



Year: 2024

Pleistocene diversification of unifoliolate-leaved *Lupinus* (Leguminosae: Papilionoideae) in Florida

Nevado, Bruno ; Atchison, Guy W ; Bridges, Edwin L ; Orzell, Steve ; Filatov, Dmitry ; Hughes, Colin E

Abstract: The importance and prevalence of recent ice-age and post-glacial speciation and species diversification during the Pleistocene across many organismal groups and physiographic settings are well established. However, the extent to which Pleistocene diversification can be attributed to climatic oscillations and their effects on distribution ranges and population structure remains debatable. In this study, we use morphologic, geographic and genetic (RADseq) data to document Pleistocene speciation and intra-specific diversification of the unifoliolate-leaved clade of Florida *Lupinus*, a small group of species largely restricted to inland and coastal sand ridges across the Florida peninsula and panhandle. Phylogenetic and demographic analyses alongside morphological and geographic evidence suggest that recent speciation and intra-specific divergence within this clade were driven by a combination of non-adaptive allopatric divergence caused by edaphic niche conservatism and opportunities presented by the emergence of new post-glacial sand ridge habitats. These results highlight the central importance of even modest geographic isolation and short periods of allopatric divergence following range expansion in the emergence of new taxa and add to the growing evidence that Pleistocene climatic oscillations may contribute to rapid diversification in a myriad of physiographic settings. Furthermore, our results shed new light on long-standing taxonomic debate surrounding the number of species in the Florida unifoliolate *Lupinus* clade providing support for recognition of five species and a set of intra-specific variants. The important conservation implications for the narrowly restricted, highly endangered species *Lupinus aridorum*, which we show to be genetically distinct from its sister species *Lupinus westianus*, are discussed.

DOI: <https://doi.org/10.1111/mec.17232>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-252941>

Journal Article

Published Version





The following work is licensed under a Creative Commons: Attribution 4.0 International (CC BY 4.0) License.

Originally published at:

Nevado, Bruno; Atchison, Guy W; Bridges, Edwin L; Orzell, Steve; Filatov, Dmitry; Hughes, Colin E (2024). Pleistocene diversification of unifoliolate-leaved *Lupinus* (Leguminosae: Papilionoideae) in Florida. *Molecular Ecology*, 33(3):e17232.

DOI: <https://doi.org/10.1111/mec.17232>

Pleistocene diversification of unifoliolate-leaved *Lupinus* (Leguminosae: Papilionoideae) in Florida

Bruno Nevado^{1,2}  | Guy W. Atchison³ | Edwin L. Bridges⁴ | Steve Orzell⁵  | Dmitry Filatov⁶ | Colin E. Hughes³

¹Faculty of Sciences, cE3c – Centre for Ecology, Evolution and Environmental Changes & CHANGE – Global Change and Sustainability Institute, University of Lisbon, Lisbon, Portugal

²Department of Animal Biology, Faculty of Sciences, University of Lisbon, Lisbon, Portugal

³Department of Systematic and Evolutionary Botany, University of Zurich, Zurich, Switzerland

⁴Botanical and Ecological Consultant, Gig Harbor, Washington, USA

⁵Avon Park Air Force Range, Avon Park, Florida, USA

⁶Department of Biology, University of Oxford, Oxford, UK

Correspondence

Bruno Nevado, Faculty of Sciences, cE3c – Centre for Ecology, Evolution and Environmental Changes & CHANGE – Global Change and Sustainability Institute, University of Lisbon, 1749-016 Lisbon, Portugal.

Email: bnevado@fc.ul.pt

Colin E. Hughes, Department of Systematic and Evolutionary Botany, University of Zurich, Zollikerstrasse 107, 8008 Zurich, Switzerland.

Email: colin.hughes@systbot.uzh.ch

Funding information

Fundação para a Ciência e a Tecnologia, Grant/Award Number: 2022.15825.CPCA, CEECIND/00229/2018 and PTDC/BIA-EVL/2398/2021; Natural Environment Research Council, Grant/Award Number: NE/K004352/1; Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: 31003A_135522 and 31003A_182453/1

Handling Editor: Richard J. Abbott

Abstract

The importance and prevalence of recent ice-age and post-glacial speciation and species diversification during the Pleistocene across many organismal groups and physiographic settings are well established. However, the extent to which Pleistocene diversification can be attributed to climatic oscillations and their effects on distribution ranges and population structure remains debatable. In this study, we use morphologic, geographic and genetic (RADseq) data to document Pleistocene speciation and intra-specific diversification of the unifoliolate-leaved clade of Florida *Lupinus*, a small group of species largely restricted to inland and coastal sand ridges across the Florida peninsula and panhandle. Phylogenetic and demographic analyses alongside morphological and geographic evidence suggest that recent speciation and intra-specific divergence within this clade were driven by a combination of non-adaptive allopatric divergence caused by edaphic niche conservatism and opportunities presented by the emergence of new post-glacial sand ridge habitats. These results highlight the central importance of even modest geographic isolation and short periods of allopatric divergence following range expansion in the emergence of new taxa and add to the growing evidence that Pleistocene climatic oscillations may contribute to rapid diversification in a myriad of physiographic settings. Furthermore, our results shed new light on long-standing taxonomic debate surrounding the number of species in the Florida unifoliolate *Lupinus* clade providing support for recognition of five species and a set of intra-specific variants. The important conservation implications for the narrowly restricted, highly endangered species *Lupinus aridorum*, which we show to be genetically distinct from its sister species *Lupinus westianus*, are discussed.

KEYWORDS

demographic modelling, Florida, *Lupinus* spp., phylogenomics, pleistocene climatic oscillations, RADseq, speciation

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

© 2024 The Authors. *Molecular Ecology* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

There is abundant evidence for recent ice-age or post-glacial speciation during the Pleistocene across diverse organismal groups and settings (Hewitt, 2000; Kadereit & Abbott, 2021). These include freshwater fish in post-glacial lakes (Hudson et al., 2011) and on Sundaland island archipelagos (Sholihah et al., 2021), birds in boreal North America (Weir & Schluter, 2004) and temperate coastal regions of New Zealand (Weir et al., 2016), grasshoppers in western North American montane sky islands in the Rockies (Knowles, 2000), mangroves in south-east Asia (He et al., 2019), the radiation of annual plants of *Nigella* across the Aegean archipelago in the Mediterranean (Comes et al., 2008), the rapid diversification of a section of *Trigonostemon* in the Malay Peninsula and Borneo (Yu & Van Welzen, 2020) and plants in the high elevation Andes (Nevado et al., 2018) and the high Arctic (Brochmann et al., 2004). While a role for Late Pleistocene glacial cycles in species diversification has long been suggested (Haffer, 1969; Simpson, 1974), the impacts of these cycles in different geological settings are likely to be variable and location specific and remain poorly understood (Haffer, 1969; Hewitt, 1996; Klicka & Zink, 1997; Papadopoulou & Knowles, 2015a, 2015b; Rull, 2011; Weir et al., 2016).

Many of the examples of speciation coinciding with the Pleistocene are from physiographic settings where climatic oscillations strongly affected habitat connectivity. Typical examples include island archipelagos, where fluctuating sea levels associated with glacial cycles caused intermittent connectivity and fragmentation between islands (e.g. Comes et al., 2008); adjacent ocean basins that were repeatedly isolated and re-connected (e.g. Filatov et al., 2021; He et al., 2019); high elevation montane 'sky-island' systems where the same glacial cycles caused shifts in elevation limits of vegetation zones and species ranges, leading to so-called flickering connectivity of high elevation habitats (Flantua et al., 2019) and continental lakes where glacial cycles caused changes in water levels that similarly affected patterns of suitable habitat connectivity (e.g. Nevado et al., 2013; Sturmbauer, 1998). These fluctuations in habitat connectivity can drive repeated cycles of geographic isolation and secondary contact between populations, a mechanism sometimes referred to as a 'species-pump', given its putative effect on speciation rates. Recent work has indeed suggested that a model of mixing-isolation-mixing (MIM) driven by such cycles of flickering connectivity, isolation and gene flow could propel rapid (exponential) speciation (He et al., 2019), and potentially even account for the exceptional hotspots of species diversity that are located in areas where such Pleistocene flickering connectivity has been especially prevalent.

It is less clear what effects Pleistocene climatic oscillations had in regions that lack the accentuated island or mountain range topography that generates cycles of flickering connectivity. In flatter areas with less accentuated topography, episodes of isolation and gene flow could still result from cycles of range expansion and migration associated the emergence of land when sea level was low and subsequent contraction as land became covered with water at sea-level maxima, or from cycles of north-south advance and retreat to glacial

refugia (Hewitt, 1996). However, whether climatic oscillations in these systems could still drive diversification remains unclear.

We examine these questions about recent Pleistocene diversification by analysing divergence and incipient speciation and quantifying historical gene flow among the species and morphological variants that make up the unifoliolate-leaved clade of peninsular Florida *Lupinus*. This clade has been taken to comprise between three and five species (Beckner, 1982; Dunn, 1971; Isely, 1986; Sholars & Riggins, 2023), but the status of most of these has been questioned at one time or another (see Appendix S1 for an account of the taxonomic history and conservation status of these taxa), and there is no current consensus about how many species should be recognized. This is especially relevant because several of the Florida taxa are rare and threatened (Appendix S1). For example, doubt about the status of *Lupinus aridorum* which has variously been treated as a distinct species, or as a variety of *Lupinus westianus*, has detracted attention from its endangered red-listing status at both a state and federal levels (Bibb et al., 2007; Contu, 2012; U.S. Fish and Wildlife Service, 1987) and confounded conservation assessments by different federal and state authorities. These conflicting conservation assessments highlight the need to revisit the taxonomy of these species with a more robust and rigorous evidence-based approach. Furthermore, intensive field collecting over the last few years has revealed evidence for additional geographically structured morphological variation among the unifoliolate lupines across peninsular Florida (Figure 1a,b) raising the possibility of recognizing one or more additional taxa.

The species of the Florida unifoliolate-leaved clade occur on xeric sands across northern and peninsular Florida, in fire-prone sandhill and Florida scrub (i.e. sand pine scrub, low oak scrub, rosemary scrub and scrubby pinelands), often dominated by re-sprouting xeric oaks, decumbent palms and clonal ericaceous shrubs (Menges & Hawkes, 1998) with a strong predilection for exposed, sandy, xerophytic soils. Two species in the clade, *Lupinus diffusus* and *Lupinus villosus*, extend north into Alabama, Georgia and North and South Carolina (Figure 1; Dunn, 1971; Isely, 1986, 1998). These habitats generally occur on fragmented and somewhat isolated areas of predominately xeric uplands and sand ridges forming a series of continental edaphic islands (Schenk et al., 2018) which harbour notable concentrations of vascular plant endemism (Christman & Judd, 1990; Estill & Cruzan, 2001; Menges et al., 2007), forming a regionally distinct local biodiversity hotspot with a unique and nationally important biota. In addition to these inland sand ridges, scrub habitat is also found along the very recent post-glacial, late-Pleistocene shoreline sand dune systems (Hine, 2013). While these endemic-rich inland sand ridges are older than the coastal sand dune systems and habitats, all of these formations are fundamentally recent, reflecting the very mobile shorelines and the dramatic impacts of late Miocene, Pliocene and especially Pleistocene sea-level fluctuations across the otherwise low-lying topography of the Florida peninsula and its extensive adjacent shallow continental shelf (Figure 1a; Hine, 2013; Locker et al., 1996; see Germain-Aubrey et al., 2014: fig. 1; Krysko et al., 2016: fig. 3). Through the Pliocene, sea level high stands of 20+ m above today's level submerged substantial parts of Florida

