

# Darwin's Overlooked Radiation: genomic evidence points to the early stages of a radiation in the Galápagos prickly pear cactus (*Opuntia*, Cactaceae)

Felipe Zapata<sup>1,2,11,^,\*</sup>, José Cerca<sup>3,^,\*</sup>, Dana McCarney<sup>1,4</sup>, Claudia L. Henriquez<sup>1</sup>, Bashir B. Tihamiyu<sup>5</sup>, John E. McCormack<sup>6</sup>, Kelsey R. Reckling<sup>1,6</sup>, Jaime A. Chaves<sup>7,8,9</sup>, and Gonzalo Rivas-Torres<sup>8,9,10\*</sup>

<sup>1</sup>Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA 90095, USA

<sup>2</sup>Center for Tropical Research, Institute of the Environment and Sustainability, University of California, Los Angeles, CA 90095, USA

<sup>3</sup>Center for Ecological and Evolutionary Synthesis (CEES), Department of Biosciences, University of Oslo, Oslo, Norway

<sup>4</sup>School of Medicine, Tufts University, Boston, MA 02155, USA

<sup>5</sup>Department of Plant Biology, University of Ilorin, Nigeria

<sup>6</sup>Moore Laboratory of Zoology, Occidental College, Los Angeles, CA 90041, USA

<sup>7</sup>Department of Biology, San Francisco State University, San Francisco, CA 94132, USA

<sup>8</sup>Colegio de Ciencias Biológicas y Ambientales and Galápagos Institute for the Arts and Sciences, Universidad de San Francisco de Quito USFQ, Quito 170901, Ecuador

<sup>9</sup>Galápagos Science Center UNC—USFQ, San Cristóbal, Galápagos, Ecuador

<sup>10</sup>Department of Geography and Environment, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

<sup>11</sup>Lead Contact

^Equal Contribution

\*Correspondence: [fzapata@ucla.edu](mailto:fzapata@ucla.edu) (F.Z.), [jose.cerca@gmail.com](mailto:jose.cerca@gmail.com) (J.C.), [grivast@usfq.edu.ec](mailto:grivast@usfq.edu.ec) (G.R.T.)

# SUMMARY

The Galápagos Islands are a prime example of a natural laboratory for the study of evolutionary radiations. While much attention has been devoted to iconic species like Darwin's finches<sup>1-4</sup>, the islands offer an equally unique but often overlooked opportunity for plant radiations<sup>5</sup>. Yet, compared to their animal counterparts, our understanding of the patterns and processes underpinning Galápagos plant radiations remains relatively limited<sup>6,7</sup>. We present evidence of the early stages of a radiation in prickly-pear cactus (*Opuntia*, Cactaceae), a plant lineage widespread across the archipelago. Phylogenomic and population genomic analyses show that notwithstanding overall low genetic differentiation across populations, there is marked geographic structure that is broadly consistent with current taxonomy and the dynamic paleogeography of the Galápagos. Because such low genetic differentiation stands in stark contrast to the exceptional eco-phenotypic diversity displayed by cacti across islands, it is plausible that phenotypic plasticity precedes genetic divergence and is the source of adaptive evolution, or that introgression between populations facilitates local adaptation. Models of population relationships including admixture indicate that gene flow is common between certain islands, likely facilitated by dispersal via animals known to feed on *Opuntia* flowers, fruits, and seeds across the archipelago. Scans of genetic differentiation between populations reveal candidate loci associated with seed traits and environmental stressors, suggesting that a combination of biotic interactions and abiotic pressures due to the harsh conditions characterizing island life in a volcanic, equatorial archipelago may underlie the diversification of prickly-pear cacti. Considered in concert, these results are relevant to both the mechanisms of plant eco-phenotypic differentiation and the evolutionary history and conservation of the Galápagos biota.

# RESULTS AND DISCUSSION

“ I certainly recognise S. America in ornithology, would a botanist? ”

C. R. Darwin (1835, Field Notes, p 30) soon after arriving to the Galápagos and noticing the mockingbirds, which he could associate with birds from mainland South America. He doubted the same could be said about the plants.

The origins, persistence, and diversification of species have been studied extensively in oceanic islands where ecological communities are often less complex and environmental conditions provide variation for evolution to flourish<sup>1,8–10</sup>. The Galápagos Islands are a prime example of such natural laboratories for studying evolution. Studies on the diversification of Galápagos tortoises<sup>11,12</sup>, iguanas<sup>13,14</sup>, land snails<sup>15</sup>, and Darwin’s finches<sup>2,16–19</sup> have provided fundamental biological insight about foundational ideas in evolutionary biology, including speciation, natural selection, and adaptive radiation.

Strikingly, and in spite of a relatively well-known flora, studies about the diversification of Galápagos plant lineages have received considerably less attention, with the exception of a few classic studies on prickly-pear cacti<sup>20–22</sup> and recent work on the salt bush<sup>23</sup> and the Darwin’s Giant Daisies<sup>6,7</sup>. This is surprising given the central role that the plants from the Galápagos, particularly the cacti, played in shaping Darwin’s ideas for his theory of evolution<sup>5</sup> and in the dietary ecology and diversification of the iconic Darwin’s finches<sup>2,3,19,24,25</sup>.

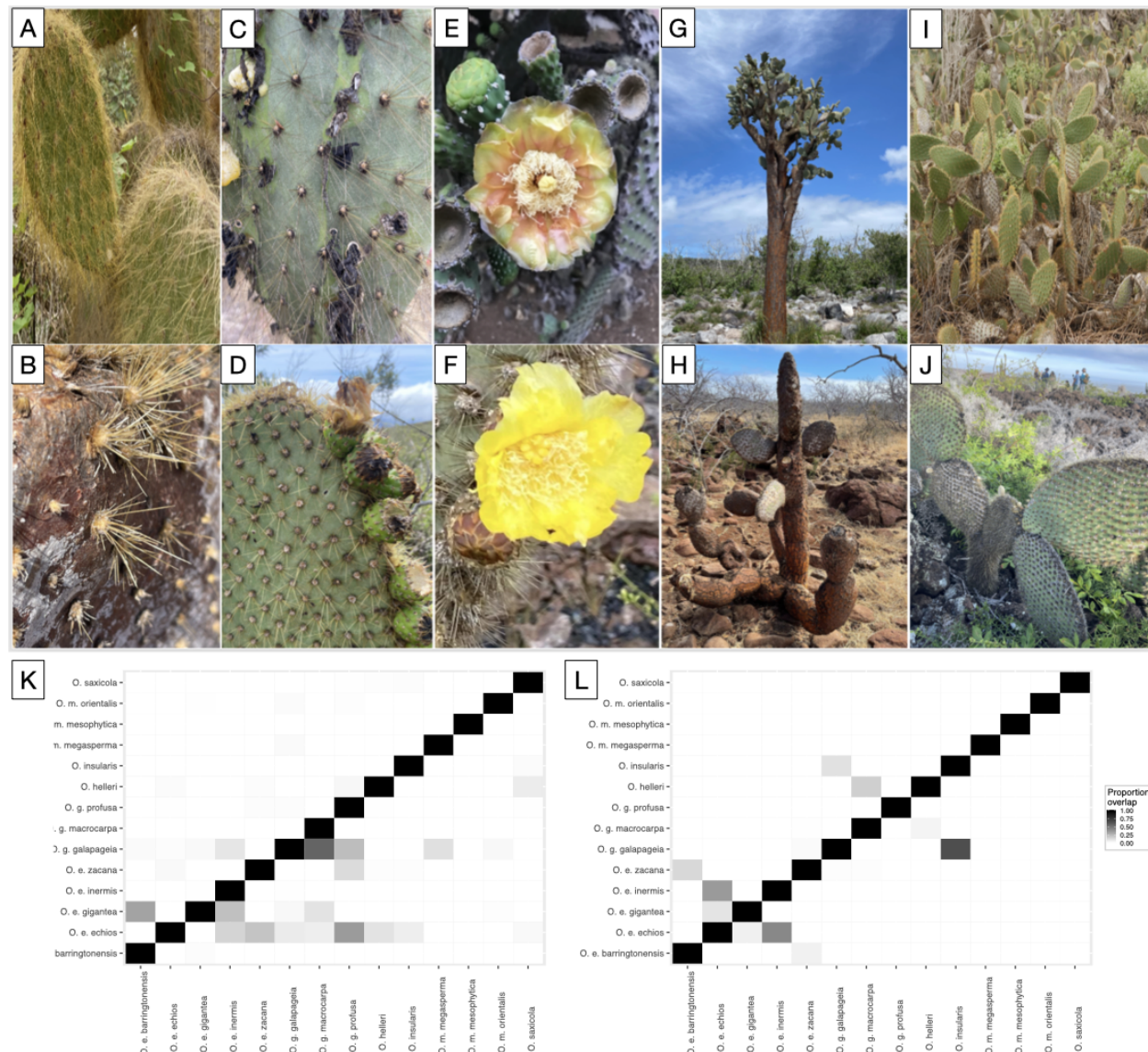
Here, we use genomic data to revisit classic work on the prickly-pear cacti (*Opuntia*) across the Galápagos archipelago to investigate the evolution of this charismatic lineage of the Galápagos flora (Fig. 1). Given that these plants furnish an important source of food and water to multiple native animals across the archipelago<sup>5,26</sup> and are critical in the ecology and persistence of some emblematic

animal radiations<sup>3</sup>, *Opuntia* can be considered a keystone element of the Galápagos biota.

