habitat suitability, influenced by shifts in host species distributions during the Last Glacial

Maximum (LGM), were the main forces shaping population genetic structure of *C. cactorum* across its native Argentine range. Shifts in the distribution of *Opuntia*host species, mediated by climatic changes during the Quaternary, are suggested to have had a direct influence on the distribution of *C. cactorum* , creating fragmented ranges that led to reduced gene flow among populations, promoting genetic differentiation. Results also supported the habitat stability hypothesis, whereby regions within which suitable environments have remained stable since the LGM to the present harbored more genetic diversity than regions of lower habitat stability. Relatively higher habitat stability within the northern region of the distribution of *C. cactorum* created refuge conditions promoting the maintenance of higher levels of genetic diversity over time, compared to other areas. The hypothesis that the geographically widespread cultivation of *O. ficus-indica* facilitated contemporary gene flow among otherwise geographically distant populations was rejected, based on population genomic analyses revealing limited genetic admixture and a hierarchical pattern of population differentiation concordant with geography.

Chacoan (North-West-Central lineages) and Pampean (East and South lineages) divergence

Genetic clustering analyses revealed three major lineages defined by geography: Central, East and South (Figure 2 and 3). Estimates from coalescent-based demographic modeling suggested that lineages began to diverge during the Late Quaternary, approximately 75 kya. The earliest divergence involved populations from the Chacoan biogeographic domain (hereafter "Chaco"), including populations from North, West and Central Argentina, and populations from the Pampean biogeographic region (hereafter "Pampa") including South and East lineages (Figure 1 and 2). Prior to this split, habitats in the Chaco biome are believed to have become more like those of the Pampa biome, due to the regional contraction of subtropical and tropical biomes (Ortiz-Jaureguizar & Cladera, 2006). Subsequently, major geological changes and climatic shifts took place in the area leading to the geographical isolation of the Chacoan xeric woodlands from the typical grasslands of the Pampean region. Such changes included a topographic configuration promoted by the Andean and sub-Andean Piedmont and the uplift of the eastern orographic systems (the Pampean sierras) associated with the Peripampasic orogenic arc (Calatayud-Mascarell, Ferretti, Enguídanos, Arnedo, 2022; Crisci, Freire, Sancho & Katinas, 2001; Ortiz-Jaureguizar & Cladera, 2006; Speranza, Seijo, Grela & Solis, 2007). These tectonic events produced a rain-shadow effect that resulted in the extremely xeric conditions presently existing in this area. Additionally, climate change during the Quaternary, which included cold, dry glacial cycles alternating with warm, moist interglacial periods, also affected these regions causing expansion and contraction of many habitats (Ortiz-Jaureguizar & Cladera, 2006). This dynamic likely affected patterns of host species distributions and habitat suitability for*C. cactorum* , which is thought to have promoted the divergence between the Central lineage (Chaco) and East, South lineages (Pampean) (Figure 2 and 3).

East and South (Pampean) lineages divergence

East and South lineages were estimated to have diverged approximately 19 kya. Regional climate models suggest that this timing coincided with an increase in precipitation along the eastern foothills of the Andes (Cook & Vizy, 2006). The topographical features together with the Quaternary climatic evolution of the region maintained a fragmented distribution of host species from *C. cactorum* in the Pampa (Mourelle & Ezcurra, 1997). This would likely have limited dispersal of *C. cactorum* , restricting the East lineage to the more humid environments with deep fertile soils typical of the Pampa, and the South lineage to the drier grasslands of the northern border of the Patagonian steppe.

Climatic changes during the Quaternary promoted complex phylogeographic patterns in many southern South American arthropods, such as the grasshopper *Dichroplus vittatus* Bruner (Rosetti et al., 2022), the beetle *Naupactus cervinus* Boheman (Rodriguero et al., 2016), mygalomorph spiders (Calatayud-Mascarell et al., 2022), among other taxa (Bonatelli, Gehara, Carstens, Colli & Moraes, 2022). Similar phylogeographic consequences of climate change over the last 21,000 years have been recently documented for *O. bonaerensis* , one of the host species for *C. cactorum* (Köhler, Esser, Font, Souza-Chies & Majure, 2020). The concordance in the timing of intraspecific diversification within *O. bonaerensis* and *C. cactorum* is strongly suggestive of a fundamental role for Quaternary climate on the evolutionary history of the cactus-moth system. Intraspecific genetic differentiation in *C. cactorum* and its close relatives *C. doddi* and *C. bucyrus* also dates back to

the late Quaternary (Poveda-Martínez et al., 2022). Diversification within *C. bucyrus* , a specialist on the columnar cactus *Trichocereus atacamensis* , is estimated to have initiated during the Marine Isotopic Stage 3 (˜42 kya), an interstadial during the last glacial period. Divergence within *C. doddi,* which feeds upon *Opuntia sulphurea,* is estimated to have occurred very close to the end of the Pleistocene (˜19 kya), coincident with estimates for the split between East and South lineages of *C. cactorum* (Poveda-Martínez et al., 2022).

Maps of environmental suitability estimated with ecological modeling suggested that the Chaco and northern Pampean regions harbored larger climatically suitable areas than other biogeographic regions such as the Yungas, Monte, Puna and Prepuna (Figure 1). Both *C. doddi* and *C. bucyrus* are found in the Prepuna and Monte regions where *C. cactorum* is absent, supporting the idea that disparate environmental/habitat conditions and disjunct host plant distributions played a role in divergence within *C. cactorum* and its closely related species (Poveda-Martínez et al., 2022). These findings suggest that intraspecific diversification within this group of closely related *Cactoblastis* species have been strongly influenced by shifts in the distribution of host species in response to Quaternary climate dynamics.