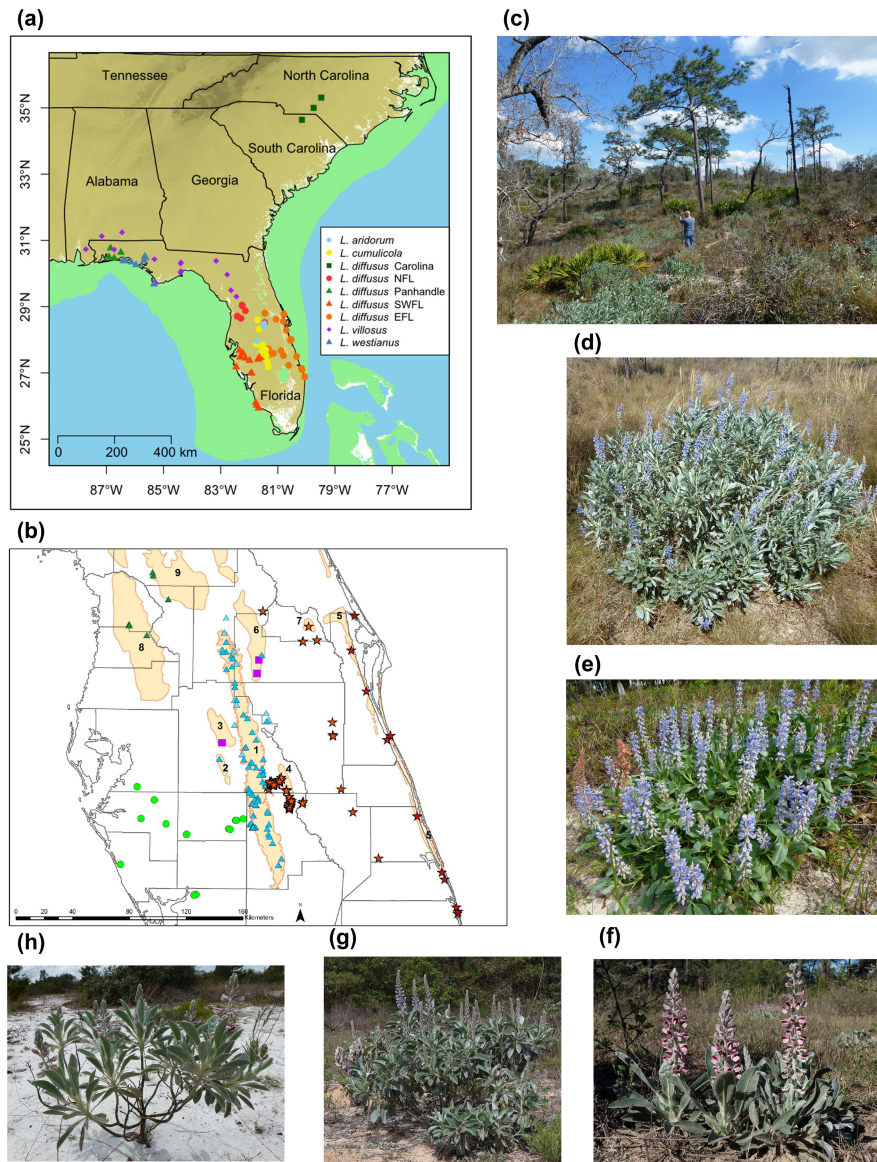


FIGURE 1 (a) Map showing the distributions of sample localities of the nine putative morphologically defined entities of the Florida unifoliolate *Lupinus* clade corresponding to putative species and intra-specific varieties which occupy essentially allopatric present-day distributions, as shown in b. The highly dynamic variation in the Florida coastline during the last 125 Kyr to the present day is indicated by delineation of the present-day coast and extent of Florida (black line), the last interglacial c. 125 Kyr when sea level was 6 m higher than at present (brown shading), and at the last glacial maximum 20 Kyr when sea level was 120 m lower than the present day and Florida was twice the area it is now (green shading). (b) Map of central peninsular Florida showing the distribution of *Lupinus* taxa in relation to the major sand ridges of this region which are numbered as follows: 1—Lake Wales Ridge; 2—Lake Henry Ridge; 3—Winter Haven Ridge; 4—Bombing Range Ridge; 5—Atlantic Coastal Ridge (discontinuous); 6—Mount Dora Ridge; 7—Geneva Hill; 8—Brooksville Ridge; 9—Sumter Upland. *Lupinus* taxa – *L. aridorum* (purple squares); *L. cumulicola* (blue triangles); *L. diffusus* EFL (red stars); *L. diffusus* SWFL (green circles); *L. diffusus* NFL (green triangles). No other *Lupinus* taxa occur within the area depicted. (c) Xeric pyrogenic sandhill habitat at Crooked Lake Sandhill on Lake Wales Ridge, Polk County, Florida with *L. cumulicola* (the silvery grey–green foliage in the foreground) and scattered *Pinus palustris*, xeric oaks, decumbent palms and clonal shrubs. (d) *L. cumulicola*. (e) *L. diffusus*. (f) *L. villosus*. (g) *L. westianus*. (h) *L. aridorum*. Photos: (c–e,h) E. Bridges, (f,g) Floyd Griffith.

with peninsular Florida reduced to a set of islands, while even at the last Pleistocene interglacial 125 Kyr, when sea level was 6 m higher than now, virtually the whole of present-day southern Florida and significant coastal areas of peninsular Florida and the Florida panhandle would have been inundated (Figure 1a), generating very recent coastal dune systems clearly visible today which also harbour

Lupinus populations (e.g. the Atlantic Coastal Ridge populations in peninsular Florida and *L. westianus* populations on sand ridges along and inland from the coast of the Florida panhandle). In contrast, at the Last Glacial Maximum c. 20 Kyr, with sea level c. 120 m lower than today, shorelines were dramatically altered such that Florida was more than twice the area it is today (Figure 1a), with the entire

modern coastline formed over just the last few 1000 years as a result of sea-level rise during the Holocene (Hine, 2013; Locker et al., 1996; see Germain-Aubrey et al., 2014, Figure 1). This recency of many of the habitats where *Lupinus* occurs in Florida is in line with the recent divergence time estimate of <1 Myr for the crown node of the Florida *Lupinus* clade (see below; Drummond et al., 2012) and provides a particular geological setting where the main impacts of sea-level fluctuations have likely been cycles of massive habitat expansion and contraction over the last 1 Myr.

In previous phylogenetic analyses of *Lupinus* (Drummond, 2008; Drummond et al., 2012; Eastwood et al., 2008; Hughes & Eastwood, 2006), the Florida unifoliolate-leaved species form a robustly supported clade which is moderately supported as sister to the Old World *Lupinus* clade, except in one study where a single Florida taxon was sampled and found to be nested within the Old World *Lupinus* clade with weak support (Keller et al., 2017). This likely sister group relationship to the Old World *Lupinus* clade is in line with chromosome numbers for Florida *Lupinus*, which at $2n=52$, show closer affinity to the Old World *L. albus*, *L. micranthus* and *L. luteus* ($2n=50-52$) than to other New World lineages ($2n=36/48$) (Conterato & Schifino-Wittmann, 2006; Eastwood et al., 2008). While the divergence time estimate for the split between the Old World and Florida clades is c. 10 Ma, diversification of the Florida clade is estimated to be very recent, with a crown node estimate of 0.9 Ma (Drummond et al., 2012). The lupines of Florida represent one of two independent derivations of unifoliolate from digitate leaves within the genus, the other in eastern South America (Eastwood et al., 2008). It is thus clear that the Florida clade is morphologically distinct, phylogenetically isolated and geographically well separated, from all other North American *Lupinus* whose diversity is heavily concentrated in western North America (Drummond et al., 2012; Sholars & Riggins, 2023).

The apparently very recent diversification of the Florida clade of *Lupinus* across the island-like system of sand ridges across Florida whose extent and connectivity have likely changed during Pleistocene sea-level fluctuations provides an excellent study system for investigating recent divergence across a continental edaphic island system and assessing to what extent Pleistocene glacial cycles may have contributed to recent speciation. We generate a densely sampled genome-wide RADseq dataset for all species and putative morphological variants of Florida *Lupinus* and undertake a series of phylogenomic and demographic analyses to understand species limits and patterns of historical gene flow and to estimate divergence times between species and morphological variants. Finally, we assess the taxonomic and conservation implications of our results for these xeric sand endemic plants in Florida.

2 | MATERIALS AND METHODS

2.1 | Field sampling

Geographic ranges and ecological and morphological diversity were assessed via survey of herbarium collections as well as extensive

fieldwork to survey and collect material from living plants as widely as possible across Florida (see Table S2 for detailed traits measured). Specimens or digital images of 596 herbarium collections were examined from FTG, FLAS, FSU, MICH, SWF and USF (acronyms follow Thiers [continuously updated]). Locality data from specimens, the Institute for Regional Conservation floristic database covering south Florida, floristic lists from Brevard County Environmental Areas and Lake Wales Ridge conservation sites were assembled and used to locate and map potential lupine populations. Maps showing xeric soils from county soil surveys and aerial imagery were used to identify xeric vegetation (sandhill, scrub and scrubby pinelands). Although sites with intact vegetation were prioritized, others were included since populations of *Lupinus* may often persist in disturbed or degraded habitats. Some sites had multiple visits due to fluctuating populations, with plants absent during unfavourable years, then subsequently displaying episodic mass flowering. We conducted field surveys of 300+ Florida sites in 35 counties, with a particular focus on central peninsular Florida where previously undocumented morphological variation corresponding to putative cryptic species had been observed. The majority of sites were visited during peak flowering, primarily in the late dry season (February to early May). At each population we collected silica-dried leaf material and voucher specimens, recorded field morphological characters (12 characters from 3 to 10 plants per population) and ecological/floristic information (soil colour, habitat, vegetation, associated plants) and photographed plant growth forms, habitats, inflorescences, leaf lamina and indumentum and stipules (Table S2). Final site visits during the early wet season allowed us to collect mature fruits and seeds from some populations. We plotted GPS locations using ARC view overlaid onto physiographic and soils data layers to determine the position of our sample sites in relation to Florida xeric ridges and soil series. Individuals (3–34) were sampled from across the geographic range of each of the nine putative morphological entities recognized during fieldwork (see below), giving a total of 106 accessions (Figure 1a). In addition, the Old World Mediterranean narrow-leaved lupine (*L. angustifolius* L.), a member of the putative sister group of the Florida unifoliolate clade (Drummond et al., 2012), was included as an outgroup. Leaf samples were collected from wild plants and dried in silica gel. Locality and voucher specimen details for all nextRAD sequenced individuals are listed in Table S1.

2.2 | NextRADseq preparation and sequencing

Total genomic DNA was extracted from silica-dried leaf material using a Qiagen DNeasy kit (Qiagen, Hombrechtikon, Switzerland) according to the manufacturer's guidelines. A Qubit Fluorometer (Thermo Fisher Scientific, Dietikon, Switzerland) was used to assess DNA quantity and gel electrophoresis was used to measure DNA quality and purity. We chose to generate RADseq data because they have proved to be powerful for analyses that span the species boundary and which, with dense sampling of species and intra-specific variation, can be used for both phylogenomic and demographic analyses of large numbers of loci scattered across

the genome (Atchison et al., 2016; Baird et al., 2008; Eaton & Ree, 2013; Nevado et al., 2018; Pante et al., 2015; Wagner et al., 2013). In addition, the large number of SNPs derived from numerous loci scattered across the genome generated using RADseq can provide powerful genetic evidence about species limits (e.g. Fujita et al., 2012; Herrera & Shank, 2016; Lamichhane et al., 2015; Wagner et al., 2013). Library preparation and sequencing of nextRAD markers from genomic DNAs was performed by SNPsaurus (SNPsaurus LLC).

To amplify genomic loci consistently between samples, the nextRAD method (Emerson et al., 2015; Russello et al., 2015) uses selective PCR primers. Genomic DNA was first fragmented with Nextera reagent (Illumina, Inc., San Diego, CA, USA), which also ligates short adapter sequences to the ends of the fragments. Variable amounts of genomic DNA were used as input into the fragmentation reaction to adjust for quality of DNA, with more input DNA for more degraded extracts. Fragmented DNA was then amplified, with one of the primers matching the adapter and extending seven nucleotides into the genomic DNA with the selective sequence (TGCAGAG). Thus, only fragments starting with a sequence that can be hybridized by the selective sequence of the primer will be efficiently amplified. The resulting fragments are fixed at the selective end, and have random lengths depending on the initial Nextera fragmentation. Because of this, amplified DNA from a particular locus is present at many different sizes and careful size selection of the library is not usually needed. Two libraries were sequenced, the first on an Illumina HiSeq 2000 to generate 100bp single end reads, and the second on an Illumina NextSeq to generate 150bp single end reads.