Understanding the dynamics of the overlooked radiation of prickly pear cacti is therefore essential, not only for revealing the biological impacts of its evolution in other lineages but also for designing conservation strategies to protect the biodiversity of the Galápagos islands as a whole.

## Genomic lineages of *Opuntia* in the Galápagos

The prickly pear cacti in the Galápagos islands display outstanding eco-phenotypic variation, including 4-fold differences in plant height, 100-fold differences in seed size and seed number per fruit, and 5-fold differences in seed hardness<sup>24,27</sup>. In addition, species and populations vary extensively in the size of pads as well as the density and morphology of spines<sup>5</sup> (Fig. 1). This pronounced phenotypic variation has led to controversies around the nature of species in Galápagos *Opuntia*. Six species are recognized, with three species further subdivided into multiple varieties, for a total of fourteen taxa<sup>21</sup>. Notably, each recognized taxon is virtually restricted to a single large to mid-size island, with the exception of *O. helleri* which is widespread across multiple islands. Previous work based on the study of variation in eight allozymes, ten microsatellites, and two sequenced loci suggested that despite the remarkable phenotypic and ecological variation of *Opuntia* in the Galápagos, there is exceptionally low genetic variation across this radiation<sup>20–22</sup>. Therefore, whether each of the different recognized taxa represents independently evolving lineages remains controversial.



**Figure 1. *Opuntia* vary extensively in phenotypic and ecological traits across the Galápagos islands.**

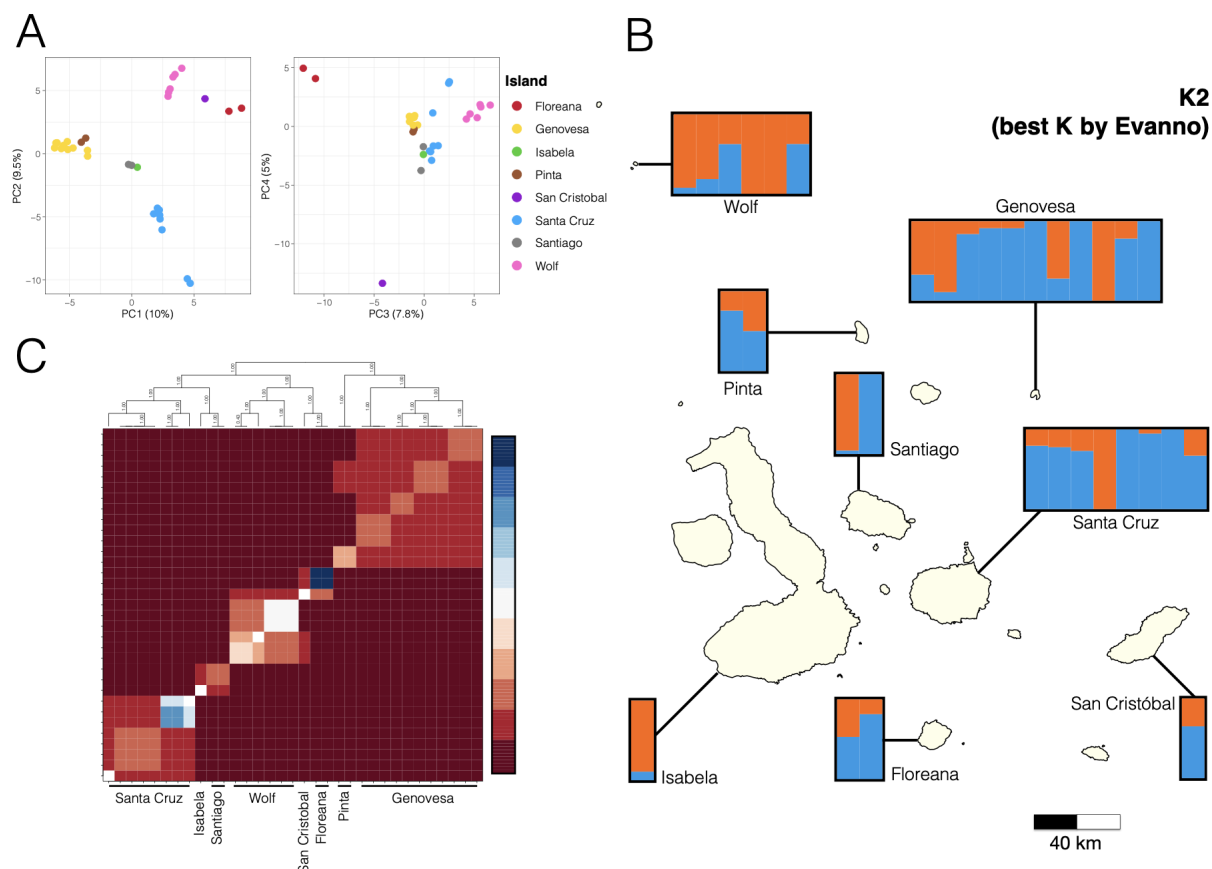
Spines: (A) *Opuntia megasperma* var. *megasperma* from Floreana (photo J.A.C.), (B) *O. echios* var. *zacana* from Seymour (photo G.R.T.), (C) *O. galapageia* var. *galapageia* from Santiago (photo G.R.T.), (D) *O. saxicola* from Isabela, Volcan Chico (photo J.E.M.); Flowers: (E) *O. megasperma* var. *mesophytica* from San Cristóbal (photo G.R.T.), (F) *O. saxicola* from Isabela, Volcan chico (photo J.A.C.); Habit: (G) *O. echios* var. *barringtonensis* from Santa Fe (photo J.E.M.), (H) *O. echios* var. *zacana* from North Seymour (photo G.R.T.), (I) *O. helleri* from Wolf (photo J.A.C.), (J) *O. insularis* from Isabela (photo G.R.T.). (K) Pairwise overlap among hypervolumes describing *Opuntia* taxa in the Galápagos in multidimensional vegetative morphospace (8 traits). Most of the overlap occurs between subspecies of the same species. (L) Pairwise overlap among hypervolumes describing *Opuntia* taxa in the Galápagos in multidimensional reproductive morphospace (5 traits). Most of the overlap occurs between subspecies of the same species

To examine the population genetic structure of *Opuntia* in the Galápagos islands, we used 3RAD sequencing<sup>28</sup> on 37 individuals representing seven lineages and eight islands, and 2 individuals belonging to an outgroup lineage. Overall, we found that while there is some genetic differentiation among cacti populations across the archipelago, such differentiation is not highly pronounced. The first two principal-component axes<sup>29</sup> accounted for only 19.5% (PC1 10% of variance explained; PC2 9.5% of variance explained; Fig 2A) of the variation in the multivariate genomic space across the Galápagos islands. Using up to four principal-component axes accounted for 32.5% of the variation (Fig 2A). Together, these findings indicate that the genomic differentiation of cacti populations across the Galápagos islands, while low, nonetheless reveals some geographic patterns. For instance, PC1 and PC2 broadly separated populations along latitude (Floreana and San Cristóbal separate from Pinta and Genovesa, and Santa Cruz separate from Wolf) while populations at the center of the archipelago (Santiago and Isabela) clustered together in PC space (Fig 2A). Genomic ancestry analyses<sup>30</sup> also revealed low genetic differentiation of populations across the archipelago (Fig. 2B). The optimal partitioning of the data indicated the existence of two demes with all the individuals displaying mixed ancestry, irrespective of island of origin or taxonomic assignment (Supplementary Figure S1). This suggests considerable admixture or recent ancestry for all sampled populations.

To characterize fine-scale population structure, we harnessed the power and resolution afforded by RAD-seq data to infer recent shared ancestry at multiple successive SNPs (haplotypes) based on nearest-neighbor relationships<sup>31,32</sup>. Clustering results readily revealed the presence of five populations, with marked substructure suggested in some of the populations (Fig 2C). The five populations corresponded to groups of samples from single islands or island pairs in close geographic proximity, namely Santa Cruz, Isabela + Santiago, Wolf, Floreana + San Cristobal, and Pinta +



Genovesa, with further substructure within island pairs matching single islands (e.g., Isabela and Santiago, Floreana, and San Cristobal). Additionally, our results showed heterogeneous degrees of differentiation of samples between and within all populations, with some localities highly differentiated from samples at other localities (e.g., all samples from Floreana and a subgroup of samples within Santa Cruz). Notably, the highly differentiated subpopulations in Santa Cruz corresponded to the two varieties of *O. echios* present on the island. Overall, these results suggest shallow genetic divergence across populations with limited gene flow and highly restricted geographically.



**Figure 2. Despite extensive ecophenotypic variation, *Opuntia* shows relatively low genetic differentiation across the Galápagos islands, yet there is a marked geographic structure.** (A) Principal component analysis (PCA) to reduce multidimensional genomic space into fewer dimensions. The left panel

shows the space defined by PC1 (10% of the variance explained) and PC2 (9.5% of the variance explained), the right panel shows the space defined by PC3 (7.8% of the variance explained) and PC4 (5% of the variance explained). (B) Genomic ancestry shows the optimal partition of the data into two genetic demes shown across islands. (C) Fine-scale population structure based on nearest-neighbor relationships across all individuals sorted by island. Five clusters are readily detected corresponding to single islands or island pairs in close geographic proximity. Further substructure exists within some clusters. Cold colors (blue, top of color legend) correspond to closely related individuals, while warm colors (red, bottom of color legend) correspond to more distantly related individuals.