Footprints of historical admixture

Despite the aforementioned distributional shifts of host species resulting from Quaternary climate dynamics, population genetic analyses revealed evidence for gene flow among Chacoan (Central lineage) and North Pampean (East lineage) populations of *C. cactorum* . In particular, extensive nuclear admixture and sharing of mtDNA haplotypes occurred across those populations located along the contact area between northern Pampa and Chaco (Figure 2). This area coincides with a transition between Chaco and Pampa environments, referred to as the Espinal (Bucher, 1982). Historical secondary contact among previously isolated Central and East populations, likely prompted by post glacial *Opuntia* host range expansion, provides a plausible explanation for the observed pattern of genetic admixture. Evidence for admixture was also detected between populations of East and South lineages, both distributed within the Pampean and Monte regions. This area has been reported as a landscape corridor facilitating connectivity between previously separated populations of the grasshopper species *D. vittatus* inhabiting grassland and savanna biomes (Rosetti et al., 2022). Our results would further emphasize the extensive habitat connectivity between these regions during the Quaternary. Overall, our results argue against the idea that the human-driven introduction and intensive cultivation of *O. ficus-indica* enable rapid expansion of *C. cactorum* , promoting contemporary gene flow among geographically distant populations.

Landscape genomic analyses

In agreement with inferences of divergence times among three major lineages as inferred by coalescent-based demographic modeling, the landscape genomic analyses revealed that the spatial pattern of population genetic differentiation was best explained by a Quaternary landscape scenario representing the distribution of climatically suitable habitats and predicted ranges of host species during the LGM (Figure 1 and 3). The influence of historical landscape composition on the contemporary genomic variation pattern was illustrated by the fact that weakly differentiated populations were predicted to have had high habitat connectivity in LGM projections. This is the case of JUJ (northwestern) and FOG (northeastern) populations ($F_{ST=}0.088$; Table S6), currently separated by 740 km but subjected to high levels of past habitat connectivity according to suitability maps. In contrast, a discontinuous habitat was expected to promote isolation and, thus, genetic differentiation even between geographically proximate populations. This would be particularly so for *C. cactorum* , given its presumably limited dispersal ability because of their short-lived adult stage (Pettey, 1948; Zimmerman, Moran & Hoffmann, 2000), and illustrated by the southern populations LPS (southwestern) and BAP (southeastern) which exhibit one of the highest F_{ST} values (F_{ST}=0.389), despite the relatively limited geographic distance separating them of approximately 290 km. Environmental projections during LGM suggest that dispersal between LPS and BAP was likely limited due to low environmental suitability, thus limiting gene flow during glacial periods. Thus, our results would indicate that environmental tolerance together with limited dispersal could have interacted with landscape features to generate population genetic structure (Broquet & Petit, 2009; Sherpa et al., 2020).

Despite signatures for shifts in host species distributions, which may have influenced contemporary population genetic structure in *C. cactorum,* host species use appeared to be less of a consequential influence. This is at odds with observations in other insect herbivores (Forbes et al., 2017; Funk et al., 2006; Poveda-Martínez et al., 2020), but in line with field and laboratory host range studies (Varone et al 2014). In areas where there is more than one *Opuntia* species, host use by moths was proportional to host species abundance. Additionally, multiple choice experiments revealed that female *C. cactorum* do not exhibit oviposition preference for *Opuntia*species. Together, these results suggest that host species is not an important selective agent for *C. cactorum* , consistent with evidence from other generalist herbivores (Vidal & Murphy, 2018; Vidal et al., 2019).

Climatic stability and patterns of genomic diversity

Linear regression analyses revealed that population genetic diversity was significantly correlated with habitat stability. The Chaco region, representing the northernmost distribution limits of *C. cactorum* , is an open vegetation biome characterized by high endemicity and diversity for both plant and animal species (Bonatelli et al., 2022; Brusquetti, Netto, Baldo, & Haddad, 2019; Nores, 1994; Werneck, 2011). Analyses of environmental habitat suitability between the present and the LGM revealed relative habitat stability within the Chaco (Figure 4). Consistent with this, nuclear genetic diversity in *C. cactorum*was higher in the Chaco compared to southern regions, supporting the hypothesis that areas characterized by a high climate stability over the last glacial and interglacial periods tend to accumulate genetic diversity (Barros et al., 2015; Carnaval et al., 2009, Hewitt, 2004; Rocha et al., 2020). The higher effective population size for the Central lineage, as estimated by coalescent demographic modeling, is in line with the existence of Pleistocene refugia in Chaco. This region also contains a higher number of *Opuntia* species than in any other region in the sampled area (Varone et al., 2014). Such host plant availability together with long-term habitat stability has likely favoured the persistence of relatively large *C. cactorum*populations in this region. However, it cannot be discounted that large effective population sizes may also be consistent with recent demographic changes in the Chaco, favored by the introduction of*O. ficus-indica* . The introduction of the novel host species*O. ficus-indica* expanded rearing resources for *C. cactorum*as a consequence of its use as a crop species over the last centuries. Population growth of *C. cactorum* would have been further favored as *O. ficus-indica* can support one more generation per year as moths develop faster compared to development on native *Opuntia*species (Varone et al., 2019).