2.3 | NextRADseq assembly

Prior to assembly, raw reads were processed through Trimmomatic v0.33 (Bolger et al., 2014) to remove bases at the ends of reads with a quality score less than 20. Reads were assembled into loci using Stacks v 2.5 (Catchen et al., 2013) using the de novo approach implemented with the wrapper *denovo_map.pl*. This approach requires all reads to have the same length, thus as a first step we used the *process_radtags* function to trim all sequences to 90bp (reads shorter than 90bp were discarded at this step). The main parameters used to control the de novo assembly of RAD loci with stacks are the minimum stack depth (-m), the number of mismatches allowed between stacks within individuals (-M) and the number of mismatches allowed between stacks between individuals (-n). The choice of parameter values can strongly affect output (e.g. Paris et al., 2017), and we thus performed a set of preliminary runs with only a subset of data (31 individuals, representing all species in the dataset) and a range of values. We kept the minimum stack depth constant (-m=4) and explored all pairwise combinations of -M=3, 4, 5 and -n=6, 8, 10. Based on these preliminary results, we selected the best values for the three parameters and re-run *denovo_map.pl* with all the individuals in the dataset.

2.4 | Population structure analyses

To identify population structure among species and clarify the position of putative early generation hybrids (Appendix S2), we used two approaches. First, we used a principal component analysis (PCA) as implemented in the program plink v 2.0 (Chang et al., 2015; Galinsky et al., 2016). Second, we used faststructure v 1.0 (Raj et al., 2014) with a simple prior and a range of *K* values (number of clusters) from 2 to 10. Model complexity was evaluated with the *chooseK.py* script (part of faststructure distribution). The input data for both approaches was obtained from the Stacks pipeline using the *populations* command, and consisted of one random SNP per RAD locus (-write-random-snp) genotyped in at least 80% of the individuals (-R 0.8) after excluding singletons (--min-mac 3) and loci with overall observed heterozygosity above 0.7 (--max-obs-het 0.70).

2.5 | Phylogenetic reconstructions

We estimated phylogenetic relationships across samples with RaxML-NG v1.0 (Kozlov et al., 2019). The input data for this analysis was obtained from the Stacks pipeline using the *populations* command and consisted of the concatenation of all RAD loci where at least 80% of the individuals were sequenced (-R 0.8) after excluding loci with overall heterozygosity above 0.7 (--max-obs-het 0.70) and including all variable and invariant positions (--phylic-var-all). We further excluded all individuals that were morphologically intermediate as they likely represent early generation hybrids (Appendix S2). We used the GTR+G nucleotide substitution model for the concatenated dataset, performed 10 independent searches from random starting trees and assessed branch support by running 200 bootstrap replicates.

Our demographic analysis revealed instances of gene flow between several species (see Discussion), suggesting that evolutionary relationships between species in this clade might be better represented by a phylogenetic network instead of a strictly bifurcating phylogenetic tree. We used the Species Networks applying Quartets (SNaQ) (Solís-Lemus & Ané, 2016) method, implemented in the software Phylonetworks (Solís-Lemus et al., 2017), to (1) test whether a phylogenetic network (allowing hybridization between species) provides a better fit to the data compared to a phylogenetic tree and (2) infer how many hybridization events occurred during the diversification of this clade. As a first step in this analysis, we estimated a phylogenetic tree for each RAD locus using RaxML-NG v1.0 with the same settings as described earlier (GTR+G nucleotide substitution model, 10 independent searches from random starting trees). We excluded both recent hybrids (Appendix S2) and the two *L. angustifolius* individuals used as outgroups. The per-locus phylogenetic trees were then used in Phylonetworks to estimate quartet concordance factors with the function *countquartetstrees*. In this step, each individual was assigned to one of nine clades based on the combined morphological,

geographical and genetic evidence (see Discussion): *L. westianus*, *L. aridorum*, *L. villosus*, *L. diffusus* Carolina, *L. diffusus* panhandle, *L. cumulicola*, *L. diffusus* NFL, *L. diffusus* SWFL and *L. diffusus* EFL. Using the phylogeny obtained with the concatenated analysis as a starting tree, we then estimated the best phylogenetic network with varying number of hybridization events allowed (h between 0 and 4) using the function `snaq!` in `Phylonetworks`. To ensure convergence, we performed 25 independent runs under each value of h . We identified the best number of hybridization events in our dataset by comparing the log pseudolikelihood of the best network for each value of h : this pseudolikelihood is expected to increase sharply with h until it reaches the optimal value, and then to increase more slowly with increasing h (Solís-Lemus & Ané, 2016).

2.6 | Demographic analysis

A series of demographic analyses of population history were carried out to test for gene flow between species; assess whether the lack of resolution in the *L. cumulicola* clade is due only to recent divergence (incomplete lineage sorting) or ongoing gene flow and estimate approximate divergence times between a series of putative species pairs. Isolation with migration (IM) models were employed using the `dadi` package (Gutenkunst et al., 2009). This approach estimates the site frequency spectrum (SFS) as the distribution of allele frequencies across SNPs sampled from a population and compares the observed SFS to that expected under alternative demographic models. Using maximum likelihood, we can thus obtain estimates for the demographic parameters of interest and compare the fit of alternative demographic models to the observed data.

For the demographic analyses, we defined five populations based on taxonomic, morphological and geographic information: *L. villosus*, $n=15$; *L. westianus*, $n=8$; *L. diffusus* SWFL, $n=17$; *L. cumulicola*, $n=11$ and *L. diffusus* EFL, $n=24$. Because we are interested in inferring the demographic history of these taxa over evolutionary timescales, we excluded from this analysis the few individuals that were morphologically intermediate (Appendix S2). These individuals are likely to represent recent hybrids, and thus have reduced relevance in inferring times of divergence or amount of gene flow during diversification of this group. In order to take into account the sensitivity of the SFS analysis to genotype calling errors and employ a polarized SFS (where an ancestral allele at each site is known), a different approach was used to assemble SNP data from that used in the phylogenetic analysis. The RADseq reads were trimmed of adaptors and low-quality ends (base quality <20) using `cutadapt` v 1.8.3 (Martin, 2011) and mapped to the published scaffolds of the *L. angustifolius* genome (Yang et al., 2013; Assembly GCA_00338175 available on GenBank) with `bwa` v. 0.7.12 (Li & Durbin, 2009) and default values with the `mem` algorithm. Reads with mapping quality <20 were excluded, and the package `Stacks` v. 1.42 (Catchen et al., 2011, 2013) was used to extract only RAD loci present in at least 2 populations, with at most two SNPs and with minimum stack size of 10 (per individual).

For analysis in `dadi` v. 1.7.0, we used one random SNP per locus with the ancestral state inferred via comparisons with the *L. angustifolius* genome. Preliminary analysis revealed an excess of high-frequency variants, indicative of ancestral state misspecification, thus we masked the two highest frequency classes before fitting the demographic models. There are six free parameters in the isolation with migration (IM) model: the relative size of the two populations after splitting (s); time of the population/species split (T_d); population sizes of population/species 1 and 2 (N_1 and N_2) and rates of effective migration in two directions ($M_{1\leftarrow 2}$ and $M_{2\leftarrow 1}$). Demographic hypotheses can be tested via likelihood ratio tests by fixing values for some of these model parameters. Two-population IM models with full migration (IM), unidirectional migration (M12 and M21) and no migration (NoMig) were run for the following population pairs: *L. villosus* and *L. westianus*, *L. diffusus* SWFL and *L. cumulicola*, *L. diffusus* and *L. cumulicola*, *L. diffusus* SWFL and *L. diffusus*. The simplest best-fitting model was selected by comparing adjusted model likelihoods (Coffman et al., 2016) of the full IM model to each of the two unidirectional migration models, and each of these to a model without migration between species after divergence.

To convert estimated relative ages into absolute divergence times, we used a mutation rate of 7.0×10^{-9} mutations per site per generation (similar to available estimates in other plant species, e.g. Krasovec et al., 2018; Ossowski et al., 2010; Xie et al., 2016), the population size estimated from per site nucleotide diversity for each population (obtained directly from `Stacks`) and estimated generation time of 2 years, in line with field observations of flowering within 2 years.

3 | RESULTS

From field and herbarium survey it was observed that there are cryptic morphological differences between populations, including among the central peninsular Florida populations of the widespread *L. diffusus* group (Table S2), and that this morphological variation is strongly partitioned geographically. Nine variants were initially delimited based on morphology: *L. aridorum*, *L. cumulicola*, *L. diffusus* Carolina, *L. diffusus* north Florida (NFL), *L. diffusus* panhandle, *L. diffusus* south-west Florida (SWFL), *L. diffusus* east Florida (EFL), *L. villosus* and *L. westianus* (Figure 1). All of these putative morphological entities were found to occupy largely allopatric distributions with only very limited and infrequent range overlap (fewer than 5% of populations) (Figure 1a,b). During fieldwork a number of putative hybrid individuals were identified based on morphological intermediacy between the nine putative morphological entities (Appendix S2).

3.1 | NextRAD data assembly

Sequencing yielded an average of 3226 K reads per sample. Preliminary analyses with `Stacks` using different parameter values

showed that increasing the number of mismatches allowed between stacks within individuals from $m=3$ to $m=4$ caused a decrease in the number of loci present in 80% of the samples as well as the number of variable sites overall (Figure S1a,b). Increasing the number of mismatches to $m=5$ further reduced these values, but had a smaller effect. We thus set $m=4$ for the final analyses with the complete dataset. Conversely, increasing the number of mismatches allowed between individuals ($n=6, 8, 10$) caused an almost linear increase in the number of sites overall, but had a smaller effect on the number of loci retained (Figure S1c,d). Similar results were observed for other datasets in a previous simulation study (Paris et al., 2017). Here, we chose to use a value for $n=8$ for the analysis of the complete dataset, because though larger values of n continued to increase the number of polymorphic loci, it also implied larger computational costs, making the analysis of the complete dataset extremely slow.

3.2 | Population structure analyses

After filtering RAD loci to exclude singletons, loci with high heterozygosity, loci sequenced in fewer than 80% of samples and retaining only a single SNP per RAD locus, our final dataset used for analyses of population structure consisted of 5592 SNPs.

The first three axes of the PCA including all samples explained c. 25% of the variance in the dataset. The resulting plots along these three axes (Figure 2a–c) show a clear separation into six clusters: *L. aridorum*, *L. westianus*, *L. villosus*, *L. diffusus* panhandle, *L. diffusus* Carolina, and a sixth cluster containing all specimens belonging to *L. cumulicola* and *L. diffusus* NFL, SWFL and EFL. To better understand the structure within this sixth cluster, we performed a second PCA using only specimens assigned to it. This second PCA (Figure 2d) shows that the four morphological entities can be separated along the first two axes. Furthermore, of the 13 putative hybrids identified based on morphology, 10 are resolved as intermediate between *L. cumulicola* and the three *L. diffusus* groups.

Analysis of population structure using faststructure revealed that the most likely number of clusters (K) is between 2 and 4. Plots of individual assignments for these values of K (Figure 3a–c) are not entirely consistent across K values, but overall support the identification of the same six clusters found with PCA: *L. aridorum* is always resolved in a separate cluster; individuals belonging to *L. cumulicola* and *L. diffusus* NFL, SWFL and EFL are also always resolved as belonging to a separate cluster; individuals of *L. westianus* and *L. villosus* are resolved as either belonging to the same cluster ($K=2, 4$) or as two separate clusters ($K=3$) and individuals of *L. diffusus* Carolina and *L. diffusus* panhandle are generally resolved as admixed between different clusters. As for the PCA, we performed a second faststructure analysis focusing only on individuals belonging to the largest cluster which includes *L. cumulicola* and *L. diffusus* NFL, SWFL and EFL. The most likely number of clusters in this analysis is between 2 and 3, and the resulting individual assignment (Figure 3d,e) shows some separation between *L. cumulicola* and remaining taxa ($K=2, 3$) and between *L. diffusus* NFL and SWFL on one hand and *L. diffusus*

EFL on the other ($K=3$). This last analysis also revealed that several individuals (including almost all of the putative hybrids inferred based on morphology – see Appendix S2) have mixed ancestry between at least two of the clusters identified.