In sum, our findings show that while genetic differentiation of prickly pear cacti in the Galápagos archipelago is low, there is a marked geographic pattern of genetic structure by islands (Fig 2A, C). This is consistent with the current taxonomic hypothesis that almost every major island harbors a unique taxon, a hypothesis worthy of further investigation with increased genomic sampling in combination with explicit analysis of phenotypic variation to examine in detail the nature of *Opuntia* species in the Galápagos<sup>33,34</sup>. Conversely, it is plausible that the geographic pattern of genetic variation displayed by *Opuntia* reflects mostly intraspecific variation and the early stages of a radiation in prickly pear cacti across the Galápagos archipelago, likely fueled by phenotypic plasticity and divergent selection resulting from island-specific environments and biotic interactions. Indeed, it has been proposed that environmental variation is an important driver of phenotypic expression in cacti more broadly<sup>35</sup> and that biotic interactions may contribute specifically to explain patterns of phenotypic diversity across *Opuntia* species in the Galápagos<sup>24,27</sup>. In the future, testing this intriguing hypothesis with increasing rigor would shed light on the genomic underpinnings of phenotypic plasticity and organismal adaptation<sup>36,37</sup>.

## **Radiation across the Galápagos archipelago**

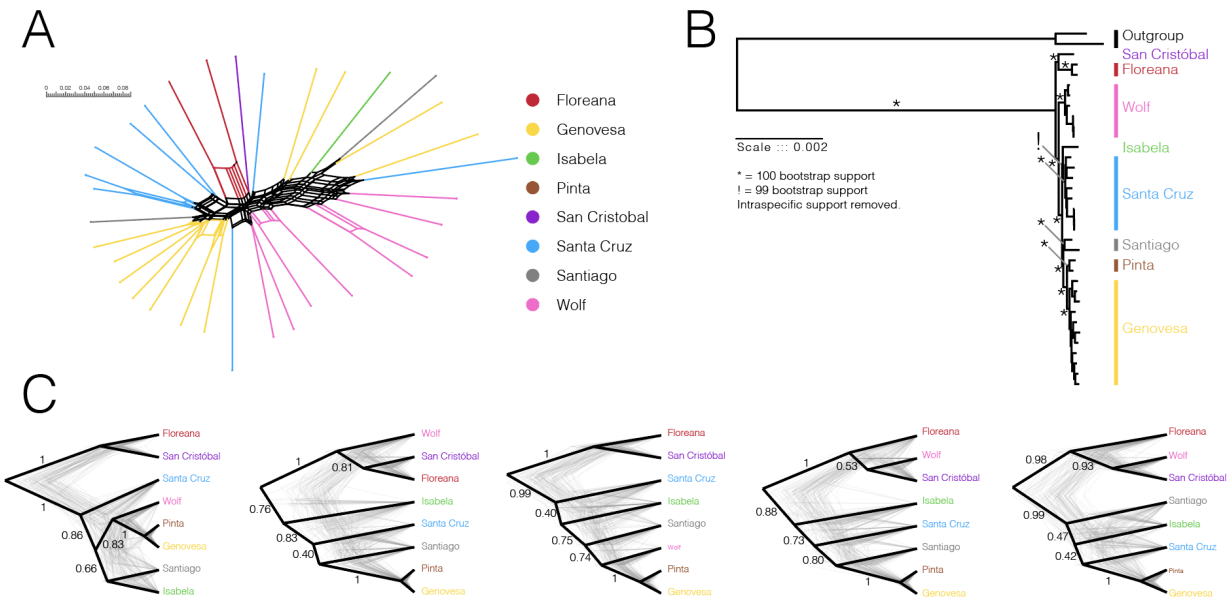


The prickly pear cacti are a prominent component of the Galápagos flora. They are widespread and occur on all major islands ( $> 10 \text{ km}^2$ ) and most islets. Although there is considerable doubt about the most closely related species on the mainland<sup>38</sup>, it is likely that *Opuntia* colonized the Galápagos only once drafting from the nearby continent via ocean currents as most other native plants of the archipelago<sup>39</sup>. Once in the archipelago, *Opuntia* diversified across islands in multiple environments, ranging from coastal dry open scrublands to higher elevation wet forests with closed canopies<sup>26</sup>. Yet, the colonization history and the biological mechanisms enabling inter-island dispersal are poorly understood.

To reconstruct the radiation of *Opuntia* throughout the Galápagos, we grouped individuals by island and used islands as units of analysis. This approach is consistent with our results on genetic structure (see above) and current taxonomy, as well as a reasonable approach to studying the colonization of organisms across archipelagos. Owing to the dynamic paleogeography of the Galápagos, including, but not limited to, volcanism linked to plate tectonics, the existence of ancient proto-islands, the subsidence and growth of extinct and extant islands, and changes in the connectivity and isolation of the current islands<sup>40</sup>, combined with the vagility of the seeds dispersers of cacti plants<sup>5,41</sup>, nearly any of the islands could be the source of colonization for other islands, and the sequence of dispersal routes could be extraordinarily complicated. Nonetheless, our results revealed two notable patterns.

First, phylogenetic reconstructions suggest a recent radiation of prickly pear cacti in the Galápagos, as seen in the tangled phylogenetic network (Fig. 3A), the short internal branches in the species tree (Fig. 3B), and the topological tree discordance using different subsets of loci (Fig. 3C). Additionally, we discovered a remarkably low number of private variants per island in our dataset

(Table 1). Together, these findings are consistent with a history of diversification influenced in part by the recent dynamic geological history of the archipelago<sup>42</sup>. Although the estimated maximum age of the oldest present Galápagos island is around 4 my (San Cristóbal), the bulk of the present islands have a maximum age of about 2.5 my<sup>21,40</sup>. This age corresponds to the beginning of the Pleistocene, a geological epoch characterized by repeated glacial-interglacial cycles that have modified dramatically the geography and connectivity of oceanic archipelagos around the globe<sup>43</sup>, including the Galápagos islands. Just within the last 3 My, the time frame of the emergent history of the present Galápagos islands, there have been approximately 36 glacial advances separated by interglacials<sup>44</sup>. Although the exact imprint of these geo-climatic events is difficult to reconstruct, Geist et al.<sup>40</sup> suggest a complex paleogeographic model for the Galápagos archipelago. They propose that only two of the nine islands present 3 Mya still exist today (San Cristóbal and Española). At 2 Mya, Santa Cruz and Floreana were a single large island and 10 other islands existed, many of which are not present today. At 1 Mya, it is likely that a massive central island was an amalgamation of multiple present islands (Santiago, Pinzón, Rábida, Santa Cruz, and Floreana), and several other islands emergent back then are submerged today. Even in more recent times, during the Last Glacial Maximum (approximately 20 Ky), there were many more islands and islets present and there were land bridges connecting three of the present big central islands (Isabela, Fernandina, and Santa Cruz). In sum, such drastic paleogeographic changes clearly affected the connectivity of populations across islands, thus possibly shaping directly the recent radiation of prickly pear cacti in the archipelago<sup>45</sup>.



**Figure 3. The radiation of *Opuntia* across the Galapagos archipelago is a recent phenomenon.** (A) The phylogenetic network shows tangled relationships among all island populations. (B) The phylogenetic tree among all island populations using a concatenated matrix of all loci shows very short internal branches subtending phylogenetic relationships. (C) The phylogenetic trees among all island populations using a coalescent approach (without concatenation) vary according to the loci included in the analysis. Each panel shows trees inferred with 5,000 random sampled loci. The clouds of trees in the background represent trees sampled from the posterior distribution, the black like represents the Maximum Credibility Tree.

Alternatively, our phylogenetic results and the distribution of private variants could be the outcome of gene flow between geographically isolated populations, perhaps facilitated by pollen and seed dispersers<sup>24</sup> and weak isolating barriers. For instance, we found that the population from San Cristóbal is sister to the population from either Floreana or Wolf, depending on the loci used to infer phylogenetic relationships (Fig. 3B, C). Such a pattern could result from differential gene flow among populations that are separated by more than 300 kilometers like in the case of San Cristobal and Wolf.

To test formally for genetic admixture between populations in our data set, we used the F4 statistic<sup>46</sup>. We detected excess allelic sharing between the populations from San Cristóbal and Wolf, Genovesa and Santa Cruz, and Wolf and the grouped populations from Genovesa and Pinta,

indicating ancient admixture between all these populations (Fig. 4). Previous studies have suggested that as mainland *Opuntia* species have weak reproductive barriers<sup>21</sup> and because the two sympatric taxa of *Opuntia* present in Santa Cruz island show little to no signal of genetic differentiation based on microsatellite data<sup>22,45</sup>, gene flow between lineages might be widespread across the Galápagos *Opuntia*. Our results are in general agreement with this idea, yet our evidence for admixture shows that gene flow is more prevalent between certain lineages than others. It remains unclear whether the instances of admixture we report here correspond to cases of hybridization or introgression and it will require more detailed analyses of patterns of reproductive isolation and ecological differentiation between lineages. Likewise, whether genetic admixture fuels or constraints differentiation is not known. Investigating the extent to which gene flow could increase the amount of genetic variation in populations to promote adaptation to local conditions will require identification of the loci underlying the traits of interest and characterization of their evolutionary histories<sup>47</sup>.