Biological control consequences

In addition to their evolutionary relevance, our findings are also relevant from a biological control perspective. Strategies have been implemented to monitor and control the cactus moth in both native and non-native ranges. These include the sterile male technique (Hight, Carpenter, Bloem & Bloem, 2005), and a pheromone-based attractant trap for males (Heath et al., 2006). These strategies are set to be complemented with biological control strategies using a parasitoid natural enemy from the native range of *C. cactorum* , currently under evaluation and being mass reared in quarantine in the United States (Mengoni-Goñalons et al., 2014; Varone et al. 2015; Varone et al. 2020).

The use of pheromone-based attractants for males, together with parasitoids, may be optimized by taking into account population genetic structure of the target species, together with patterns of gene flow, and the climatic factors that underpin range changes and population genetic differentiation. In the context of *C. cactorum* , a word of caution is warranted for the implementation of pheromone traps, as potential pheromone specificity may be associated with divergent lineages. In this case, the pheromone currently used in monitoring*C. cactorum* was developed based on virgin females from the East lineage (Heath et al., 2006). With the evidence of strong genetic structure within *C. cactorum* , additional field and laboratory experiments may be necessary to test the effectiveness of pheromone monitoring, specifically the Central and South lineages which are well differentiated and where the impact of the moth is especially high in Argentina (Varone et al., 2014). In a similar vein, although bioinsecticides are not currently recommended for the control of*C. cactorum* (Bloem, Mizell, Bloem, Hight & Carpenter, 2005), differentiated populations may also tend to exhibit differential susceptibility to bioinsecticides or synthetic insecticides (Ríos-Díez

& Saldamando-Benjumea, 2011; Arias et al., 2019). Taking into account population structure within *C. cactorum* may enhance the effectiveness of management strategies considering the specific lineages identified herein.

Acknowledgements

We are grateful to Mariel Guala and Malena Fuentes Corona for fieldwork support. We are also grateful to the Centro de Cómputo de Alto Rendimiento (CeCAR) and Biocódices S.A. for granting use of computational resources. Funding was obtained from FONCyT through grant PICT1447/2016 awarded to G.L. and USDA APHIS-PPQ, Farm Bill 10201. D.P.M. is the recipient of a PhD scholarship awarded by CONICET. V.N was supported by a Juan de la Cierva-Formación postdoctoral fellowship (grant FJC2018-035611-I) funded by MCIN/AEI/10.13039/501100011033. L.V. and E.H. are members of Carrera del Investigador CONICET.

References

Agrain, F. A., Domínguez, C. M., Carrara, R., Griotti, M., & Roig-Junent, S. A. (2021). Exploring the role of climatic niche changes in the evolution of the southern South American genus *Baripus*(Coleoptera: Carabidae): optimization of non-hereditary climatic variables and phylogenetic signal measurement. *Cladistics, 37* (6), 816-828.

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

Andrews, S., Krueger, F., Segonds-Pichon, A., Biggins, L., Krueger, C., & Wingett, S. (2010). FastQC. A quality control tool for high throughput sequence data, 370.

Arias, O., Cordeiro, E., Correa, A. S., Domingues, F. A., Guidolin, A. S., & Omoto, C. (2019). Population genetic structure and demographic history of *Spodoptera frugiperda* (Lepidoptera: Noctuidae): implications for insect resistance management programs. *Pest Management Science, 75* (11), 2948-2957.

Bandelt, H. J., Forster, P., & Rohl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution, 16* (1), 37-48.

Barros, M. J., Silva-Arias, G. A., Fregonezi, J. N., Turchetto-Zolet, A. C., Iganci, J. R., Diniz-Filho, J. A. F., & Freitas, L. B. (2015). Environmental drivers of diversity in Subtropical Highland Grasslands.*Perspectives in Plant Ecology, Evolution and Systematics, 17* (5), 360-368.

Bloem, S., R.F. Mizell, III, K.A. Bloem, S.D. Hight, & J.E. Carpenter. 2005. Laboratory evaluation of insecticides for control of the invasive Cactoblastis cactorum (Lepidoptera: Pyralidae). Florida Entomologist 88(4): 395-400.

Bonatelli, I. A., Gehara, M., Carstens, B. C., Colli, G. R., & Moraes, E. M. (2022). Comparative and predictive phylogeography in the South American diagonal of open formations: Unravelling the biological and environmental influences on multitaxon demography. *Molecular Ecology, 31* (1), 331-342.

Borer, M., Arrigo, N., Buerki, S., Naisbit, R. E., & Alvarez, N. (2012). Climate oscillations and species interactions: large-scale congruence but regional differences in the phylogeographic structures of an alpine plant and its monophagous insect. *Journal of Biogeography, 39* (8), 1487-1498.

Bouckaert, R., Heled, J., Kuhnert, D., Vaughan, T., Wu, C. H., Xie, D., & Drummond, A. J. (2014). BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology, 10* (4), e1003537.

Bouckaert, R. R., & Heled, J. (2014). DensiTree 2: seeing trees through the forest. BioRxiv, 012401.

Braconnot, P., Otto-Bliesner, B., Harrison, S., Joussaume, S., Peterchmitt, J. Y., Abe-Ouchi, A., ... & Zhao, Y. (2007). Results of PMIP2 coupled simulations of the Mid-Holocene and Last Glacial Maximum–Part 1: experiments and large-scale features. *Climate of the Past, 3* (2), 261-277.

Brooks, C. P., Ervin, G. N., Varone, L., & Logarzo, G. A. (2012). Native ecotypic variation and the role of host identity in the spread of an invasive herbivore, Cactoblastis cactorum. *Ecology, 93* (2), 402-410.