3.3 | Phylogeny

Phylogenetic analysis of the concatenated dataset (489,897 sites and 44,449 SNPs) using RaxML-NG recovered three main clades (Figure 4, Figure S2). The first clade (hereafter AWV clade) includes all individuals of *L. aridorum*, *L. westianus* and *L. villosus* and has high internal support and relatively long internal branches with each of the three species recovered as monophyletic with high support. In addition, there is phylogeographic structure within *L. westianus* with three inland accessions (F91, F92, F93) forming a well-supported sub-clade that is sister to a sub-clade comprising the coastal accessions. The second clade includes *L. diffusus* samples from North and South Carolina and the Florida panhandle. The third and largest clade comprises *L. cumulicola*, *L. diffusus* NFL, *L. diffusus* SWFL and *L. diffusus* EFL (hereafter referred to as the *L. cumulicola* clade). This clade has comparatively weaker internal support and shorter internal branches than the AWV clade, yet accessions of the putative morphological/geographical entities are almost always resolved as monophyletic (*L. cumulicola*, *L. diffusus* NFL and *L. diffusus* SWFL), with the only exception being three *L. diffusus* EFL samples (F67, F69 and F68) which are nested within a clade comprising *L. diffusus* NFL accessions, albeit with low support (Figure 4), a result not mirrored in the PCA and STRUCTURE analyses where these three accessions cluster unambiguously with the remaining EFL accessions (Figure 3).

The analysis of quartet concordance factors with the SNaQ method revealed that a phylogenetic network (allowing hybridization between species) provided a better fit to our data compared to a phylogenetic tree: the log pseudolikelihood increased sharply from $h=0$ (no hybridization allowed) to $h=2$ (2 hybridization events detected) (Figure S3). The best-fitting phylogenetic network with $h=2$ (Figure S3) shows the same topology as the phylogenetic tree obtained by concatenation (Figure 4). The first hybridization event detected with $h=2$ involved a relatively large introgression of *L. villosus* genetic material into the gene pool of *L. diffusus* panhandle (affecting c. 20% of the gene pool of the latter species), while the second event corresponded to a relatively minor introgression of genetic material of *L. diffusus* EFL into *L. cumulicola* (c. 2.5% of the genome affected).

3.4 | Demographic analyses

The best demographic model for each pair of populations (Table 1; Tables S3 and S4) always involves migration, suggesting that there is significant gene flow between all the populations and species tested. For one pairwise comparison (*L. diffusus* SWFL vs. *L. diffusus* EFL) independent runs did not converge, with highly variable parameter estimates across runs. This might be due to the model not accounting

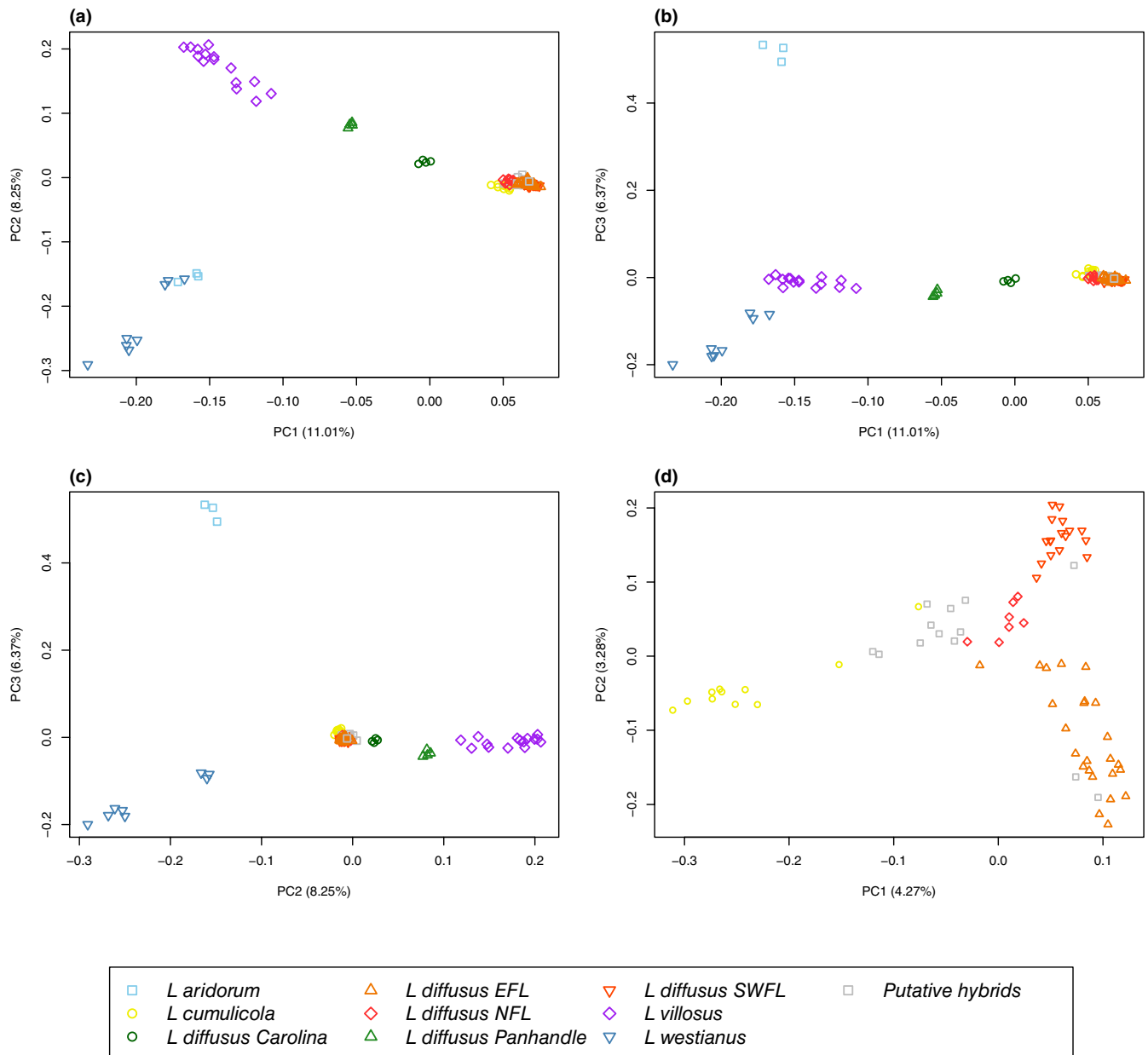


FIGURE 2 Principal component analysis of the nextRADseq dataset. (a–c) PCA using the entire dataset, with different projections over the first three axes which together explain c. 25% of the total variation observed in the dataset. (d) PCA including only specimens assigned to the *Lupinus cumulicola* clade (*L. cumulicola*, *L. diffusus NFL*, *L. diffusus SWFL*, *L. diffusus EFL* and putative hybrids based on morphological intermediacy) and showing only the first two axes (total variation explained c. 7.5%).

for gene flow between these two entities and *L. cumulicola* (which is geographically intermediate between these two entities and shows significant migration in pairwise models). From the remaining pairwise comparisons, migration levels are higher among populations within central peninsular Florida (*L. cumulicola* vs. *L. diffusus SWFL* and *L. cumulicola* vs. *L. diffusus EFL*) than between *L. villosus* and *L. westianus*, suggesting that the lack of phylogenetic support within the *L. cumulicola* clade is most likely due at least in part to higher levels of gene flow compared to the better resolved AWV clade. Estimated split times between species/populations (Table 2) vary depending on which population is used to calibrate the ancestral population size, and are also strongly dependent on what mutation

rate and generation time are assumed, suggesting that these estimates should be treated with caution. The estimated divergence time between *L. villosus* and the *L. westianus/L. aridorum* clade is somewhat older (77–235 Kyr) than between populations in the *L. cumulicola* clade which are estimated to range from 50 to 138 Kyr. Split times between *L. diffusus SWFL* versus *L. diffusus EFL* were not calculated as this pairwise analysis did not converge. Even taking into account the uncertainties surrounding these split time estimates, these estimates suggest that species diversification across the entire Florida unifoliolate-leaved clade is very recent indeed, most likely occurring towards the late Pleistocene, and largely within the last 150–250 Kyr.

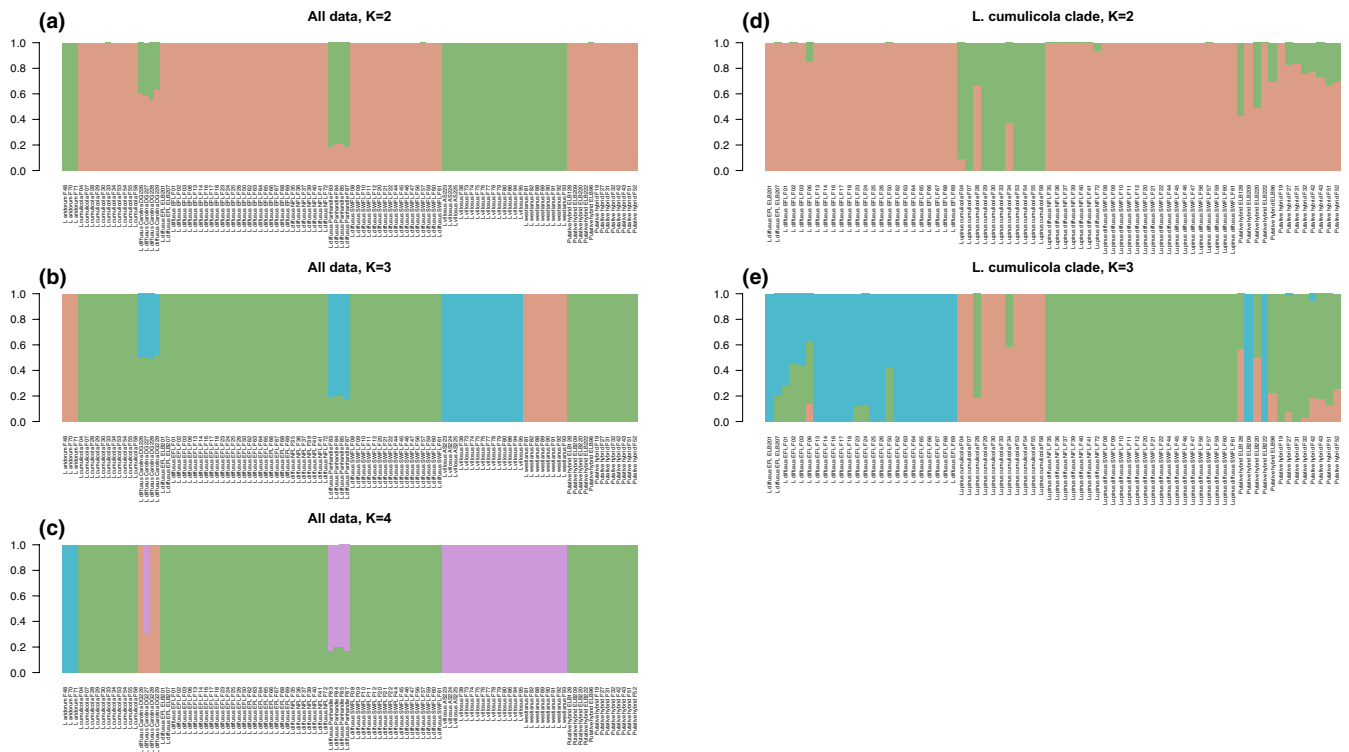


FIGURE 3 Barplots depicting genetic structure based on the nextRADseq dataset (a–c) across all Florida *Lupinus* specimens and (d, e) across specimens assigned to *L. cumulicola* clade. Each bar represents an individual, and the colours represent the fraction of its genome with ancestry within each cluster. Shown are results for different number of clusters assumed (from $K=2$ to $K=4$).