**Table 1. Private variant analysis shows low differentiation of *Opuntia* across the Galápagos.** For each island, we report the number of island-specific variants. Private variants correspond to alleles that only exist on a particular island. This analysis is based on a dataset of 52,869 total SNPs.

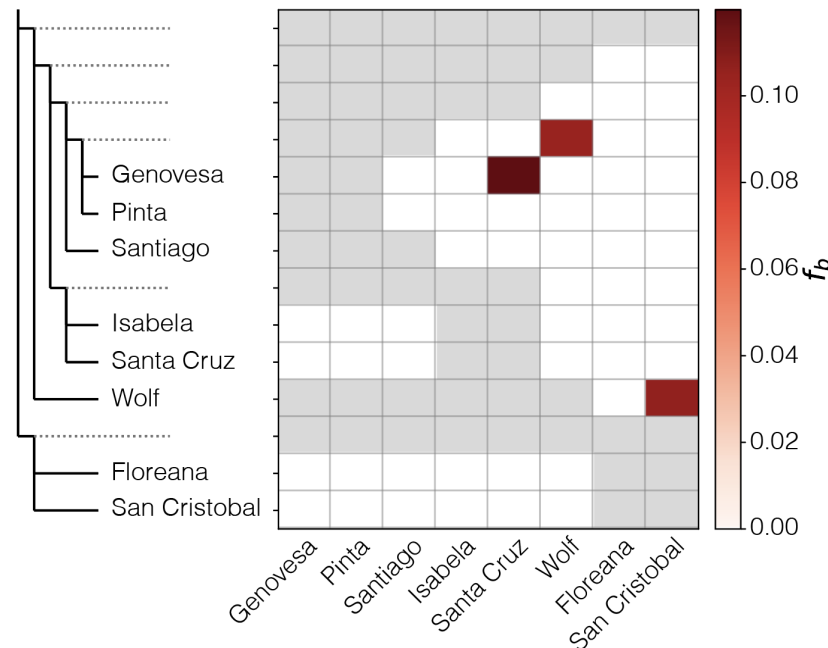
Island	Private Variants	% Private Variants
Floreana	1,805	3.4%
Genovesa	2,831	5.4%
Isabela	891	1.7%
Pinta	738	1.4%
San Cristóbal	949	1.8%
Santa Cruz	3,916	7.4%
Santiago	1,314	2.5%
Wolf	1,661	3.1%
<b>Total</b>	<b>14,105</b>	<b>26.7%</b>

The second notable phylogenetic result concerns the geographic structure of the radiation of Galápagos *Opuntia*. Both the concatenated and coalescent-based species tree analyses recovered consistently the oldest divergence of the radiation to be the separation between the populations from San Cristóbal and Floreana, and sometimes Wolf, from the remaining populations (Fig. 3B, C). Because San Cristóbal is the oldest current island (2.4 - 4.0 my) <sup>40</sup>, the deep divergence event separating the *Opuntia* population from San Cristóbal is consistent with the history of the islands. However, the age of Floreana (1.5 - 2.3 my) is on par with the ages of several other current islands included in our dataset, such as Santa Cruz (1.1 - 2.3 my) and Wolf (1.6 - 1.7), as well as with the ages of Pinzón (1.3 - 1.7 my), Rábida (1.3 - 1.6 my), and Santa Fé (2.9 - 2.9) not included in our sampling. Furthermore, Floreana was part of a larger landmass that included Santa Cruz at least 2 mya <sup>40</sup>. Therefore, although the deep split separating the *Opuntia* population from San Cristóbal and the other islands is consistent with the age of this island, it is unclear if the colonization of Floreana involved a stepping stone dispersal via other older islands such as Española, Santa Fé, or Pinzón, for which we are missing *Opuntia* samples. Including samples from such islands in combination with explicit biogeographic analyses is required to shed light on this possibility.

Sister to the San Cristóbal and Floreana (and sometimes Wolf) clade, the relationships among the remaining populations are broadly consistent with the estimates of island ages and geographic proximity (Fig. 3B, C). The populations from Genovesa and Pinta are sister to each other and are nested within a clade which includes the populations from Santiago, Santa Cruz, and Isabela. Genovesa and Pinta are some of the youngest islands in our dataset, with ages ranging between 0.3 - 0.8 my and 0.7 - 0.8 my, respectively <sup>40</sup>. Because these islands are also far apart (~80 km) from the centrally located islands (e.g., Santa Cruz, Isabela, Santiago) and have never been connected to other islands in

the archipelago in geological time, our result suggests a recent, long-distance dispersal colonization of *Opuntia* to the islands of Genovesa, Pinta, and possibly Marchena for which we are missing samples in the current study. The ages of Santa Cruz and Santiago islands range between 1.1 -2.3 my and 0.8 - 1.4 my, respectively. In contrast to other current islands included in this study, these two islands have been part of the same landmass at some point over the last million years<sup>40</sup>. Isabela is one of the youngest islands of the archipelago, with an age estimate of 0.5 - 0.8 my. Further, Isabela is in extremely close geographic proximity to Santa Cruz and Santiago, and it is likely that there was a landbridge connecting Isabela and Santa Cruz during the Last Glacial Maximum (20 ky)<sup>40</sup>. Therefore, our results are consistent with the colonization of the centrally located islands (Santiago, Santa Cruz, and Isabela) from an older island (extinct or extant), with subsequent, repeated movements of plant propagules among all the central islands, likely facilitated by the shared geology and close geographic proximity. Evidence for this scenario comes from the uncertain and incongruent relationships among the populations of these islands (Fig 3C).





**Figure 4. Genetic admixture between populations of *Opuntia* across the Galápagos.** The phylogenetic tree obtained using the concatenated matrix (see Figure 3B) is shown on the left in an “expanded” form, so that each branch, including internal branches, points to a corresponding row in the matrix with the inferred  $f_b$ -branch statistics. The values in the matrix refer to excess allele sharing between the branch identified on the expanded tree on the y- axis (relative to its sister branch) and the population on the x-axis. Warm colors (red, top of the color legend) refer to excess allele sharing.

Intriguingly, the relationship between the *Opuntia* population from Wolf and populations from all other islands is the most uncertain. Although such uncertainty is largely explained by instances of gene flow (Fig. 4), Wolf is a small (1.3 km<sup>2</sup>) and distant island (>150 km apart from other islands with the exception of Darwin island) in the northwest corner of the archipelago, and it is not known how *Opuntia* could have colonized this island. Wolf is approximately as old as Floreana (1.6 - 1.7 my; <sup>40</sup>), hence it might have been available for colonization for a long period of time, yet how *Opuntia* propagules reached this remote location in the first place remains elusive. While the direction of sea currents in the Galapagos archipelago is westward and northwestward <sup>23</sup> and this could have facilitated the dispersal of *Opuntia* to Wolf, it is more likely that animals helped disperse these plants. The

presence in Wolf Island of birds known to disperse *Opuntia* seeds elsewhere, such as Darwin's finches (*G. septentrionalis* and *G. magnirostris*: JAC pers. comm.) and mockingbirds, could support this colonization hypothesis. Lava lizards on other islands also act as seed dispersers for *Opuntia* (*O. echios* in Santa Cruz and *O. galapageia* in Pinta)<sup>48</sup>, but no lava lizards are currently found in Wolf or Darwin islands. A general trend of unidirectional gene flow and historical migration on several species within the archipelago<sup>12,49–52</sup> resulting from prevailing south–south-east trade winds in the Galápagos<sup>53–55</sup> could explain the colonization of *Opuntia* propagules or seeds dispersed by animals on these northern islands. Complementing our current sampling with populations from all other current Galápagos islands will certainly help clarify the biogeography and potential dispersal routes of *Opuntia* across the archipelago.