Broquet, T. & Petit, E. J. (2009). Molecular estimation of dispersal for ecology and population genetics. *Annual Review of Ecology, Evolution, and Systematics, 40* , 193–216

Brusquetti, F., Netto, F., Baldo, D., & Haddad, C. F. (2019). The influence of Pleistocene glaciations on Chacoan fauna: genetic structure and historical demography of an endemic frog of the South American Gran Chaco. *Biological Journal of the Linnean Society, 126* (3), 404-416.

Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. A., & RoyChoudhury, A. (2012). Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis.*Molecular Biology and Evolution, 29* (8), 1917-1932.

Bucher, E. H. (1982). Chaco and Caatinga—South American arid savannas, woodlands and thickets. In Ecology of tropical savannas (pp. 48-79). Springer, Berlin, Heidelberg.

Calatayud-Mascarell, A., Ferretti, N., Enguidanos, A., & Arnedo, M. A. 2022. Same place, different stories: Disparate evolutionary trends of mygalomorph spiders from the Peripampasic orogenic arc. *Journal of Biogeography* , *49* (7), 1234-1247.

Carnaval, A. C., Hickerson, M. J., Haddad, C. F., Rodrigues, M. T., & Moritz, C. (2009). Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. *Science, 323* (5915), 785-789.

Caterino, M. S., & Sperling, F. A. (1999). PapilioPhylogeny based on mitochondrial cytochrome oxidase I and II genes. *Molecular Phylogenetics and Evolution, 11* (1), 122-137.

Cobos, M. E., Peterson, A. T., Barve, N., & Osorio-Olvera, L. (2019). kuenm: an R package for detailed development of ecological niche models using Maxent. *PeerJ, 7* , e6281.

Cook, K. H., & Vizy, E. K. (2006). South American climate during the Last Glacial Maximum: delayed onset of the South American monsoon.*Journal of Geophysical Research: Atmospheres, 111* (D2).

Crisci-V, J., Freire-E, S., Sancho, G., & Katinas, L. (2001). Historical biogeography of the Asteraceae from Tandilia and Ventania mountain ranges (Buenos Aires, Argentina). *Caldasia, 23* (1), 21–41.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics, 27* (15), 2156-2158.

Driscoe, A. L., Nice, C. C., Busbee, R. W., Hood, G. R., Egan, S. P., & Ott, J. R. (2019). Host plant associations and geography interact to shape diversification in a specialist insect herbivore. *Molecular Ecology, 28* (18), 4197-4211.

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution, 29* (8), 1969-1973.

Eaton, D. A., & Overcast, I. (2020). ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics, 36* (8), 2592-2594.

Ervin, G. N. (2012). Indian fig cactus (*Opuntia ficus-indica* (L.) Miller) in the Americas: an uncertain history. *Haseltonia, 2012* (17), 70-81.

Ewels, P., Magnusson, M., Lundin, S., & Kaller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics, 32* (19), 3047-3048.

Excoffier, L., Dupanloup, I., Huerta-Sanchez, E., Sousa, V. C., & Foll, M. (2013). Robust demographic inference from genomic and SNP data.*PLoS Genetics, 9* (10), e1003905.

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.

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.

Forbes, A. A., Devine, S. N., Hippee, A. C., Tvedte, E. S., Ward, A. K., Widmayer, H. A., & Wilson, C. J. (2017). Revisiting the particular role of host shifts in initiating insect speciation. *Evolution, 71* (5), 1126-1137.

Frichot, E., & Francois, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution, 6* (8), 925-929.*https://doi.org/10.1111/2041*

Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G., & Francois, O. (2014). Fast and efficient estimation of individual ancestry coefficients. *Genetics, 196* (4), 973-983.*https://doi.org/10.1534/genetics.113.160572*

Funk, D. J., Nosil, P., & Etges, W. J. (2006). Ecological divergence exhibits consistently positive associations with reproductive isolation across disparate taxa. *Proceedings of the National Academy of Sciences, 103* (9), 3209-3213.*https://doi.org/10.1073/pnas.0508653103*

Goudet, J. (2005). Hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes, 5* (1), 184-186.

Guo, Q. (2012). Incorporating latitudinal and central–marginal trends in assessing genetic variation across species ranges. *Molecular Ecology, 21* (22), 5396-5403.

Hall, K. R., Anantharaman, R., Landau, V. A., Clark, M., Dickson, B. G., Jones, A., ... & Shah, V. B. (2021). Circuitscape in Julia: empowering dynamic approaches to connectivity assessment. *Land, 10* (3), 301.

Heath, R. R., Teal, P. E., Epsky, N. D., Dueben, B. D., Hight, S. D., Bloem, S., ... & Bloem, K. A. (2006). Pheromone-based attractant for males of *Cactoblastis cactorum* (Lepidoptera: Pyralidae).*Environmental Entomology, 35* (6), 1469-1476.

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages.*Nature, 405* (6789), 907-913.*https://doi.org/10.1038/35016000*

Hewitt, G. M. (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 359* (1442), 183-195.

Hight, S. D., & Carpenter, J. E. (2009). Flight phenology of male*Cactoblastis cactorum* (Lepidoptera: Pyralidae) at different latitudes in the southeastern United States. *Florida Entomologist, 92* (2), 208-216.

Hight, S. D., Carpenter, J. E., Bloem, S., & Bloem, K. A. (2005). Developing a sterile insect release program for *Cactoblastis cactorum* (Berg) (Lepidoptera: Pyralidae): Effective overflooding ratios and release-recapture field studies. *Environmental Entomology, 34* (4), 850-856.

Hijmans, R. J., & van Etten, J. (2016). raster: Geographic data analysis and modeling. R package, 734.