4 | DISCUSSION

4.1 | Late Pleistocene speciation in Florida

Here we deploy an extensive RADseq dataset to generate the first densely sampled and robustly supported phylogeny for the Florida unifoliolate-leaved *Lupinus* clade, estimate clade split times, detect putative hybrids and estimate levels of historical gene flow among putative species and morphological variants. Across peninsular Florida and adjacent areas of the southeast USA, our results reveal a set of four robustly supported reciprocally monophyletic clades which are congruent with genetic clusters found in multivariate and STRUCTURE analyses of SNP data (Figures 2–4), and which we equate with species (see taxonomic implications below): *L. westianus*, *L. aridorum*, *L. villosus* and the *L. diffusus* s.s., i.e., accessions of *L. diffusus* from outside the Florida peninsula. A fifth clade, here referred to as the *L. cumulicola* clade, containing all the peninsular accessions of *L. diffusus* and all individuals of *L. cumulicola* was also recovered albeit with low bootstrap support. The inferred phylogenetic relationships between these five clades are the same whether a concatenated approach (Figure 4) or a network approach that accounts for both incomplete lineage sorting and hybridization (Figure S3) is used. Furthermore, within the *L. cumulicola* clade, our data support recognition of a set of four morphological variants that are notably structured phylogenetically, albeit not all of them robustly supported as monophyletic, and which occupy largely allopatric geographic distributions almost completely confined to the

sand ridges of peninsular Florida (Figures 1b, 3e and 4). These almost entirely allopatric geographical distributions of *L. cumulicola* and the three *L. diffusus* variants (NFL, SWFL, EFL) across central peninsular Florida (Figure 1b) and the occurrences of these groups largely restricted to the major sand ridge systems: Lake Wales Ridge (*L. cumulicola*), Atlantic Coastal Ridge and Bombing Range Ridge (*L. diffusus* EFL) and Brooksville Ridge (*L. diffusus* NFL) which are isolated from each other by largely lupine-free habitats (Figure 1b) are striking. For instance, *L. cumulicola* – the first branching group within the *L. cumulicola* clade (Figure 4) – is almost never found outside the Lake Wales Ridge (LWR, labelled 1 in Figure 1b), whereas *L. diffusus* EFL is rarely found there but is common on the nearby Bombing Range Ridge (BRR, labelled 4 in Figure 1b). These two ridges share the same sand ridge soil type and are at some points separated by only a few kilometres, yet their ages are very different: BRR is thought to be Plio-Pleistocene in age, whereas parts of the LWR date back to the Pliocene (Green et al., 2019; Hardin, 2019; Hine, 2013; Scott, 2001; Webb, 1990; Weekley et al., 2008). This phylogenetic pattern is compatible with a scenario whereby the *L. diffusus* NFL, EFL and SWFL variants radiated across younger sand ridges from the older central Florida LWR sand ridge *L. cumulicola*, a scenario predicted by Schenk et al. (2018) and depicted in their figure 1. Similarly striking biogeographic differences in the biota of Florida sand ridges occur in other xeric endemic taxa (e.g. Branch & Hokit, 2000; Christman & Judd, 1990; Deyrup, 2005; Deyrup & Cover, 2004; Hill, 2023; Lamb et al., 2018; Schoonover McClelland et al., 2023). All this suggests a central role for sand ridge edaphic niche conservatism and allopatric



FIGURE 4 Phylogeny of the Florida *Lupinus* clade, reconstructed using the concatenated RADseq dataset excluding putative hybrid individuals. The phylogeny was rooted with the two *L. angustifolius* individuals (not shown). Bootstrap support denoted with symbols (squares BS = 100; triangles $90 \leq BS < 100$; circles $80 \leq BS < 90$). Scale bar is in expected substitutions per site.

non-adaptive divergence in the differentiation of these variants, and indeed across the Florida unifoliolate species clade as a whole, as found for other Floridian sand ridge plant clades (e.g. Naranjo et al., 2023; Schenk et al., 2018).

Analyses of divergence times from both previous phylogenies (crown node of the Florida clade < 1 Myr: Drummond et al., 2012) and split times from demographic analyses presented here (Table 2) suggest that all of this species and intra-specific diversification

TABLE 1 Parameter estimates for best demographic model for each pair of species/populations.

| Pop1 | Pop2 | Best model | s | nu1 | nu2 | T | M12 | M21 | Theta |
|-------------------------|------------------------|------------|-------|-------|-------|-------|-------|--------|-----------------------|
| <i>Lupinus villosus</i> | <i>L. westianus</i> | M21 | 0.559 | 4.390 | 1.368 | 0.396 | 0.000 | 0.452 | 357.387 |
| <i>L. diffusus</i> SWFL | <i>L. cumulicola</i> | M21 | 0.589 | 3.778 | 6.385 | 0.325 | 0.000 | 1.116 | 377.052 |
| <i>L. diffusus</i> SWFL | <i>L. diffusus</i> EFL | M21 | 0.111 | 0.830 | 0.951 | 0.542 | 0.000 | 11.147 | 1063.780 ^a |
| <i>L. cumulicola</i> | <i>L. diffusus</i> EFL | IM | 0.502 | 4.021 | 7.575 | 0.564 | 0.528 | 1.070 | 412.658 |

Note: Models fitted to data names are: IM – migration allowed in both directions; NoMig – no migration allowed; M12 – migration allowed only from population 2 to population 1; M21 – migration allowed only from population 1 to population 2. Parameters abbreviations are: s – size of population 1 at time of split relative to size of ancestral population (pop2 size was 1–s). nu1 and nu2 – current population sizes of populations 1 and 2 (relative to population size before splitting). T – time of split (in units of $2 \times$ ancestral population size [Nanc]). M12 and M21 – migration value from population 2 to 1 and from population 1 to 2 respectively (in units of $2 \times$ Nanc * m, with m = proportion of population consisting of immigrants in each generation, and Nanc = population size before split, which is not a free parameter in the models).

^aPairwise analysis of SWFL and EFL did not converge, with highly variable parameter estimates across runs. This might be due to the model not accounting for gene flow between these 2 entities and *L. cumulicola* (which is geographically intermediate between these two entities, and shows significant migration in pairwise models).

TABLE 2 Estimated split times between species/populations.

| Pop1 | Pop2 | T | Pop1 π /nt | Pop2 π /nt | Ne1 | Ne2 | N_ANC (pop1) | N_ANC (pop2) | ABS TIME (pop1) | ABS TIME (pop2) |
|-------------------------|------------------------|--------|----------------|----------------|---------|---------|--------------|--------------|-----------------|-----------------|
| <i>Lupinus villosus</i> | <i>L. westianus</i> | 0.3956 | 0.006 | 0.0057 | 214,286 | 203,571 | 48,812 | 148,810 | 77,240 | 235,476 |
| <i>L. diffusus</i> SWFL | <i>L. cumulicola</i> | 0.3254 | 0.0056 | 0.0069 | 200,000 | 246,429 | 52,938 | 38,595 | 68,904 | 50,235 |
| <i>L. cumulicola</i> | <i>L. diffusus</i> EFL | 0.5641 | 0.0069 | 0.0081 | 246,429 | 289,286 | 61,285 | 38,190 | 138,284 | 86,171 |

Note: Ne1 and Ne2 are current population sizes for pop1 and pop2, calculated from the π /nt for each population divided by $4 \times$ mutation rate. N_ANC(pop1) and N_ANC(pop2): are the population sizes of the ancestral population before the split between pop1 and pop2, calculated using the current population size of pop1 or pop2. Example: best model for *villosus-westianus* has nu1 = 4.39 and nu2 = 1.368. This means the current population size of population 1 (*villosus*) is estimated as $4.39 \times$ the ancestral population size before the split. If we use our estimate for current population size of *villosus* (214,286) we see that ancestral population size was $214,286/4.39 = 48,812$. We can do the same using population 2 (*westianus*) and get $203,571/1.368 = 148,810$.

occurred during the late Pleistocene. This recency of the Florida *Lupinus* species is supported by divergence time estimates from the demographic analyses which range from 77 to 235 Kyr for the *L. westianus*/*L. villosus* split which is somewhat older than the 50 to 138 Kyr estimates for splits between the *L. diffusus* NFL, EFL, SWFL morphological variants within the *L. cumulicola* clade. While age estimates for the splits between these peninsular variants are inevitably tentative depending on the mutation rate and generation time specified in the demographic model, they are likely confined to just the last 100–250 Kyr. This is very much in line with the weakly supported pattern of differentiation of these varieties from the older Lake Wales Ridge *L. cumulicola* towards both coasts and including the *L. diffusus* EFL populations on the most recently formed Bombing Range Ridge and Atlantic Coastal Ridge.

These age estimates were obtained by modelling pairwise speciation events (Table 2). This approach could result in biased estimates if significant indirect gene flow exists, i.e. gene flow between the focal pair of species occurred via a third, unsampled species. However, our joint analysis of all species in a network framework (Figure S3) detected only two hybridization events. One of these events is consistent with our results from demographic modelling, with gene flow between *L. diffusus* EFL and *L. cumulicola*. The second event is unlikely to affect our demographic analyses as it involves

gene flow from *L. villosus* into *L. diffusus* panhandle, and the latter species was not used in the demographic modelling. Thus, our dating estimates are not likely to be strongly biased by indirect gene flow.

These divergence time estimates, alongside the higher resolution and longer internal branches on the phylogeny in the AWV clade (Figure 4), suggest that these species are potentially somewhat older than the *L. cumulicola* clade spanning central peninsular Florida. Nevertheless, divergence time estimates across the whole of the Florida *Lupinus* clade, whether from phylogenies (Drummond et al., 2012) or the demographic analyses presented here, suggest that diversification across the entire Florida unfoliolate species clade occurred during the late Pleistocene. While previous work suggested that some endemic central Florida sand scrub taxa had pre-Pleistocene origins (Germain-Aubrey et al., 2014), our results are in line with studies of other plant groups which suggest that the majority of endemic Florida sandhill and sand scrub endemics diversified during the Pleistocene (Edwards et al., 2008; Naranjo et al., 2023; Oliveira et al., 2007; Schenk et al., 2018), as found in some animal clades (e.g. Krysko et al., 2016).

Our results thus provide an example of Pleistocene speciation that is not associated with classical flickering habitat connectivity caused by Pleistocene glacial cycles in island, mountain and lake settings where the majority of examples of Pleistocene speciation

have been documented (e.g. Comes et al., 2008; He et al., 2019; Kadereit & Abbott, 2021; Knowles, 2000; Nevado et al., 2018; Sholihah et al., 2021). In Florida, it is clear that most of the sand ridge islands where *Lupinus* species variants grow would not have been submerged by sea-level rises at least during the most recent glacial cycles. The overall topography of this area suggests that these sea-level changes may not have caused major vicariance events (Figure 1a), but would nevertheless have contributed to the isolation of sand ridge islands, e.g. the disjunction between *L. diffusus* s.s. and the *L. cumulicola* clade corresponds to the largely low-lying 'Suwannee Straits' that span NE Florida (Webb, 1990). Instead of flickering connectivity as a driver of speciation and diversification, in Florida, the most notable impact of the Pleistocene glacial cycles was cyclical expansion and contraction of the available land area that involved a doubling of the area of peninsular Florida c. 20 Kyr (Figure 1a) and recurrent formation of new sand ridges, resulting from sea-level fluctuations. These newly emerging post-glacial sand ridges potentially provided opportunities for colonization and divergence. We document divergence of two entities that occur on recently formed post-glacial coastal sand ridges: the *L. diffusus* EFL populations along the Atlantic Coastal Ridge dating to just the last few thousand years (Hine, 2013; Lane, 1994), and *L. westianus* which is endemic along the coastal sand ridges of the Florida panhandle that were also inundated at the last interglacial and on sand ridges immediately inland. Notably, for *L. westianus*, there is a phylogeographic split between a coastal sub-clade and a sub-clade confined to nearby sand ridges 30km inland indicative of divergence following colonization of the post-glacial coastal sand ridge systems, a geographic pattern replicated in the genus *Paronychia* with *P. erecta* on the coastal panhandle sand ridges and *P. minima* very narrowly endemic on the inland sand ridge (Schenk et al., 2018). This suggests a significant degree of isolation of these two nearly adjacent sand ridges. Our results suggest that divergence of all these endemic sand ridge entities coincided with the late Pleistocene when the land area of peninsular Florida expanded from a minimum at the last interglacial (c. 125Kyr), to almost twice its current area at the last glacial maximum (c. 20Kyr) and then progressively contracted again to the present-day coastlines as sea-level rose spawning new sand ridges. These split time estimates are thus compatible with a scenario whereby the massive expansion of land in peninsular Florida and of potential sand ridge habitats for Lupines between 125 and 20Kyr as sea levels dropped to their low point at the last glacial maximum provided opportunities for morphological and genetic differentiation of the *L. diffusus* NFL, EFL and SWFL variants as their ranges expanded across emerging sand ridges becoming isolated from the core central Lake Wales Ridge range of *L. cumulicola*. Subsequent retreat back to modern coastlines accompanying range contraction in post-glacial times has likely brought these variants back into secondary contact with *L. cumulicola* in a few specific locations in the immediate vicinity of the core *L. cumulicola* distribution, spawning putative hybrids and gene flow, as we observe for populations on Lake Henry Ridge (Appendix S2). This suggests that shifting availability of xeric ridges across peninsular Florida brought about by late Pleistocene sea-level

fluctuations, while not the main driver of speciation and diversification, also played an important role in driving diversification of xeric sand specialists.