## A potential genomic link to adaptive phenotypes

One of the most remarkable attributes of the Galápagos prickly pear cacti radiation is the extraordinary phenotypic diversity they display in vegetative and reproductive traits (Fig. 1). Previous studies have suggested that both environmental and biotic factors may be responsible for such striking phenotypic diversity. In examining the patterns of morphological variation of *Opuntia* in three centrally located islands (Santa Fé, Santa Cruz, and Pinzón), Racine and Downhower<sup>41</sup> suggested that variation in tree height-diameter growth and pad production was largely explained by the combined effects of inter and intraspecific competition for light and wind velocity. In addition to the extensive variation observed in vegetative growth form, these authors also proposed that the pronounced variation in seed size and seed coat observed among island populations was related to seed dispersal and predation. While mockingbirds, tortoises, and iguanas disperse seeds by consuming either the fruits or

the fleshy aril around the seed, the cactus finch (*Geospiza scandens*) is a major seed predator that cracks the seed coat to extract the endosperm<sup>27</sup>. Hence, it is likely that seed predation is an important factor in the diversification of *Opuntia*, and the cactus finch is a prime agent of selection. Subsequent studies are consistent with this hypothesis and show that the interaction between Darwin's finches and prickly pear cacti extends, in fact, to multiple species of ground finches and cacti<sup>24</sup>. The intricate relationship between these two lineages is largely modulated by environmental changes such as increased and prolonged periods of drought, which affect plant survival and thus seed availability. Long-term studies of Darwin's finches populations have shown how such changes have had drastic effects on the morphology of the beaks of the finches, which are critical for seed manipulation<sup>3</sup>. Thus to some extent, it is plausible there is some degree of coevolution between cactus and *Geospiza* finches in the Galápagos mediated by flowers and seed features which requires rigorous study<sup>2,18,19,56–59</sup>

In order to detect potential genetic contributions linked to phenotypic variation in *Opuntia* across islands, we conducted parallel scans of genetic differentiation (fixation index,  $F_{ST}$ ) between pairs of populations with more than five individuals, namely Genovesa, Santa Cruz, and Wolf. We selected the 5% divergent RADseq tags and mapped these to *Arabidopsis thaliana* gene sequences using conservative cut-offs (Supplementary Table S2). Three genomic regions stood out, as these showed differentiation across all pairwise comparisons. The first region was the *CLocus\_1489*, which is associated with the MATE efflux protein family (TT12 or ATTT12). This protein family encodes a multidrug and toxic compound extrusion (MATE) transporter involved in seed coat pigmentation<sup>60</sup> and transportation of proanthocyanidins which are pigmentation proteins<sup>61,62</sup>. The second region was the *CLocus\_59059*, which is likely involved in DNA replication initiation and elongation<sup>63</sup>. The third region was detected as two loci, *CLocus\_4279* and *CLocus\_4683*, which are associated with a periderm

regulator (RIK) <sup>64</sup>. The periderm is a critical tissue protecting the vascular tissue from biotic and abiotic stresses. When considering each pairwise comparison separately, we detected differentiation in genomic regions associated with environmental stress, including temperature, drought, and salt, genomic regions associated with light responses and circadian clock, as well as genomic regions associated with plant pigmentation, among others (Supplementary Table S2).

Although significant differentiation detected via  $F_{ST}$  scans does not necessarily imply differentiation via natural selection <sup>65</sup> and further functional validation is desirable <sup>66</sup>, some results are noteworthy. Genomic differentiation linked to chemical compounds involved in seed pigmentation might be related to seed predation, supporting the hypothesis of eco-evolutionary feedback between *Opuntia* and Darwin's finches <sup>24</sup>. Alternatively, it is plausible that variable pigmentation could confer better camouflage to different soil types, a scenario congruent with the proposal of seed dormancy in *Opuntia* <sup>67</sup>. Further sampling and experimental data are necessary before these hypotheses can be confronted confidently. Genomic differentiation associated with regulation and development of the periderm as well as with multiple regulators involved in environmental stresses might be related to the harsh conditions that characterize the Galápagos islands, and possibly biotic interactions. The Galápagos islands lie directly on the equator, where sun exposure may be a major environmental stress for island dwellers. One of the Darwin daisies from the Galápagos, *Scalesia atractylloides*, shows signatures of selection along the genome associated with light reception and growth regulation <sup>7</sup>. Additionally, some *Opuntia* plants grow on harsh volcanic soils and some individuals are exposed constantly to salt spray, indicating that prickly pear cacti deal with considerable sustained abiotic stresses. Likewise, selective pressures due to biotic stress may occur from tortoises, which are known to be the main herbivores of prickly pear cacti in the Galápagos <sup>5</sup>. Taken together, our results suggest that

a combination of environmental and biotic selective pressures may have left an imprint on the cacti's genomes, an observation worthy of closer investigation to unravel the genomic underpinnings and molecular mechanisms of the *Opuntia* radiation in the Galápagos archipelago.

## Conclusions

This study presents genomic evidence of decoupling between phenotypic and genomic differentiation in the Galápagos *Opuntia*. While cactus populations from different islands show extreme variation in vegetative and reproductive traits (Fig. 1), the evidence for concomitant genomic variation is weak (Figs. 2, 3). This is consistent with the hypothesis that these cacti are at the early stages of a radiation and that gene flow between populations may facilitate local adaptation<sup>68</sup>. Alternatively, it is plausible that phenotypic plasticity precedes genetic change and is the source of adaptive divergence<sup>69</sup>. Further geographic sampling, deeper genome sequencing, and experimental data are critically needed before these hypotheses can be confronted confidently and the biological mechanisms underpinning this radiation are better understood. The observed patterns of parallel differentiation in genomic regions associated with seed pigmentation and periderm development suggest seed predation and abiotic stresses as potential mechanisms of adaptation that may be targets of selection between populations. Further study is needed to determine whether these patterns are widespread across all islands in the archipelago to understand how these genomic changes contribute to plant fitness and how selection acts during local adaptation.

# MATERIALS AND METHODS

## Taxon Sampling and DNA Extraction

Samples were collected from 8 islands across the Galapagos archipelago (Figure 2; Supplementary Table S1). The localities selected represented most of the islands across the geographic range of *Opuntia* in the Galapagos. We assigned all specimens to species following the taxonomy of Helsen et al.<sup>21</sup> (Supplementary Table S1). We collected approximately 1-2 cm<sup>2</sup> of pad tissue for each specimen, stored it in silica gel, and then transported it to the University of California, Los Angeles for further analyses. Genomic DNA extraction and purification were performed following a modified version of the CTAB extraction protocol<sup>70,71</sup> that incorporates a pre-wash step [li2007optimized] to aid in the removal of polyphenols and proteins prior to extraction. Genomic DNA quantification was performed using a Qubit fluorometer v.3.0 (Invitrogen by Thermo Fisher Scientific, Carlsbad, CA, USA) and an Agilent 2200 TapeStation (Agilent Technologies, Santa Clara, CA, USA).

## Library Preparation and Sequencing

To generate sequence data, we prepared quadruple-indexed, triple-enzyme RADseq libraries using the *EcoRI*, *XbaI*, and *NheI* restriction enzymes<sup>28</sup>. Prior to sequencing, size selection for fragment length of 375-525 bp was performed on a PippenPrep (Sage Science, Beverly, MA, USA). Libraries were pooled and sequenced in one lane of 100PE sequencing on the Illumina HiSeq4000 Sequencing Platform at the Broad Stem Cell Research Center at the University of California, Los Angeles.



## Bioinformatics and SNP discovery

We demultiplexed, removed adapters, and filtered reads for data quality using the iPyrad package v 0.9.72<sup>72</sup>. Because of the potential polyploid nature of the Galápagos *Opuntia*<sup>38</sup>, we carefully explored and processed the data. Specifically, we ran Stacks v2.60<sup>73</sup> using different M/N combinations as suggested by<sup>74</sup> to optimize rad-assembly parameters. This involved running stacks 10 times using -M (number of mismatches allowed between stacks within individuals) -N (number of mismatches allowed between stacks between individuals) values between 1 and 10. We ran this protocol four times, specifying different population maps (i.e. groups of populations), namely by islands, by species, by islands excluding outgroups, and by islands excluding outgroups. This led to an optimal combination of -M / -N of 2 for analysis involving the outgroup and -M 1 -N 1 for analysis not including the outgroup.

Because samples with a high degree of missing data tend to compromise the quality of the final RADseq variant dataset, we followed the protocol established by<sup>75</sup> to remove “bad apples”. This led to the removal of three outliers with high levels of missing data, resulting in a final dataset of 33 individuals of Galápagos *Opuntia* and two outgroup specimens. For the final dataset, we ran stacks with and without outgroups, specifying -R 0.8 (minimum percentage of individuals across populations required to process a locus).

## Population genetics and phylogeographic analyses

We performed a phylogenetic reconstruction and a Dsuite analysis using the Stacks run with outgroups. For the phylogenetic reconstruction, we retrieved a collection of FASTA rad loci (--fasta-samples). Since we had a dataset of 35 individuals (70 alleles), we kept loci with more than 59

data points (30 individuals), retaining a total of 10,484 loci (2,134,262 bp). We obtained a consensus sequence between the two loci using EMBOSS v6.6.0 consambig as done in <sup>76</sup> and concatenated the consensus loci using FASconCAT-G v1.04 (<https://github.com/PatrickKueck/FASconCAT-G/>). We used the concatenated loci to run IQ-TREE v2.1.3 <sup>77,78</sup> specifying 1,000 ultrafast bootstraps.

We ran Principal Component Analysis (PCA), STRUCTURE, a private allele analysis, fineRAD, SplitsTree, and SNAPP with the dataset without outgroups. For the PCA and STRUCTURE analysis, we obtained a vcf file where linkage was reduced by specifying only a single and random SNP per radtag (--write-random-snp), and then cleaned the vcf for coverage and for missing data as specified for the F4 analysis, obtaining a total of 23,566 SNPs. To run the PCA, we used the libraries vcfr <sup>79</sup>, adegenet <sup>80</sup>, and ggplot2 <sup>81</sup>. For STRUCTURE analysis we used plink <sup>82</sup> to convert the VCF into a structure file, and ran STRUCTURE v2.3.4 <sup>30</sup> specifying Ks between 2 and 10 (Burn-in of 100,000 and an MCMC chain of 100,000). For plotting the results of this analysis, we used the online server CLUMPAK <sup>83</sup> and its plug-in which calculates the Evanno method to determine the optimal K <sup>84</sup>. For the remaining analyses (private allele analysis, SplitsTree, and SNAPP) we used a VCF where filtering included only coverage (--maxDP 200 --max-meanDP 200 --minDP 10 --min-meanDP 10), missing data (--max-missing 0.25) and loci shared (-R 0.8). For the private alleles, we used vcfr and the function 'genetic\_diff' to calculate measures of genetic differentiation (method = nei) by population, retrieving a table with a description of heterozygotes and homozygotes, which was then parsed to determine alleles private to the different populations. We then ran fineRADstructure v0.3.2 <sup>32</sup> using default settings and annex plotting scripts (<https://www.milan-malinsky.org/fineradstructure>). For the SplitsTree analysis, we converted the vcf into a nexus file using the script vcf2phyl.py (<https://github.com/edgardomortiz/vcf2phylip>), and used SplitsTree v5 to estimate a network <sup>85-87</sup>.