Hoffmann, J. H., Moran, V. C., Zimmermann, H. G., & Impson, F. A. (2020). Biocontrol of a prickly pear cactus in South Africa: Reinterpreting the analogous, renowned case in Australia. *Journal of Applied Ecology, 57* (12), 2475-2484.

Julien, MH and Griffiths, MW (1998), Biological Control of Weeds a World Catalogue of Agents and their Target Weeds, 4th ed., CABI Publishing, Wallingford, UK

Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W., & Prodohl, P. A. (2013). diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors.*Methods in Ecology and Evolution, 4* (8), 782-788.

Keightley, P. D., Pinharanda, A., Ness, R. W., Simpson, F., Dasmahapatra, K. K., Mallet, J., ... & Jiggins, C. D. (2015). Estimation of the spontaneous mutation rate in *Heliconius melpomene* . *Molecular Biology and Evolution, 32* (1), 239-243.

Kohler, M., Esser, L. F., Font, F., Souza-Chies, T. T., & Majure, L. C. (2020). Beyond endemism, expanding conservation efforts: What can new distribution records reveal?. *Perspectives in Plant Ecology, Evolution and Systematics, 45* , 125543.

Laukkanen, L., Mutikainen, P., Muola, A., & Leimu, R. (2014). Plant-species diversity correlates with genetic variation of an oligophagous seed predator. *PLoS One, 9* (4), e94105.

Lavinia, P. D., Barreira, A. S., Campagna, L., Tubaro, P. L., & Lijtmaer, D. A. (2019). Contrasting evolutionary histories in Neotropical birds: Divergence across an environmental barrier in South America. *Molecular Ecology, 28* (7), 1730-1747.

Leigh, J. W., & Bryant, D. (2015). POPART: full-feature software for haplotype network construction. *Methods in Ecology and Evolution, 6* (9), 1110-1116.

Lynch, M., & Conery, J. S. (2003). The origins of genome complexity.*Science, 302* (5649), 1401–1404.

Marsico, T. D., Wallace, L. E., Ervin, G. N., Brooks, C. P., McClure, J. E., & Welch, M. E. (2011). Geographic patterns of genetic diversity from the native range of *Cactoblastis cactorum* (Berg) support the documented history of invasion and multiple introductions for invasive populations. *Biological Invasions, 13* (4), 857-868.

McFadyen, R. E. (1985). Larval characteristics of Cactoblastis spp.(Lepidoptera: Pyralidae) and the selection of species for biological control of prickly pears (*Opuntia* spp.). *Bulletin of Entomological Research, 75* (1), 159-168.

McRae, B. H., & Beier, P. (2007). Circuit theory predicts gene flow in plant and animal populations. *Proceedings of the National Academy of Sciences, 104* (50), 19885-19890.

McRae, B. H. (2006). Isolation by resistance. *Evolution, 60* (8), 1551-1561.

Mengoni-Gonalons, C. M., Varone, L., Logarzo, G., Guala, M., Rodriguero, M., Hight, S. D., & Carpenter, J. E. (2014). Geographical range and laboratory studies on *Apanteles opuntiarum* (Hymenoptera: Braconidae) in Argentina, a candidate for biological control of*Cactoblastis cactorum* (Lepidoptera: Pyralidae) in North America.*Florida Entomologist* , 1458-1468.

Morrone, J. J. (2014). Biogeographical regionalisation of the Neotropical region. *Zootaxa, 3782* (1), 1-110.

Mourelle, C., & Ezcurra, E. (1997). Differentiation diversity of Argentine cacti and its relationship to environmental factors.*Journal of Vegetation Science, 8* (4), 547-558.

Noguerales, V., Cordero, P. J., & Ortego, J. (2018). Inferring the demographic history of an oligophagous grasshopper: Effects of climatic niche stability and host-plant distribution. *Molecular Phylogenetics and Evolution, 118* , 343-356.

Noguerales, V., Meramveliotakis, E., Castro-Insua, A., Andujar, C., Arribas, P., Creedy, T. J., ... & Papadopoulou, A. (2021). Community metabarcoding reveals the relative role of environmental filtering and spatial processes in metacommunity dynamics of soil microarthropods across a mosaic of montane forests. *Molecular Ecology .https://doi.org/10.1111/mec.16275*

Nores, M. (1994). Quaternary vegetational changes and bird differentiation in subtropical South America. *The Auk, 111* (2), 499-503.

Ochoa, M. J., Targa, M. G., Abdala, G., Leguizamon, G. (2007, October). Extending fruiting season of cactus pear (*Opuntia ficus-indica*(l.) miller) in Santiago del estero, Argentina. In VI International Congress on Cactus Pear and Cochineal 811 (pp. 87-90).

Oliveira, E. A. D., Perez, M. F., Bertollo, L. A. C., Gestich, C. C., Rab, P., Ezaz, T., ... & Cioffi, M. B. (2020). Historical demography and climate driven distributional changes in a widespread Neotropical freshwater species with high economic importance. *Ecography, 43* (9), 1291-1304.

Ortego, J., Cespedes, V., Millan, A., & Green, A. J. (2021). Genomic data support multiple introductions and explosive demographic expansions in a highly invasive aquatic insect. *Molecular Ecology, 30* (17), 4189-4203.

Ortego, J., Gugger, P. F., Riordan, E. C., & Sork, V. L. (2014). Influence of climatic niche suitability and geographical overlap on hybridization patterns among southern Californian oaks. J*ournal of Biogeography, 41* (10), 1895-1908.