The results of our demographic and phylogenetic analyses are very much in line with what might be expected for recent speciation and formation of incipient species, where reproductive isolation is incomplete and is compatible with the idea of ephemeral species (Rabosky, 2013; Rosenblum et al., 2012) whereby speciation is common and rapid, but the majority of produced species do not necessarily persist, but instead go extinct or are re-absorbed into parental forms. The idea that the variants within the *L. cumulicola* clade represent incipient species is reinforced by the lack of support along the backbone of the clade (Figure 4), which is probably due in part to recency and also likely to be a function of on-going gene flow, as shown by the demographic analyses and as manifest by the occurrence of putative hybrid individuals, between the various *L. diffusus* morphological variants and *L. cumulicola* (Appendix S2). For example, the admixed genotypes between the SWFL *L. diffusus* variant accessions and *L. cumulicola* accessions from the Lake Henry Ridge (Figures 2 and 3) show approximately equal proportions of these groups, supporting the putative recent hybridity of these individuals. This genetic admixture coincides with morphological intermediacy in plant habit, varying from prostrate to upright within populations, and intermediate leaf pubescence length, density and orientation between these putative variants. Similarly, the admixture seen in the putative *L. cumulicola* x *L. diffusus* EFL hybrids (Figures 2 and 3) corroborates the observed morphological intermediacy of these individuals between these two varieties (Appendix S2). Significant gene flow between all these varieties and the occurrence of these hybrids shows clearly that these variants are not reproductively isolated.

The one exception to the monophyly of the four *L. cumulicola* clade variants (*L. cumulicola*, and *L. diffusus* NFL, EFL, SWFL) is the placement of three putative *L. diffusus* EFL accessions (F67, F68 and F69) in a clade comprising the *L. diffusus* NFL accessions (Figure 4). It is notable in the PCA and STRUCTURE analyses these three accessions cluster unambiguously with other EFL accessions, and their phylogenetic placement with accessions of NFL in the phylogeny is weakly supported. Thus, there is limited genetic support to suggest that these accessions belong with NFL. Re-examination of the morphology of these three accessions in the light of their anomalous placement in the phylogeny suggests some degree of morphological intermediacy between NFL and EFL, but this remains inconclusive and is not reflected in the STRUCTURE plots (Figure 3). It is perhaps notable that these plants were collected from small, long un-burned populations in deeply shaded habitats. Additional sampling will be required from these populations to ascertain their true identities.

Overall, we conclude that diversification of the Florida unifoliolate clade species and variants of *Lupinus* was driven by a combination of isolation and allopatric speciation across the sand ridge island system mediated by strong edaphic niche conservatism, and the recent emergence of new sand ridges associated with Pleistocene

sea-level fluctuations. This has two important implications regarding the effect of Pleistocene climatic oscillations on diversification processes. First, it highlights the importance of even modest geographic isolation in dispersal limitation and suggests a central role for even short periods of allopatric divergence following range expansion in the emergence of new taxa. Second, it adds to the growing body of literature (reviewed in Kadereit & Abbott, 2021) in acknowledging that Pleistocene climatic oscillations may contribute to rapid diversification in a myriad of physiographic settings, not only in the more widely studied island archipelagos, 'sky' island and continental lake systems, but wherever geographic, edaphic and climatic factors combine to promote periods of geographical isolation and opportunities associated with expansion and emergence of new areas of suitable habitat.

4.2 | Biogeography of the Florida *Lupinus* clade

Research to understand the origins of the narrowly endemic Florida sand ridge scrub biota has ascertained the sister group relationships, divergence times and biogeographic affinities of specific endemic sand ridge taxa (Germain-Aubrey et al., 2014), and assessed whether they are related to species and lineages in the western USA originating via vicariance associated with mid-Pliocene scrub habitat fragmentation, or instead to eastern USA species and lineages that diverged more recently during the Pleistocene. The study by Germain-Aubrey et al. (2014) suggested that three of the four narrow Florida sand ridge endemics that they studied show affinities to eastern North American species and lineages, and estimated divergence times spanning both the Pliocene and Pleistocene. Here we show that the biogeographic affinities of *L. aridorum*, the narrowly restricted central Florida sand ridge endemic and indeed the other peninsular Florida endemic species and variants of *Lupinus*, are with *Lupinus* species that occur elsewhere in the south-eastern United States, in line with the predominant pattern of easterly origins of Florida sand ridge endemic plants (Germain-Aubrey et al., 2014) and with the easterly Pleistocene hypothesis of Schenk et al. (2018). This pattern of recent Pleistocene diversification observed for Florida *Lupinus* is strongly reminiscent of patterns seen in the genera *Conradina* (Edwards et al., 2008), *Dicerandra* (Oliveira et al., 2007) (both members of the Scrub Mint Clade sensu Naranjo et al., 2023 of Lamiaceae) and *Paronychia* (Caryophyllaceae) (Schenk et al., 2018). Each of these plant clades comprises between 5 and 10 Floridian species, is largely or completely restricted to Florida scrub and sandhill habitats of Florida and the southeastern USA and includes several highly localized narrowly restricted endemics often restricted to individual sand ridges which evolved recently (Edwards et al., 2008; Naranjo et al., 2023; Oliveira et al., 2007; Schenk et al., 2018). This suggests that similar processes of allopatric isolation and species differentiation across a set of edaphic islands formed by the sandhill and Florida scrub habitats, some of which emerged only during the late Pleistocene, shaped the divergence of species in all of these plant clades.

4.3 | Taxonomic implications

The lack of taxonomic consensus over how many species should be recognized in the Florida *Lupinus* clade, ranging from one to five over the last 200+ years (see Appendix S1), is stark, and has spawned conflicting assessments of the conservation threat status of endangered species in this clade. Our results reveal a novel hypothesis of relationships and species limits which sheds new light on previous taxonomies and raises several issues related to species delimitations.

First, we find robust support for *L. aridorum* and *L. westianus* as distinct, reciprocally monophyletic sister clades. Our analyses show that *L. aridorum* is strongly differentiated genetically from *L. westianus* with multiple accessions of each of these species forming robustly supported (BS 100%) sister clades subtended by long branches indicative of substantial genetic divergence. These two clades occupy disjunct distributions geographically isolated from each other, *L. westianus* restricted to coastal and inland sand ridges in the Florida panhandle and *L. aridorum* restricted to inland sand ridges in central peninsular Florida (Figure 1) and are distinguished by a suite of minor but consistent morphological differences in plant stature, branching habit, flower colour and leaflet size. Taken together, the combined phylogenetic, geographical and morphological evidence supports recognition of two distinct species.

Second, we show that *L. diffusus*, which has traditionally been considered to occur widely across both peninsular Florida, the Florida panhandle and further north into Alabama, Georgia, South and North Carolina, is non-monophyletic, with the *L. diffusus* populations from peninsular Florida (NFL, SWFL and EFL) more closely related to *L. cumulicola* than to the more northerly populations of *L. diffusus* (Figure 4). This close relationship between *L. cumulicola* and material from peninsular Florida currently assigned to *L. diffusus* is also supported by multivariate and genetic structure analyses of SNP data (Figures 2 and 3), and sheds new light on the conflicting taxonomic status of these two species (Duncan & McCartney, 1992; Isely, 1990, 1998). This result could be taken to support Isely's (1998) view that *L. cumulicola* presents no more than a peninsular form of *L. diffusus*, or alternatively, Dunn's (1971) placement of some elements of *L. diffusus* from peninsular Florida within *L. cumulicola*. As pointed out by Duncan and McCartney (1992), *L. diffusus* and *L. cumulicola* can be separated by a suite of at least six morphological characters supporting recognition of these two clades as distinct species. It is notable that this split between *L. diffusus* s.s. to the north and the *L. cumulicola* clade which encompassed the remainder of *L. diffusus* s.l., that is, the NFL, SWFL and EFL variants, coincides with a significant geographical disjunction corresponding the low-lying 'Suwannee Straits'. Denser sampling of *L. diffusus* across Alabama, Georgia and South Carolina to complement our current sampling which was restricted to the Carolinas and the Florida panhandle would be desirable to confirm the monophyly of that species in its re-circumscribed form.

Third, within the *L. cumulicola* clade, there is striking evidence of phylogeographic structure largely corresponding to our field-based

morphological variants (*L. diffusus* EFL, NFL and SWFL) that is congruent with the almost entirely non-overlapping geographic distributions of these variants (Figures 1b, 2 and 4). This near-complete phylogenetic, morphological and geographical congruence suggests that these variants also merit taxonomic recognition. However, evidence of inter-variant hybrids, limited phylogenetic support for some of these sub-clades and evidence of significant gene flow indicative of incomplete reproductive isolation argue for recognition of these entities at intra-specific rather than species rank, as possible named varieties of *L. cumulicola*.

Our results thus provide genetic, geographical and morphological evidence for delimiting five species within the Florida *Lupinus* clade: *L. diffusus* (re-defined to include only material from northern Florida and adjacent States to the north), *L. aridorum*, *L. westianus*, *L. villosus* and *L. cumulicola* (expanded to include all the morphological variants from central peninsular Florida) (Figure 1d–h). These findings will be presented in a forthcoming taxonomic account of the clade (E. L. Bridges & S. Orzell, unpublished data).

4.4 | Conservation implications

The Florida xeric sand ridges and uplands contain one of the highest concentrations of narrowly restricted endemic plants in the south-eastern USA and are considered one of the most threatened habitats in North America (Richardson et al., 2014). Within Florida 407 plant species have been classified as endangered (Ward et al., 2003). Within the south-eastern United States, the Florida panhandle and central peninsular Florida are epicentres of xeric plant endemism, with many being either state or federally listed as threatened or endangered species (Christman & Judd, 1990; Estill & Cruzan, 2001; Menges et al., 2007; Richardson et al., 2014). The degradation and destruction of habitat due to encroaching *Citrus* agriculture, and especially urban development – the sub-urbanization of central Florida – is a major threat to these globally rare, narrow endemics, including some of the unifoliolate *Lupinus* taxa. Within these habitats, *L. aridorum* is one of the most critically endangered plant species in Florida, occupying just the Winter Haven and Mount Dora ridges in Polk and Orange counties in central Florida (Figure 1a,b), with the number of known localities declining from 15 (U. S. Fish and Wildlife Service, 1987) to just 9 (Peterson, 2016; Peterson et al., 2012; Richardson et al., 2014). Furthermore, all these populations are small (0.1–1 ha), mostly declining and mostly unprotected, prompting recent efforts to more closely monitor populations, propagate plants to augment populations (Peterson, 2016; Peterson et al., 2012; Richardson et al., 2014) and investigate the conservation genetics of this species (Peterson, 2016; Peterson et al., 2012; Ricono et al., 2015).