For SNAPP, we subsampled the VCF to 4,000 random SNPs and ran SNAPP as included in BEAST2

88,89

For the F4 analyses, we used Dsuite<sup>90</sup>. This involved retrieving a vcf from Stacks (-R 0.8), and filtering it further by coverage (--maxDP 200 --max-meanDP 200 --minDP 10 --min-meanDP 10) and missing data (--max-missing 0.25) using vcftools<sup>91</sup>. We then ran the algorithm Fbranch of Dsuite specifying the tree obtained above and plotting it using a collection of ruby scripts from Michael Matschiner

([https://github.com/mmatschiner/tutorials/blob/master/analysis\\_of\\_introgession\\_with\\_snp\\_data/README.md](https://github.com/mmatschiner/tutorials/blob/master/analysis_of_introgession_with_snp_data/README.md)).

## Functional genetics

For the functional analysis, we ran FST with a window size of 1,000 bp and window step of 1,000 bp (independent for each rad locus) using the Weir and Cockheram estimate as implemented in VCF tools<sup>91</sup>. Using the weighted estimate, we extracted the top 5% most differentiated between Genovesa vs Wolf (446 loci at the 5% cut-off), Genovesa vs Santa Cruz (490 loci at the 5% cut-off), and Santa Cruz vs Wolf (466 loci at the 5% cut-off), and used blast 2.9.0 to compare with the *Arabidopsis thaliana* protein list

([https://www.arabidopsis.org/download\\_files/Proteins/TAIR10\\_protein\\_lists/TAIR10\\_pep\\_20101214](https://www.arabidopsis.org/download_files/Proteins/TAIR10_protein_lists/TAIR10_pep_20101214)). To do this, we created a protein database (makeblastdb -dbtype prot) and ran blastx with the top 5% outlier radtags. We then extracted e-values below 0.001 and kept only amino acids with a sequence length of 20. We searched the literature for functional information associated with each locus hit.

## Phenotypic variation

We used the most recent taxonomic monograph of Galpágos *Opuntia* to tabulate the minimum and maximum values reported for 13 phenotypic traits used to describe and delimit each of the 14 recognized taxa<sup>26</sup>. Vegetative traits included plant height, pad length, pad width, pad thickness, areoles width, distance between adjacent areoles, number of spines per areole, and spine length. Reproductive traits included fruit length, fruit width, seed length, seed width, and seed thickness. The combination of minimum and maximum values of these traits delimits a hypervolume of  $n$  dimensions ( $n$  = the number of phenotypic traits) in the phenotypic space corresponding to each species. We constructed separate hypervolumes for vegetative traits ( $n$  = 8) and reproductive traits ( $n$  = 5) because some *Opuntia* species seem to differ more in one kind of trait than the other. To determine the distinctiveness of each taxon, we log-transformed all data and estimated the pairwise asymmetric proportion of overlap of all vegetative and reproduction hypervolumes. We used the library *geometry*<sup>92</sup> in R to carry out this analysis.

## Data and Code Availability

All sequence data have been deposited to the Sequence Read Archive (NCBI): BioProject ID PRJNA1022515 (<https://www.ncbi.nlm.nih.gov/bioproject/1022515>). All processed data files to generate results are available at [https://github.com/zapata-lab/ms\\_opuntia\\_genomics](https://github.com/zapata-lab/ms_opuntia_genomics) All code to run the analyses is available at: [https://github.com/zapata-lab/ms\\_opuntia\\_genomics](https://github.com/zapata-lab/ms_opuntia_genomics)

## SUPPLEMENTAL INFORMATION

Supplemental Information includes:

- Summary of the study in Spanish.
- Supplementary Figure S1: alternative results to partition the data into demes (K=2, K=6, K= 8).
- Supplementary Table S1: taxon sampling metadata
- Supplementary Table S2: results of scans of genetic differentiation (fixation index,  $F_{ST}$ ) between pairs of populations with more than five individuals.

## ACKNOWLEDGMENTS

The authors thank the Galápagos National Park. All plant materials collected and used in this investigation were regulated by permits no. MAE-DNB-CM-2016-0041-M-0002 and PC 87-22 provided by the Galápagos National Park-Ministerio del Ambiente, Ecuador to G.R.-T and J.C. This work was partially funded by a UCLA - Faculty Research Grant to F.Z. This work used computational and storage services associated with the Hoffman2 Shared Cluster provided by UCLA Institute for Digital Research and Education's Research Technology Group, and the Norwegian national infrastructure for high-performance computing and storage. J.C. thanks Angel Rivera-Colon and Nicolas Rochette for invaluable comments on RADseq analyses of polyploids.

## AUTHOR CONTRIBUTIONS

Conceptualization, F.Z., J.C., J.E.M., J.A.C., and G.R.-T.; Investigation, F.Z., J.C., D.M., C.L.H., B.B.T., J.E.M., K.R.R., J.A.C., and G.R.-T.; Resources, F.Z., J.C., J.E.M., J.A.C., and G.R.-T.; Data Curation, F.Z., D.M., C.L.H., J.E.M., and K.R.R.; Writing – Original Draft, F.Z. and J.C.; Writing – Review and Editing, all authors; Visualization, J.C.; Supervision and Project Administration, F.Z. and G.R.-T.; Funding Acquisition, F.Z.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

## REFERENCES

1. Darwin, C. (1859). On the origin of species by means of natural selection, or, The preservation of favoured races in the struggle for life /. 10.5962/bhl.title.68064.
2. Lack, D. (1945). The Galapagos Finches: (Geospizinae): a study in variation (Calif. Acad. of Sciences).
3. Grant, P.R., and Grant, R.B. (2014). 40 Years of Evolution: Darwin's Finches on Daphne Major Island (Princeton University Press).
4. Enbody, E.D., Sendell-Price, A.T., Sprehn, C.G., Rubin, C.-J., Visscher, P.M., Grant, B.R., Grant, P.R., and Andersson, L. (2023). Community-wide genome sequencing reveals 30 years of Darwin's finch evolution. *Science* 381, eadf6218. 10.1126/science.adf6218.
5. Stöcklin, J. (2009). Darwin and the Plants of the Galápagos-Islands. *Bauhinia*, 33–48.
6. Fernández-Mazuecos, M., Vargas, P., McCauley, R.A., Monjas, D., Otero, A., Chaves, J.A., Andino, J.E.G., and Rivas-Torres, G. (2020). The Radiation of Darwin's Giant Daisies in the Galápagos Islands. *Curr. Biol.* 30, 1–18. 10.1016/j.cub.2020.09.019.
7. Cerca, J., Petersen, B., Lazaro-Guevara, J.M., Rivera-Colón, A., Birkeland, S., Vizueta, J., Li, S., Li, Q., Loureiro, J., Kosawang, C., et al. (2022). The genomic basis of the plant island syndrome in Darwin's giant daisies. *Nat. Commun.* 13, 3729. 10.1038/s41467-022-31280-w.
8. Wallace, A.R. (1869). The Malay Archipelago: The Land of the Orang-utan and the Bird of Paradise; a Narrative of Travel, with Studies of Man and Nature (Courier Corporation).
9. Macarthur, R.H., and Wilson, E.O. (1967). The Theory of Island Biogeography REV - Revised. (Princeton University Press).
10. Cerca, J., Cotoras, D.D., Bieker, V.C., De-Kayne, R., Vargas, P., Fernández-Mazuecos, M., López-Delgado, J., White, O., Stervander, M., Geneva, A.J., et al. (2023). Evolutionary genomics of oceanic island radiations. *Trends Ecol. Evol.* 10.1016/j.tree.2023.02.003.
11. Miller, J.M., Quinzin, M.C., Edwards, D.L., Eaton, D.A.R., Jensen, E.L., Russello, M.A., Gibbs, J.P., Tapia, W., Rueda, D., and Caccone, A. (2018). Genome-Wide Assessment of Diversity and