Ortiz-Jaureguizar, E., & Cladera, G. A. (2006). Paleoenvironmental evolution of southern South America during the Cenozoic. *Journal of Arid Environments, 66* (3), 498-532.

Pembleton, L. W., Cogan, N. O., & Forster, J. W. (2013). StAMPP: An R package for calculation of genetic differentiation and structure of mixed ploidy level populations. *Molecular Ecology Resources, 13* (5), 946-952.

Peterman, W. E., Connette, G. M., Semlitsch, R. D., & Eggert, L. S. (2014). Ecological resistance surfaces predict fine-scale genetic differentiation in a terrestrial woodland salamander. *Molecular Ecology, 23* (10), 2402-2413.

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PloS ONE, 7* (5), e37135.

Pettey, F. W. (1948) The biological control of prickly pear in South Africa. Science Bulletin, Department of Agriculture of the Union of South Africa 271: 1–163

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling, 190* (3-4), 231-259.

Poveda-Martinez, D., Aguirre, M. B., Logarzo, G., Hight, S. D., Triapitsyn, S., Diaz-Sotero, H., ... & Hasson, E. (2020). Species complex diversification by host plant use in an herbivorous insect: The source of Puerto Rican cactus mealybug pest and implications for biological control. *Ecology and Evolution, 10* (19), 10463-10480.

Poveda-Martinez, D., Varone, L., Fuentes Corona, M., Hight, S., Logarzo, G., & Hasson, E. (2022). Spatial and host related genomic variation in partially sympatric cactophagous moth species. *Molecular Ecology, 31* (1), 356-371.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., ... & Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics, 81* (3), 559-575.

Raghu, S., & Walton, C. (2007). Understanding the ghost of *Cactoblastis* past: historical clarifications on a poster child of classical biological control. *BioScience, 57* (8), 699-705.

Rios-Diez, J. D., & Saldamando-Benjumea, C. I. (2011). Susceptibility of *Spodoptera frugiperda* (Lepidoptera: Noctuidae) strains from central Colombia to two insecticides, methomyl and lambda-cyhalothrin: a study of the genetic basis of resistance. *Journal of Economic Entomology, 104* (5), 1698-1705.

Rocha, A. V., Cabanne, G. S., Aleixo, A., Silveira, L. F., Tubaro, P., & Caparroz, R. (2020). Pleistocene climatic oscillations associated with landscape heterogeneity of the South American dry diagonal explains the phylogeographic structure of the narrow-billed woodcreeper (Lepidocolaptes angustirostris, Dendrocolaptidae). *Journal of Avian Biology, 51* (9), 1–13.

Rodriguero, M. S., Lanteri, A. A., Guzman, N. V., Carus Guedes, J. V., & Confalonieri, V. A. (2016). Out of the forest: past and present range expansion of a parthenogenetic weevil pest, or how to colonize the world successfully. *Ecology and Evolution, 6* (15), 5431-5445.

Rosetti, N., Krohling, D., & Remis, M. I. (2022). Evolutionary history and colonization patterns of the wing dimorphic grasshopper*Dichroplus vittatus* in two Argentinean biomes. *Scientific Reports, 12* (1), 1-17.

Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., & Sanchez-Gracia, A. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution, 34* (12), 3299-3302.

Sherpa, S., Renaud, J., Gueguen, M., Besnard, G., Mouyon, L., Rey, D., & Despres, L. (2020). Landscape does matter: Disentangling founder effects from natural and human aided post introduction dispersal during an ongoing biological invasion. *Journal of Animal Ecology, 89* (9), 2027-2042.

Simmonds, F. J., & Bennett, F. D. (1966). Biological control of*Opuntia* spp. by *Cactoblastis cactorum* in the Leeward Islands (West Indies). *Entomophaga 11* : 183-189.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., & Flook, P. (1994). Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America, 87* (6), 651-701.

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

Soley-Guardia, M., Carnaval, A. C., & Anderson, R. P. (2019). Sufficient versus optimal climatic stability during the Late Quaternary: using environmental quality to guide phylogeographic inferences in a Neotropical montane system. *Journal of Mammalogy, 100* (6), 1783-1807.

Speranza, P. R., Seijo, J. G., Grela, I. A., & Solis Neffa, V. G. (2007). Chloroplast DNA variation in the *Turnera sidoides* L. complex (Turneraceae): biogeographical implications. *Journal of Biogeography, 34* (3), 427-436.

Turchetto-Zolet, A. C., Pinheiro, F., Salgueiro, F., & Palma-Silva, C. (2013). Phylogeographical patterns shed light on evolutionary process in South America. *Molecular Ecology, 22* (5), 1193-1213.

Varone, L., Acosta, M. M., Logarzo, G. A., Briano, J. A., Hight, S. D., & Carpenter, J. E. (2012). Laboratory performance of *Cactoblastis cactorum* (Lepidoptera: Pyralidae) on South and North American*Opuntia* species occurring in Argentina. *Florida Entomologist* , 1163-1173.

Varone, L., Aguirre, M. B., Lobos, E., Ruiz Perez, D., Hight, S. D., Palottini, F., ... & Logarzo, G. A. (2019). Causes of mortality at different stages of *Cactoblastis cactorum* in the native range.*BioControl, 64* (3), 249-261.

Varone, L., Logarzo, G., Martinez, J. J., Navarro, F., Carpenter, J. E., & Hight, S. D. (2015). Field host range of *Apanteles opuntiarum*(Hymenoptera: Braconidae) in Argentina, a potential biocontrol agent of*Cactoblastis cactorum* (Lepidoptera: Pyralidae) in North America.*Florida Entomologist, 98* (2), 803-806.