The lack of consensus surrounding the taxonomic status of *L. aridorum* and its treatment as a variety of *L. westianus* have detracted attention from its endangered red-listing status at both state and federal levels (Bibb et al., 2007; Contu, 2012; U.S. Fish and Wildlife Service, 1987). Here we show that *L. aridorum* is genetically strongly

differentiated from *L. westianus*, supporting its treatment as a distinct species and bringing renewed focus on its endangered status and conservation. Furthermore, recognition of *L. aridorum* and *L. westianus* as distinct species further highlights the threatened status of *L. westianus* which is also a Florida endemic, narrowly restricted within the Florida panhandle, in an area with several other narrowly restricted endemic legumes including *Rhynchosia cytisoides*, *Tephrosia mohrii* and *Baptisia hirsuta*. Within *L. westianus* there is robust phylogenomic support for the separation of the inland accessions (F91–F93) from those along the coastal dune systems (F81–F90) (Figure 4), indicative of limited dispersal and gene flow between these sand ridges separated by just 30 km. The genetic distinctiveness of these inland and coastal sub-clades within *L. westianus*, a pattern mirrored in the genus *Paronychia* with two distinct species, *P. erecta* on the coastal sand ridges and *P. minima* on the inland sand ridges (Schenk et al., 2018), suggests that it will be important to protect both areas to conserve the genetic diversity of these species and areas.

AUTHOR CONTRIBUTIONS

Study design: BN, GWA, CEH. Data collection: GWA, ELB, SO, CEH. Data analysis: BN, GWA. Manuscript preparation: BN, SO, CEH. All authors contributed comments to and agree with the final version of the manuscript.

ACKNOWLEDGEMENTS

This work was supported by funds from the Swiss National Science Foundation (Grants 31003A_135522 and 31003A_182453/1 to C.E.H.), the Natural Environment Research Council UK (Grant NE/K004352/1 to D.A.F) and the Fundação para a Ciência e a Tecnologia (Grants CEECIND/00229/2018, 2022.15825.CPCA and <https://doi.org/10.54499/PTDC/BIA-EVL/2398/2021> to B.N.). The authors thank all those who assisted with locating Florida lupine populations: Chris Becker, Tabitha Biehl, Reed Bowman, Glen Bupp, Kris DeLaney, Rodney Felix, Alex Griffel, Jean Huffman, Eric Menges, Steve (Sticky) Morrison, Manny Perez, Cheryl Peterson, Juliet Rynear, Paul Schmalzer, Jack Stout, Sam Van Hook, Wayne Taylor, George Wilder and Brett Williams, the numerous private landowners that kindly granted us permission to collect lupines on their property, Jack Stout and Kris DeLaney for field assistance, alerting us to several lupine localities, and kindly sharing their extensive knowledge of Florida lupines, George Wilder for assistance in the field in Collier County and Paul Schmalzer in Brevard County, Doug Goldman and Al Schotz for material from outside of Florida and all those who processed federal and state collection permits and staff at various Florida State Parks for their assistance with access. We thank herbarium curators Bruce Hansen (USF), Kent Perkins (FLAS) and George Wilder (Naples Botanical Garden) for digital images of specimens and handling other requests, and Richard Raebler at (MICH) for help with examining McFarlin's lupine specimens. We thank Floyd Griffith for allowing us to use some of his digital images of lupines. Finally, we thank Paul Etter and Eric Johnson at SNPsaurus, Oregon for guidance and help with RAD sequencing, the ScienceCloud service of S3IT, University

of Zurich for computational support, Florian Boucher and Yanis Bouchenak-Khelladi for constructive comments on the manuscript and Jens Ringelberg for cartographic assistance.

DATA AVAILABILITY STATEMENT

Raw Illumina reads for all samples were submitted to NCBI's SRA depository under BioProject ID PRJNA1024924.

ORCID

Bruno Nevado  <https://orcid.org/0000-0002-9765-2907>

Steve Orzell  <https://orcid.org/0000-0001-6183-4015>

REFERENCES

- Atchison, G. W., Nevado, B., Eastwood, R. J., Contreras-Ortiz, N., Reynel, C., Madriñán, S., Filatov, D. A., & Hughes, C. E. (2016). Lost crops of the Incas: Origins of domestication of the Andean pulse crop 'tarwi', *Lupinus mutabilis*. *American Journal of Botany*, *103*, 1592–1606.
- Baird, N., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., Selker, E. U., Cresko, W. A., & Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One*, *3*, e3376.
- Beckner, J. (1982). *Lupinus aridorum* J. B. McFarlin ex Beckner (Fabaceae), a new species from Central Florida. *Phytologia*, *50*, 209–211.
- Bibb, K., Jennings, M., & Knight, M. (2007). 5-Year review, scrub lupine/*Lupinus aridorum*. U.S. Fish and Wildlife Service, Southeast region, Jacksonville Ecological Services Field Office.
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*, 2114–2120.
- Branch, L. C., & Hokit, D. G. (2000). A comparison of scrub herpetofauna on two central Florida sand ridges. *Florida Scientist*, *63*, 108–117.
- Brochmann, C., Brysting, A. K., Alsos, I. G., Borgen, L., Grundt, H. H., Scheen, A. C., & Elven, R. (2004). Polyploidy in arctic plants. *Biological Journal of the Linnean Society*, *82*(4), 521–536.
- Catchen, J., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. (2011). Stacks: Building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics*, *1*, 171–182.
- Catchen, J., Hohenlohe, P., Bassham, S., Amores, A., & Cresko, W. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*, 3124–3140.
- Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, *4*, 7.
- Christman, S. P., & Judd, W. S. (1990). Notes on plants endemic to Florida scrub. *Florida Scientist*, *53*, 52–73.
- Coffman, A. J., Hsieh, P. H., Gravel, S., & Gutenkunst, R. N. (2016). Computationally efficient composite likelihood statistics for demographic inference. *Molecular Biology and Evolution*, *33*, 591–593.
- Comes, H. P., Tribsch, A., & Bittkau, C. (2008). Plant speciation in continental Island floras as exemplified by *Nigella* in the Aegean Archipelago. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1506), 3083–3096.
- Conterato, I. F., & Schifino-Wittmann, M. T. (2006). New chromosome numbers, meiotic behaviour and pollen fertility in American taxa of *Lupinus* (Leguminosae): Contributions to taxonomic and evolutionary studies. *Botanical Journal of the Linnean Society*, *150*, 229–240.
- Contu, S. (2012). *Lupinus westianus*. The IUCN Red List of Threatened Species 2012: e.T19891535A20101361. <https://doi.org/10.2305/IUCN.UK.2012.RLTS.T19891535A20101361.en>
- Deyrup, M. (2005). A new species of flightless pygmy mole cricket from a Florida sand ridge (Orthoptera: Tridactylidae). *Florida Entomologist*, *88*(2), 141–145.
- Deyrup, M., & Cover, S. (2004). A new species of *Odontomachus* ant (Hymenoptera: Formicidae) from inland ridges of Florida, with a key to *Odontomachus* of the United States. *Florida Entomologist*, *87*(2), 136–144.
- Drummond, C. S. (2008). Diversification of *Lupinus* (Leguminosae) in the western New World: Derived evolution of perennial life history and colonization of montane habitats. *Molecular Phylogenetics and Evolution*, *48*, 408–421.
- Drummond, C. S., Eastwood, R. J., Miotto, S. T. S., & Hughes, C. E. (2012). Multiple continental radiations and correlates of diversification in *Lupinus* (Leguminosae): Testing for key innovation with incomplete taxon sampling. *Systematic Biology*, *61*, 443–460.
- Duncan, W. H., & McCartney, R. B. (1992). About *Lupinus cumulicola* (Fabaceae). *Sida*, *15*, 346–347.
- Dunn, D. B. (1971). A case of long-range dispersal and "rapid speciation" in *Lupinus*. *Transactions Missouri Academy of Sciences*, *5*, 26–38.
- Eastwood, R. J., Drummond, C. S., Schifino-Wittmann, M. T., & Hughes, C. E. (2008). Diversity and evolutionary history of Lupins – Insights from new phylogenies. In *Lupins for health and wealth. Proceedings of the 12th International Lupin conference, Freemantle, Western Australia, 14–18 September 2008*. International Lupin Association.
- Eaton, D. R., & Ree, R. (2013). Inferring phylogeny and introgression using RADseq data: An example from flowering plants (*Pedicularis*: Orobanchaceae). *Systematic Biology*, *62*, 689–706.
- Edwards, C. E., Lefkowitz, D., Soltis, D. E., & Soltis, P. S. (2008). Phylogeny of *Conradina* and related southeastern scrub mints (Lamiaceae) based on GapC gene sequences. *International Journal of Plant Sciences*, *169*, 579–594.
- Emerson, K. J., Conn, J. E., Bergo, E. S., Randel, M. A., & Sallum, A. M. (2015). Brazilian *Anopheles darlingi* root (Diptera: Culicidae) clusters by major biogeographical region. *PLoS One*, *10*, e0130773.
- Estill, J. C., & Cruzan, M. B. (2001). Phylogeography of rare plant species endemic to the southeastern United States. *Castanea*, *3*–23.
- Filatov, D. A., Bendif, E. M., Archontikis, O. A., Hagino, K., & Rickaby, R. E. (2021). The mode of speciation during a recent radiation in open-ocean phytoplankton. *Current Biology*, *31*(24), 5439–5449.
- Flantua, S. G., O'Dea, A., Onstein, R. E., Giraldo, C., & Hooghiemstra, H. (2019). The flickering connectivity system of the north Andean páramos. *Journal of Biogeography*, *46*(8), 1808–1825.
- Fujita, M. K., Leaché, A. D., Burbrink, F. T., McGuire, J. A., & Moritz, C. (2012). Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology & Evolution*, *27*, 480–488.
- Galinsky, K. J., Bhatia, G., Loh, P. R., Georgiev, S., Mukherjee, S., Patterson, N. J., & Price, A. L. (2016). Fast principal-component analysis reveals convergent evolution of ADH1B in Europe and East Asia. *American Journal of Human Genetics*, *98*(3), 456–472.
- Germain-Aubrey, C. C., Soltis, P. S., Neubig, K. M., Thurston, T., Soltis, D. E., & Gitzendanner, M. A. (2014). Using comparative biogeography to retrace the origins of an ecosystem: The case of four plants endemic to the central Florida scrub. *International Journal of Plant Sciences*, *175*, 418–431.
- Green, R. C., Williams, C. P., Lamarche, A. G., Hebets, C. L., Albritton, C. K., & Evans, G. S. (2019). *Text to accompany geologic map of the USGS Bartow 30 × 60 minute quadrangle, central Florida*. Open-File Report 108.
- Gutenkunst, R. N., Hernandez, R. D., Williamson, S. H., & Bustamante, C. D. (2009). Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genetics*, *5*, e1000695.
- Haffer, J. (1969). Speciation in Amazonian forest birds: Most species probably originated in forest refuges during dry climatic periods. *Science*, *165*(3889), 131–137.
- Hardin, J. O. (2019). A GIS-based synthesis of relict shorelines in peninsular Florida. In *Ocean Engineering and Marine Sciences student publications 2*. Florida Institute for Technology.