- Divergence Among Extant Galapagos Giant Tortoise Species. *J. Hered.* *109*, 611–619.
12. Caccone, A., Gentile, G., Gibbs, J.P., Fritts, T.H., Snell, H.L., Betts, J., and Powell, J.R. (2002). Phylogeography and history of giant Galápagos tortoises. *Evolution* *56*, 2052–2066. 10.1111/j.0014-3820.2002.tb00131.x.
  13. Rassmann, K. (1997). Evolutionary age of the Galápagos iguanas predates the age of the present Galápagos islands. *Mol. Phylogenet. Evol.* *7*, 158–172. 10.1006/mpev.1996.0386.
  14. MacLeod, A., Rodríguez, A., Vences, M., Orozco-terWengel, P., García, C., Trillmich, F., Gentile, G., Caccone, A., Quezada, G., and Steinfartz, S. (2015). Hybridization masks speciation in the evolutionary history of the Galápagos marine iguana. *Proceedings of the Royal Society B: Biological Sciences* *282*, 20150425–20150429.
  15. Phillips, J.G., Linscott, T.M., Rankin, A.M., Kraemer, A.C., Shoobs, N.F., and Parent, C.E. (2020). Archipelago-Wide Patterns of Colonization and Speciation Among an Endemic Radiation of Galápagos Land Snails. *J. Hered.* *111*, 92–102. 10.1093/jhered/esz068.
  16. Grant, P.R., and Grant, B.R. (2007). *How and why species multiply* (Princeton University Press).
  17. Lamichhaney, S., Berglund, J., Almén, M.S., Maqbool, K., Grabherr, M., Martinez-Barrio, A., Promerová, M., Rubin, C.-J., Wang, C., Zamani, N., et al. (2015). Evolution of Darwin’s finches and their beaks revealed by genome sequencing. *Nature*, 1–16. 10.1038/nature14181.
  18. Lack, D. (1947). *Darwin’s Finches* (Cambridge University Press).
  19. Bowman, R.I. (1963). Evolutionary patterns in Darwin’s finches. *Occas. Pap. Calif. Acad. Sci.* *44*, 107–140.
  20. Browne, R.A., Anderson, D.J., White, M.D., and Johnson, M.A. (2003). Evidence for low genetic divergence among Galápagos *Opuntia* cactus species. *Notícias de Galápagos*.
  21. Helsen, P., Browne, R.A., Anderson, D.J., Verdyck, P., and Van Dongen, S. (2009). Galápagos’ *Opuntia* (prickly pear) cacti: extensive morphological diversity, low genetic variability. *Biol. J. Linn. Soc. Lond.* *96*, 451–461. 10.1111/j.1095-8312.2008.01141.x.
  22. Helsen, P., Verdyck, P., Tye, A., and Van Dongen, S. (2009). Low levels of genetic differentiation between *Opuntia echios* varieties on Santa Cruz (Galapagos). *Osterr. bot. Z.* *279*, 1–10. 10.1007/s00606-008-0064-5.
  23. Arjona, Y., Fernández-López, J., Navascués, M., Alvarez, N., Nogales, M., and Vargas, P. (2020). Linking seascape with landscape genetics: Oceanic currents favour colonization across the Galápagos Islands by a coastal plant. *J. Biogeogr.* *47*, 2622–2633. 10.1111/jbi.13967.
  24. Grant, B.R., and Grant, P.R. (1981). Exploitation of *Opuntia* cactus by birds on the Galápagos. *Oecologia* *49*, 179–187. 10.1007/BF00349186.

25. Olesen, J.M., Damgaard, C.F., Fuster, F., Heleno, R.H., Nogales, M., Rumeu, B., Trøjelsgaard, K., Vargas, P., and Traveset, A. (2018). Disclosing the double mutualist role of birds on Galápagos. *Sci. Rep.* 8, 57. 10.1038/s41598-017-17592-8.
26. Anderson, E.F., and Walkington, D.L. (1971). Cactaceae. In *Flora of the Galápagos islands*, I. L. Wiggins and D. M. Porter, eds. (Stanford University Press), pp. 533–546.
27. Racine, C.H., and Downhower, J.F. (1974). Vegetative and Reproductive Strategies of *Opuntia* (Cactaceae) in the Galapagos Islands'. *Biotropica* 6, 175–186. 10.2307/2989650.
28. Bayona-Vásquez, N.J., Glenn, T.C., Kieran, T.J., Pierson, T.W., Hoffberg, S.L., Scott, P.A., Bentley, K.E., Finger, J.W., Louha, S., Troendle, N., et al. (2019). Adapterama III: Quadruple-indexed, double/triple-enzyme RADseq libraries (2RAD/3RAD). *PeerJ* 7, e7724–e7725.
29. Patterson, N., Price, A.L., and Reich, D. (2006). Population Structure and Eigenanalysis. *PLoS Genet.* 2, e190–20.
30. Pritchard, J.K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.
31. Lawson, D.J., Hellenthal, G., Myers, S., and Falush, D. (2012). Inference of Population Structure using Dense Haplotype Data. *PLoS Genet.* 8, e1002453–16.
32. Malinsky, M., Trucchi, E., Lawson, D.J., and Falush, D. (2018). RADpainter and fineRADstructure: Population Inference from RADseq Data. *Mol. Biol. Evol.* 35, 1284–1290.
33. Cadena, C.D., Zapata, F., and Jiménez, I. (2018). Issues and Perspectives in Species Delimitation using Phenotypic Data: Atlantean Evolution in Darwin's Finches. *Syst. Biol.* 67, 181–194.
34. Cadena, C.D., and Zapata, F. (2021). The genomic revolution and species delimitation in birds (and other organisms): Why phenotypes should not be overlooked. *Ornithology* 138, 1–18.
35. Arthur, G., and Nobel, P. (1986). *The Cactus Primer*.
36. Berner, D., and Salzburger, W. (2015). The genomics of organismal diversification illuminated by adaptive radiations. *Trends Genet.* 31, 491–499. 10.1016/j.tig.2015.07.002.
37. Parsons, K.J., Concannon, M., Navon, D., Wang, J., Ea, I., Groveas, K., Campbell, C., and Albertson, R.C. (2016). Foraging environment determines the genetic architecture and evolutionary potential of trophic morphology in cichlid fishes. *Mol. Ecol.* 25, 6012–6023. 10.1111/mec.13801.
38. Majure, L.C., Puente, R., Griffith, M.P., Judd, W.S., Soltis, P.S., and Soltis, D.E. (2012). Phylogeny of *Opuntias*.s. (Cactaceae): Clade delineation, geographic origins, and reticulate evolution. *Am. J. Bot.* 99, 847–864.

39. Fajardo, J., Mateo, R.G., Vargas, P., Fernandez-Alonso, J.L., Gómez Rubio, V., Felicísimo, Á.M., and Muñoz, J. (2019). The role of abiotic mechanisms of long-distance dispersal in the American origin of the Galápagos flora. *Glob. Ecol. Biogeogr.* 28, 1610–1620.
40. Geist, D.J., Snell, H., Snell, H., Goddard, C., and Kurz, M.D. (2014). A paleogeographic model of the Galápagos islands and biogeographical and evolutionary implications. In *The Galapagos A Natural Laboratory for the Earth Sciences*, Geophysical Monograph, K. S. Harpp, E. Mittelstaedt, N. d'Ozouville, and D. W. Graham, eds. (John Wiley & Sons, Inc), pp. 145–166. 10.1002/9781118852538.ch8.
41. Racine, C.H., and Downhower, J.F. (1974). Vegetative and Reproductive Strategies of *Opuntia* (Cactaceae) in the Galapagos Islands. *Biotropica* 6, 175–186. 10.2307/2989650.
42. Whittaker, R.J., School of Geography Robert J Whittaker, and Fernandez-Palacios, J.M. (2007). *Island Biogeography: Ecology, Evolution, and Conservation* (OUP Oxford).
43. Spratt, R.M., and Lisiecki, L.E. (2016). A Late Pleistocene sea level stack. *Clim. Past* 12, 1079–1092. 10.5194/cp-12-1079-2016.
44. Huybers, P. (2007). Glacial variability over the last two million years: an extended depth-derived age model, continuous obliquity pacing, and the Pleistocene progression. *Quat. Sci. Rev.* 26, 37–55. 10.1016/j.quascirev.2006.07.013.
45. Helsen, P., Verdyck, P., and Van Dongen, S. (2011). The influence of historical gene flow, bathymetry and distribution patterns on the population genetics of morphologically diverse Galápagos' *Opuntia echios*. *J. Mol. Evol.* 72, 315–325. 10.1007/s00239-011-9434-7.
46. Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., and Reich, D. (2012). Ancient admixture in human history. *Genetics* 192, 1065–1093. 10.1534/genetics.112.145037.
47. Todesco, M., Owens, G.L., Bercovich, N., Légaré, J.-S., Soudi, S., Burge, D.O., Huang, K., Ostevik, K.L., Drummond, E.B.M., Imerovski, I., et al. (2020). Massive haplotypes underlie ecotypic differentiation in sunflowers. *Nature* 584, 602–607. 10.1038/s41586-020-2467-6.
48. Hervías-Parejo, S., Heleno, R., Rumeu, B., Guzmán, B., Vargas, P., Olesen, J.M., Traveset, A., Vera, C., Benavides, E., and Nogales, M. (2019). Small size does not restrain frugivory and seed dispersal across the evolutionary radiation of Galápagos lava lizards. *Curr. Zool.* 65, 353–361. 10.1093/cz/zoy066.
49. Arbogast, B.S., Drovetski, S.V., Curry, R.L., Boag, P.T., Seutin, G., Grant, P.R., Grant, B.R., and Anderson, D.J. (2006). The origin and diversification of Galapagos mockingbirds. *Evolution* 60, 370–382. 10.1111/j.0014-3820.2006.tb01113.x.
50. Caccone, A., Gibbs, J.P., Ketmaier, V., Suatoni, E., and Powell, J.R. (1999). Origin and evolutionary relationships of giant Galápagos tortoises. *Proc. Natl. Acad. Sci. U. S. A.* 96,