Varone, L., Logarzo, G. A., Briano, J. A., Hight, S. D., & Carpenter, J. E. (2014). *Cactoblastis cactorum* (Berg) (Lepidoptera: Pyralidae) use of *Opuntia* host species in Argentina.*Biological Invasions, 16* (11), 2367-2380.*https://doi.org/10.1007/s10530-014-0670-9*

Varone, L., Mengoni Gonalons, C., Faltlhauser, A. C., Guala, M. E., Wolaver, D., Srivastava, M., & Hight, S. D. (2020). Effect of rearing*Cactoblastis cactorum* on an artificial diet on the behaviour of*Apanteles opuntiarum. Journal of Applied Entomology, 144* (4), 278-286.

Vidal, M. C., & Murphy, S. M. (2018). Bottom-up vs. top-down effects on terrestrial insect herbivores: A meta-analysis. *Ecology Letters, 21* (1), 138–150.*https://doi.org/10.1111/ele.12874*

Vidal, M. C., Quinn, T. W., Stireman III, J. O., Tinghitella, R. M., & Murphy, S. M. (2019). Geography is more important than host plant use for the population genetic structure of a generalist insect herbivore.*Molecular Ecology, 28* (18), 4317-4334.

Wang, I. J., & Bradburd, G. S. (2014). Isolation by environment. *Molecular Ecology, 23* (23), 5649-5662.

Wang, I. J. (2013). Examining the full effects of landscape heterogeneity on spatial genetic variation: a multiple matrix regression approach for quantifying geographic and ecological isolation. *Evolution, 67* (12), 3403-3411.

Wang, I. J. (2020). Topographic path analysis for modelling dispersal and functional connectivity: calculating topographic distances using the topoDistance r package. *Methods in Ecology and Evolution, 11* (2), 265-272.

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* , 1358-1370.

Werneck, F. P. (2011). The diversification of eastern South American open vegetation biomes: historical biogeography and perspectives. *Quaternary Science Reviews, 30* (13-14), 1630-1648.

Zimmerman, H. G., Moran, V. C., & Hoffmann, J. A. (2000) The renowned cactus moth, *Cactoblastis cactorum* : its natural history and threat to native Opuntia floras in Mexico and the United States of America. *Diversity and Distributions 6* : 259–269

Data Availability

The datasets generated during the current study for each one of the analyses are available in Figshare (*https://figshare.com/s/a72fadc5a273aecbf346*). Raw reads are available at NCBI as BioProject PRJNA666743. The reference draft genome is available at the NCBI under accession number JADGIL010000000. Mitochondrial haplotypes are available at NCBI under accession numbers OM176592-OM176657. *Cactoblastis cactorum* and host species occurrences were retrieved from proper field records as well as Gbif records (*https://www.gbif.org*).

Supplementary data

Supplementary File

Author contributions

D.P.M., L.V., G.L., and E.H. designed the project. D.P.M., L.V., E.H., S.D.H., and G.L. performed field collections. D.P.M., V.N., B.E., and E.H. delineate the strategy of population genomics, demographic inference and landscape genomics analyses. D.P.M. conducted molecular work and data analyses. D.P.M., and E.H. drafted the manuscript. All coauthors made contributions to the final version of the manuscript.

Tables

Table 1. Parameters inferred from coalescent simulations with Fastsimcoal2 under the best-supported demographic model (Model 8; Figure 3). For each parameter, point estimate and lower and upper 95% confidence intervals are shown. T_{DIV1} and T_{DIV2}, population divergence; NE_{ANC1} and NE_{ANC2}, historical and contemporary effective population sizes; M_{SC}, M_{CS}, M_{ES}, M_{SE}, M_{CE}, and M_{EC}, asymmetric rates of gene flow among lineages.

Model 8	**Model 8**	**95% Confidence interval**	**95% Confidence interval**
Parameter	**Point estimate**	**Lower bound**	**Upper bound**
$\mathbf{NE_{ANC1}}$	3,504,318	2,576,620	3,858,949
$\mathbf{NE_{ANC2}}$	11,746	3,773	21,102
$\mathbf{NE_{CENTER}}$	245,957	208,570	380,730
$\mathbf{NE_{EAST}}$	41,551	35,596	54,151
$\mathbf{T_{DIV2}}$	75,513	61,524	120,375
$\mathbf{T_{DIV1}}$	19,207	17,202	23,310

Model 8	**Model 8**	**95% Confidence interval**	**95% Confidence interval**
$\mathbf{M_{ES}}$	1.18 X 10^{-6}	8.86 X 10^{-7}	1.75 X 10^{-6}
$\mathbf{M_{SE}}$	4.68 X 10^{-8}	5.41 X 10^{-11}	1.65 X 10^{-6}
$\mathbf{M_{CS}}$	2.71 X 10^{-8}	3.55 X 10^{-11}	6.83 X 10^{-8}
$\mathbf{M_{SC}}$	6.37 X 10^{-6}	4.37 X 10^{-6}	7.32 X 10^{-6}
$\mathbf{M_{CE}}$	1.50 X 10^{-5}	1.23 X 10^{-5}	1.67 X 10^{-5}
$\mathbf{M_{EC}}$	2.98 X 10^{-5}	3.78 X 10^{-11}	1.42 X 10^{-6}

Table 2. Results of multiple matrix regression with randomization (MMRR) testing for the relationships between population genetic differentiation (F_{ST}) and explanatory variables representing weighted topographic distance (isolation by distance, IBD_{WTD}), climatic dissimilarity (isolation by environment, IBE_{CLI}) and alternative isolation by resistance (IBR) scenarios. IBR scenarios consist of current and past habitat suitability based on (i) only climatic variables (*$Current_{ENV}$, and LGM_{ENV}), (ii) climate-based habitat suitability maps for host species ($Current_{HOST}$, LGM_{HOST}), and (iii) combining climatic variables climate-based habitat suitability maps for host species (*$Current_{ENV\text{-}HOST}$ and $LGM_{ENV\text{-}HOST}$). IBR_{NULL} represents an IBR scenario where all pixel values are equal to 1 (flat landscape). β, regression coefficient; t , t -statistic; p , significance level.