- He, Z., Li, X., Yang, M., Wang, X., Zhong, C., Duke, N. C., Wu, C. I., & Shi, S. (2019). Speciation with gene flow via cycles of isolation and migration: Insights from multiple mangrove taxa. *National Science Review*, 6(2), 275–288.
- Herrera, S., & Shank, T. M. (2016). RAD sequencing enables unprecedented phylogenetic resolution and objective species delimitation in recalcitrant divergent taxa. *Molecular Phylogenetics and Evolution*, 100, 70–79.
- Hewitt, G. M. (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, 58(3), 247–276.
- Hewitt, G. M. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, 405, 907–913.
- Hill, J. G. (2023). A new *Floritettix* (Orthoptera, Acrididae, Melanoplinae) from the Bombing Range Ridge, Florida, USA. *Journal of Orthoptera Research*, 32(2), 133–142.
- Hine, A. C. (2013). *Geologic history of Florida: Major events that formed the sunshine state*. University Press of Florida.
- Hudson, A. G., Vonlanthen, P., & Seehausen, O. (2011). Rapid parallel adaptive radiations from a single hybridogenetic ancestral population. *Proceedings of the Royal Society B: Biological Sciences*, 278, 58–66.
- Hughes, C., & Eastwood, R. (2006). Island radiation on a continental scale: Exceptional rates of plant diversification after uplift of the Andes. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 10334–10339.
- Isely, D. (1986). Notes on Leguminosae: Papilionoideae of the Southeastern United States. *Brittonia*, 38, 356.
- Isely, D. (1990). *Vascular Flora of the Southeastern United States: Leguminosae (Fabaceae)*. University of North Carolina Press.
- Isely, D. (1998). *Native and naturalized Leguminosae (Fabaceae) of the United States (exclusive of Alaska and Hawaii)*. Monte L. Bean Life Science Museum.
- Kadereit, J. W., & Abbott, R. J. (2021). Plant speciation in the Quaternary. *Plant Ecology & Diversity*, 14, 105–142.
- Keller, J., Rousseau-Guétin, M., Martin, G., Morice, J., Boutte, J., Coissac, E., Ourari, M., Ainouche, M., Salmon, A., Cabello-Hurtado, F., & Ainouche, A. (2017). The evolutionary fate of the chloroplast and nuclear *rps16* genes as revealed through the sequencing and comparative analyses of four novel legume chloroplast genomes from *Lupinus*. *DNA Research*, 24(4), 343–358.
- Klicka, J., & Zink, R. M. (1997). The importance of recent ice ages in speciation: A failed paradigm. *Science*, 277(5332), 1666–1669.
- Knowles, L. L. (2000). Tests of Pleistocene speciation in montane grasshoppers (genus *Melanoplus*) from the sky islands of western North America. *Evolution*, 54(4), 1337–1348.
- Kozlov, A. M., Darriba, D., Flouri, T., Morel, B., & Stamatakis, A. (2019). RAxML-NG: A fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*, 35(21), 4453–4455.
- Krasovec, M., Chester, M., Ridout, K., & Filatov, D. A. (2018). The mutation rate and the age of the sex chromosomes in *Silene latifolia*. *Current Biology*, 28, 1832–1838.
- Krysko, K. L., Nuñez, L. P., Lippi, C. A., Smith, D. J., & Granatosky, M. C. (2016). Pliocene-Pleistocene lineage diversification in the Eastern Indigo Snake (*Drymarchon couperi*) in the Southeastern United States. *Molecular Phylogenetics and Evolution*, 98, 111–122.
- Lamb, T., Justice, T. C., Brewer, M. S., Moler, P. E., Hopkins, H., & Bond, J. E. (2018). A biogeographical profile of the sand cockroach *Arenivaga floridensis* and its bearing on origin hypotheses for Florida scrub biota. *Ecology and Evolution*, 8(11), 5254–5266.
- Lamichhaney, S., Berglund, J., Almén, M. S., Maqbool, K., Grabherr, M., Martinez-Barrio, A., Promerová, M., Rubin, C. J., Wang, C., Zamani, N., & Grant, B. R. (2015). Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature*, 518, 371–375.
- Lane, E. (1994). *Florida's geological history and geological resources*. Florida Geological Survey Special Publication 35.
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler Transform. *Bioinformatics*, 25, 1754–1760.
- Locker, S. D., Hine, A. C., Tedesco, L. P., & Shinn, E. A. (1996). Magnitude and timing of episodic sea-level rise during the last deglaciation. *Geology*, 24, 827–830.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal*, 17, 10–12.
- Menges, E. S., & Hawkes, C. V. (1998). Interactive effects of fire and microhabitat on plants of Florida scrub. *Ecological Applications*, 8, 935–946.
- Menges, E. S., Weekley, C. W., Hamzé, S. I., & Pickert, R. L. (2007). Soil preferences for federally-listed plants on the Lake Wales Ridge in Highlands County, Florida. *Florida Scientist*, 70, 24.
- Naranjo, A. A., Folk, R. A., Gitzendanner, M. A., Soltis, D. E., & Soltis, P. S. (2023). Ancestral area analyses reveal Pleistocene-influenced evolution in a clade of coastal plain endemic plants. *Journal of Biogeography*, 50, 393–405.
- Nevado, B., Contreras-Ortiz, N., Hughes, C., & Filatov, D. A. (2018). Pleistocene glacial cycles drive isolation, gene flow and speciation in the high-elevation Andes. *New Phytologist*, 219, 779–793.
- Nevado, B., Mautner, S., Sturmbauer, C., & Verheyen, E. (2013). Water-level fluctuations and metapopulation dynamics as drivers of genetic diversity in populations of three Tanganyikan cichlid fish species. *Molecular Ecology*, 22(15), 3933–3948.
- Oliveira, L. O., Huck, R. B., Gitzendanner, M. A., Judd, W. S., Soltis, D. E., & Soltis, P. S. (2007). Molecular phylogeny, biogeography, and systematics of *Dicerandra* (Lamiaceae), a genus endemic to the southeastern United States. *American Journal of Botany*, 94, 1017–1027.
- Ossowski, S., Schneeberger, K., Lucas-Llledo, J. I., Warthmann, N., Clark, R. M., Shaw, R. G., Weigel, D., & Lynch, M. (2010). The rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana*. *Science*, 327, 92–94.
- Pante, E., Abdelkrim, J., Viricel, A., Gey, D., France, S. C., Boisselier, M. C., & Samadi, S. (2015). Use of RAD sequencing for delimiting species. *Heredity*, 114, 450–459.
- Papadopoulou, A., & Knowles, L. L. (2015a). Species-specific responses to island connectivity cycles: Refined models for testing phylogeographic concordance across a Mediterranean Pleistocene aggregate island complex. *Molecular Ecology*, 24(16), 4252–4268.
- Papadopoulou, A., & Knowles, L. L. (2015b). Genomic tests of the species-pump hypothesis: Recent island connectivity cycles drive population divergence but not speciation in Caribbean crickets across the Virgin Islands. *Evolution*, 69(6), 1501–1517.
- Paris, J. R., Stevens, J. R., & Catchen, J. M. (2017). Lost in parameter space: A road map for stacks. *Methods Ecology and Evolution*, 8, 1360–1373.
- Peterson, C. L. (2016). *An Integrated Conservation Program for the protection of rare plant species of Florida: north and central region*. Final report to the State of Florida Department of Agriculture and Consumer Services, Contract 22645.
- Peterson, C. L., Bupp, G., & Rynear, J. (2012). *An Integrated Conservation Program for the protection of rare plant species of Florida: north and central region*. Final report to the State of Florida Department of Agriculture and Consumer Services, Contract 17165.
- Rabosky, D. L. (2013). Diversity-dependence, ecological speciation, and the role of competition in macroevolution. *Annual Review of Ecology, Evolution, and Systematics*, 44, 481–502.
- Raj, A., Stephens, M., & Pritchard, J. K. (2014). fastSTRUCTURE: Variational inference of population structure in large SNP data sets. *Genetics*, 197(2), 573–589.
- Richardson, M. L., Rynear, J., & Peterson, C. L. (2014). Microhabitat of critically endangered *Lupinus aridorum* (Fabaceae) at wild and introduced locations in Florida scrub. *Plant Ecology*, 215, 399–410.
- Ricono, A., Bupp, G., Peterson, C., Nunziata, S. O., Lance, S. L., & Pruett, C. L. (2015). Development and characterization of microsatellite

- loci for the endangered scrub lupine, *Lupinus aridorum* (Fabaceae). *Applications in Plant Sciences*, 3, 3–5.
- Rosenblum, E. R., Sarver, B. A., Brown, J. W., Des Roches, S., Hardwick, K. M., Hether, T. D., Eastman, J. M., et al. (2012). Goldilocks meets Santa Rosalia: An ephemeral speciation model explains patterns of diversification across time scales. *Evolutionary Biology*, 39, 255–261.
- Rull, V. (2011). Neotropical biodiversity: Timing and potential drivers. *Trends in Ecology & Evolution*, 26(10), 508–513.
- Russello, M. A., Waterhouse, M. D., Etter, P. D., & Johnson, E. A. (2015). From promise to practice: Pairing non-invasive sampling with genomics in conservation. *PeerJ*, 3, e1106.
- Schenk, J. J., Kontur, S., Wilson, H., Noble, M., & Derryberry, E. (2018). Allopatric speciation drives diversification of ecological specialists on sandhills. *International Journal of Plant Sciences*, 179, 325–339.
- Schoonover McClelland, R. K., Weakley, A. S., & Poindexter, D. B. (2023). Seven new species of *Trichostema* (Lamiaceae: Ajugoideae) from the North American Coastal Plain biodiversity hotspot. *Phytotaxa*, 603(2), 95–149.
- Scott, T. M. (2001). Text to accompany the geologic map of Florida. *Florida Geological Survey Open File Report*, 80, 1–28.
- Sholars, T., & Riggins, R. (2023). *Lupinus*. In *Flora of North America Editorial Committee* (Ed.), *Flora of North America*. Vol 11, *Magnoliophyta: Fabaceae Part 1* (pp. 193–248). Oxford University Press.
- Sholihah, A., Delrieu-Trottin, E., Condamine, F. L., Wowor, D., Rüber, L., Pouyau, L., Agnès, J. F., & Hubert, N. (2021). Impact of Pleistocene eustatic fluctuations on evolutionary dynamics in southeast Asian biodiversity hotspots. *Systematic Biology*, 70(5), 940–960.
- Simpson, B. B. (1974). Glacial migrations of plants: Island biogeographical evidence. *Science*, 185(4152), 698–700.
- Solís-Lemus, C., & Ané, C. (2016). Inferring phylogenetic networks with maximum pseudolikelihood under incomplete lineage sorting. *PLoS Genetics*, 12, e1005896.
- Solís-Lemus, C., Bastide, P., & Ané, C. (2017). PhyloNetworks: A package for phylogenetic networks. *Molecular Biology and Evolution*, 34, 3292–3298.
- Sturmbauer, C. (1998). Explosive speciation in cichlid fishes of the African Great Lakes: A dynamic model of adaptive radiation. *Journal of Fish Biology*, 53, 18–36.
- U.S. Fish and Wildlife Service. (1987). Endangered and Threatened Wildlife and Plants; endangered status for *Lupinus aridorum* (Scrub Lupine). *Federal Register*, 52, 11172–11175.
- Wagner, C. E., Keller, I., Wittwer, S., Selz, O. M., Mwaiko, S., Greuter, L., Sivasundar, A., & Seehausen, O. (2013). Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Molecular Ecology*, 22, 787–798.
- Ward, D. B., Austin, D. F., & Coile, N. C. (2003). Endangered and threatened plants of Florida, ranked in order of rarity. *Castanea*, 68, 160–174.
- Webb, S. D. (1990). Historical biogeography. In R. L. Myers & J. J. Ewell (Eds.), *Ecosystems of Florida* (pp. 70–100). University of Central Florida Press.
- Weekley, C. W., Menges, E. S., & Pickert, R. L. (2008). An ecological map of Florida's Lake Wales Ridge: A new boundary delineation and an assessment of post-Columbian habitat loss. *Florida Scientist*, 71(1), 45–64.
- Weir, J. T., Haddrath, O., Robertson, H. A., Colbourne, R. M., & Baker, A. J. (2016). Explosive ice age diversification of kiwi. *Proceedings of the National Academy of Sciences of the United States of America*, 113(38), E5580–E5587.
- Weir, J. T., & Schluter, D. (2004). Ice sheets promote speciation in boreal birds. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271(1551), 1881–1887.
- Xie, Z., Wang, L., Wang, L., Wang, Z., Lu, Z., Tian, D., Yang, S., & Hurst, L. D. (2016). Mutation rate analysis via parent-progeny sequencing of the perennial peach. I. A low rate in woody perennials and a higher mutagenicity in hybrids. *Proceedings of the Royal Society B: Biological Sciences*, 283, 20161016.
- Yang, H., Tao, Y., Zheng, Z., Zhang, Q., Zhou, G., Sweetingham, M. W., Howieson, J. G., & Li, C. (2013). Draft genome sequence, and a sequence-defined genetic linkage map of the legume crop species *Lupinus angustifolius* L. *PLoS One*, 8, e64799.
- Yu, R. Y., & Van Welzen, P. C. (2020). Historical biogeography of *Trigonostemon* and *Dimorphocalyx* (Euphorbiaceae). *Botanical Journal of the Linnean Society*, 192, 333–349.

SUPPORTING INFORMATION

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

How to cite this article: Nevado, B., Atchison, G. W., Bridges, E. L., Orzell, S., Filatov, D., & Hughes, C. E. (2024). Pleistocene diversification of unifoliolate-leaved *Lupinus* (Leguminosae: Papilionoideae) in Florida. *Molecular Ecology*, 00, 1–17. <https://doi.org/10.1111/mec.17232>