- 13223–13228. 10.1073/pnas.96.23.13223.
51. Benavides, E., Baum, R., Snell, H.M., Snell, H.L., and Sites, J.W., Jr (2009). Island biogeography of Galápagos lava lizards (Tropiduridae: Microlophus): species diversity and colonization of the archipelago. *Evolution* 63, 1606–1626. 10.1111/j.1558-5646.2009.00617.x.
52. Chaves, J.A., Parker, P.G., and Smith, T.B. (2012). Origin and population history of a recent colonizer, the yellow warbler in Galápagos and Cocos Islands. *J. Evol. Biol.* 25, 509–521. 10.1111/j.1420-9101.2011.02447.x.
53. Power, D.M. (1975). Similarity among avifaunas of the Galapagos islands. *Ecology* 56, 616–626. 10.2307/1935495.
54. Geist, D.J. (1992). Pyroclastic rocks as paleoclimate indicators, Galapagos Islands. In Abstracts and Program, Chapman Conference on Climate, Volcanism and Global Change, R. P. Turco, ed. (American Geophysical Union, Washington, DC.), p. 18.
55. Colinvaux, P.A. (1984). The Galápagos climate: past and present. In Galápagos, R. Perry, ed. (Pergamon. Oxford, U.K.), pp. 55–69.
56. Grant, P.R., and Grant, B.R. (1980). The breeding and feeding characteristics of Darwin’s finches on Isla genovesa, Galapagos. *Ecol. Monogr.* 50, 381–410. 10.2307/2937257.
57. Bowman, R.I. (1961). Morphological differentiation and adaptation in the Galapagos finches. *Univ. Calif. Publ. Zool.* 58, 1–302.
58. Abbott, I., Abbott, L.K., and Grant, P.R. (1977). Comparative Ecology of Galapagos Ground Finches (Geospiza Gould): Evaluation of the Importance of Floristic Diversity and Interspecific Competition. *Ecol. Monogr.* 47, 151–184.
59. Smith, J.N.M., Grant, P.R., Grant, B.R., Abbott, I.J., and Abbott, L.K. (1978). Seasonal variation in feeding habits of Darwin’s ground finches. *Ecology* 59, 1137–1150. 10.2307/1938228.
60. Chai, Y.-R., Lei, B., Huang, H.-L., Li, J.-N., Yin, J.-M., Tang, Z.-L., Wang, R., and Chen, L. (2009). TRANSPARENT TESTA 12 genes from Brassica napus and parental species: cloning, evolution, and differential involvement in yellow seed trait. *Mol. Genet. Genomics* 281, 109–123. 10.1007/s00438-008-0399-1.
61. Frank, S., Keck, M., Sagasser, M., Niehaus, K., Weisshaar, B., and Stracke, R. (2011). Two differentially expressed MATE factor genes from apple complement the Arabidopsis transparent testa12 mutant. *Plant Biol.* 13, 42–50. 10.1111/j.1438-8677.2010.00350.x.
62. Gao, J.-S., Wu, N., Shen, Z.-L., Lv, K., Qian, S.-H., Guo, N., Sun, X., Cai, Y.-P., and Lin, Y. (2016). Molecular cloning, expression analysis and subcellular localization of a Transparent Testa 12 ortholog in brown cotton (Gossypium hirsutum L.). *Gene* 576, 763–769.

10.1016/j.gene.2015.11.002.

63. Rizvi, I., Choudhury, N.R., and Tuteja, N. (2016). Arabidopsis thaliana MCM3 single subunit of MCM2-7 complex functions as 3' to 5' DNA helicase. *Protoplasma* 253, 467–475. 10.1007/s00709-015-0825-2.
64. Boher, P., Soler, M., Fernández-Piñán, S., Torrent, X., Müller, S.Y., Kelly, K.A., Serra, O., and Figueras, M. (2021). Silencing of StRIK in potato suggests a role in periderm related to RNA processing and stress. *BMC Plant Biology* 21. 10.1186/s12870-021-03141-z.
65. Ravinet, M., Faria, R., Butlin, R.K., Galindo, J., Bierne, N., Rafajlović, M., Noor, M.A.F., Mehlig, B., and Westram, A.M. (2017). Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. *J. Evol. Biol.* 30, 1450–1477. 10.1111/jeb.13047.
66. Gudmunds, E., Wheat, C.W., Khila, A., and Husby, A. (2022). Functional genomic tools for emerging model species. *Trends Ecol. Evol.* 10.1016/j.tree.2022.07.004.
67. Jackson, M.H. (1993). *Galapagos, a Natural History* (University of Calgary Press).
68. Suarez-Gonzalez, A., Lexer, C., and Cronk, Q.C.B. (2018). Adaptive introgression: a plant perspective. *Biol. Lett.* 14. 10.1098/rsbl.2017.0688.
69. West-Eberhard, M.J. (1989). Phenotypic plasticity and the origins of diversity. *Annu. Rev. Ecol. Syst.* 20, 249–278. 10.1146/annurev.es.20.110189.001341.
70. Doyle, J.J., and Doyle, J.L. (1987). A rapid DNA isolation procedure from small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19, 11–15.
71. Cullings, K.W. (1992). Design and testing of a plant-specific PCR primer for ecological and evolutionary studies. *Mol. Ecol.* 1, 233–240. 10.1111/j.1365-294x.1992.tb00182.x.
72. Eaton, D.A.R., and Overcast, I. (2020). ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics* 36, 2592–2594.
73. Rochette, N.C., Rivera-Colón, A.G., and Catchen, J.M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Mol. Ecol.* 28, 4737–4754. 10.1111/mec.15253.
74. Paris, J.R., Stevens, J.R., and Catchen, J.M. (2017). Lost in parameter space: a road map for stacks. *Methods Ecol. Evol.* 8, 1360–1373. 10.1111/2041-210x.12775.
75. Cerca, J., Maurstad, M.F., and Rochette, N.C. (2021). Removing the bad apples: A simple bioinformatic method to improve loci-recovery in de novo RADseq data for non-model organisms. *Methods Ecol. Evol.*
76. Cerca, J., Rivera-Colón, A.G., Ferreira, M.S., Ravinet, M., Nowak, M.D., Catchen, J.M., and Struck, T.H. (2021). Incomplete lineage sorting and ancient admixture, and speciation without

- morphological change in ghost-worm cryptic species. *PeerJ* 9, e10896. 10.7717/peerj.10896.
77. Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., and Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* 37, 1530–1534. 10.1093/molbev/msaa015.
  78. Minh, B.Q., Hahn, M.W., and Lanfear, R. (2020). New Methods to Calculate Concordance Factors for Phylogenomic Datasets. *Mol. Biol. Evol.* 37, 2727–2733. 10.1093/molbev/msaa106.
  79. Knaus, B.J., and Grünwald, N.J. (2017). vcfr: a package to manipulate and visualize variant call format data in R. *Mol. Ecol. Resour.* 17, 44–53. 10.1111/1755-0998.12549.
  80. Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403–1405. 10.1093/bioinformatics/btn129.
  81. Wickham, H. (2011). Ggplot2. *Wiley Interdiscip. Rev. Comput. Stat.* 3, 180–185. 10.1002/wics.147.
  82. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. 10.1086/519795.
  83. Kopelman, N.M., Mayzel, J., Jakobsson, M., Rosenberg, N.A., and Mayrose, I. (2015). Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.* 15, 1179–1191. 10.1111/1755-0998.12387.
  84. Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14, 2611–2620. 10.1111/j.1365-294X.2005.02553.x.
  85. Bryant, D., and Moulton, V. (2004). Neighbor-net: an agglomerative method for the construction of phylogenetic networks. *Mol. Biol. Evol.* 21, 255–265. 10.1093/molbev/msh018.
  86. Hamming, R.W. (1950). Error detecting and error correcting codes. *The Bell System Technical Journal* 29, 147–160. 10.1002/j.1538-7305.1950.tb00463.x.
  87. Dress, A.W.M., and Huson, D.H. (2004). Constructing splits graphs. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 1, 109–115. 10.1109/TCBB.2004.27.
  88. Mccann, J., Jang, T.-S., Macas, J., Schneeweiss, G.M., Matzke, N.J., Novák, P., Stuessy, T.F., Villaseñor, J.L., and Weiss-Schneeweiss, H. (2018). Dating the Species Network: Allopolyploidy and Repetitive DNA Evolution in American Daisies (*Melampodium* sect. *Melampodium*, Asteraceae). *Syst. Biol.* 67, 1010–1024. 10.1093/sysbio/syy024.
  89. Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N.A., and RoyChoudhury, A. (2012). Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full



coalescent analysis. *Mol. Biol. Evol.* 29, 1917–1932. 10.1093/molbev/mss086.

90. Malinsky, M., Matschiner, M., and Svardal, H. (2021). Dsuite - Fast D-statistics and related admixture evidence from VCF files. *Mol. Ecol. Resour.* 21, 584–595. 10.1111/1755-0998.13265.
91. 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., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. 10.1093/bioinformatics/btr330.
92. Roussel, J.R., Barber, C.B., Habel, K., and Grasman, R. geometry: Mesh generation and surface tessellation. R package version 0.4.