Variable	β	t	p
Explanatory term	**Explanatory term**	**Explanatory term**	**Explanatory term**
$LGM_{ENV\text{-}HOST}$	0.416	16.365	0.001
Rejected terms	**Rejected terms**	**Rejected terms**	**Rejected terms**
Climatic dissimilarity (IBE_{CLI})		-0.476	0.831
Topographic distance (IBD_{WTD})		-2.058	0.301
IBR - $Current_{HOST}$		3.251	0.278
IBR - LGM_{ENV}		4.808	0.117
IBR - LGM_{HOST}		-5.834	0.087
Flat landscape - IBR_{NULL}		-3.654	0.278

*$Current_{ENV}$ and $Current_{ENV\text{-}HOST}$ were excluded from the initial full model due to these terms resulted non-significant in their respective univariate analysis.

Table 3. Results of linear regressions testing for the relationships between population genetic diversity, and habitat stability models during the last 21,000 years, according to ENMs based on (i) only climatic variables ($Stability_{ENV}$), (ii) climate-based habitat suitability maps for host species ($Stability_{HOST}$), and (iii) combining climatic variables with climate-based habitat suitability maps for host species ($Stability_{ENV\text{-}HOST}$). Correlations with geographical variables (latitude and longitude) were also assessed. Genetic diversity was calculated for both nuclear (nDNA) and mitochondrial (mtDNA) data. R^2, regression coefficient; p , significance level.

	$Stability_{ENV}$	$Stability_{ENV}$	$Stability_{ENV}$	$Stability_{HOST}$	$Stability_{HOST}$	$Stability_{HOS}$
nDNA	R^2	p	R^2	R^2	p	R^2
Expected heterozygosity (H_E)	0.466	0.000	0.292	0.292	0.002	0.245
Nucleotide diversity (π)	0.509	0.000	0.266	0.266	0.004	0.219
mtDNA						
Haplotype diversity (H_D)	0.084	0.133	0.021	0.021	0.453	0.031
Nucleotide diversity (π)	0.076	0.154	0.005	0.005	0.722	0.008

Figure legends

Figure 1. Geographic location and host species of the *Cactoblastis cactorum* populations sampled in this study (A). Regions represent biogeographic provinces according to Morrone et al. (2014). Panels (B) and (C) represent the maps of habitat suitability for *C. cactorum* during the present (B) and during the Last Glacial Maximum (LGM, ca. 21ky) (C) as inferred by the most-supported ENM model in MAXENT. For visual representation, we only displayed ENM results based on the model constructed with the climate-based habitat suitability maps for host species ($\text{Current}_{\text{HOST}}$), which resulted with the lowest AIC score in model comparison. Inferences from alternative models are shown in Figure S5.

Figure 2 . Population genetic structure and phylogenetic relationships of the *C. cactorum* populations sampled in this study, using nuclear and mitochondrial genetic data. Each genetic group is represented by the same color across panels. (A) Pie charts represent the average ancestry coefficient of individuals belonging to each of 28 sampling sites distributed along three biogeographic regions according to Morrone (2014). (B) Species tree reconstructed by the coalescent method SNAPP using a matrix of 2,086 unlinked SNPs; numbers in nodes denote posterior probability for the most supported topology based on the maximum credibility tree. (C) Results of individual assignment in genetic clusters using sNMF method. Vertical bars represent the ancestry coefficient of each individual to the corresponding genetic cluster. (D) Principal component analysis (PCAs) using the two major axes of genomic variation. (E) Median-joining network obtained with mtDNA data. The mitochondrial network consisted of 66 haplotypes grouped in at least six haplogroups.

Figure 3. Visual representation of the most supported demographic model (Model 8) estimated using Fastsimcoal2. Parameter estimates included timing of population divergence (T_{DIV1} and T_{DIV2}), historical and contemporary effective population sizes (NE_{ANC1} and NE_{ANC2}), and asymmetric rates of gene flow (M_{SC}, M_{CS}; M_{ES}, M_{SE}; M_{CE}, and M_{EC}).

Figure 4. Spatial pattern of population genetic diversity for *Cactoblastis cactorum* . The map shows the projection of climatic and habitat stability through time, from LGM (˜21 Kya) to the present, for the three ENM models based on (i) only climatic variables ($\text{Stability}_{\text{ENV}}$) (A), (ii) climate-based habitat suitability maps for host species ($\text{Stability}_{\text{HOST}}$) (B), and (iii) combining climatic variables climate-based habitat suitability maps for host species ($\text{Stability}_{\text{ENV-HOST}}$) (C). Circle sizes denote varying levels of population genetic diversity as indicated by nuclear [(A). H_{E}, expected heterozygosity; (B). π, nucleotide diversity) and mitochondrial (C). H_{D}, haplotype diversity)]. Bottom panels show the significantly positive relationships between a given genetic diversity index and latitude.