Received: 29 January 2024 | Accepted: 6 December 2024

DOI: 10.1002/ajb2.70005

RESEARCH ARTICLE

Botany

When the sand blossoms: Phylogeny, trait evolution, and geography of speciation in *Linanthus*

Ioana G. Anghel[1] | **Lydia L. Smith**[2,3] | **Isaac H. Lichter-Marck**[1,4] | **Felipe Zapata**[1,5]

[1]Department of Ecology and Evolutionary Biology, University of California, Los Angeles, California 90095, USA

[2]Museum of Vertebrate Zoology, University of California, Berkeley, California 94720, USA

[3]Department of Integrative Biology, University of California, Valley Life Sciences Building, Berkeley, California 94720, USA

[4]Institute for Biodiversity Science and Sustainability, California Academy of Sciences, San Francisco, California 94118, USA

[5]Center for Tropical Research, Institute of the Environment and Sustainability, University of California, Los Angeles, California 90095, USA

Correspondence

Ioana G. Anghel, Department of Ecology and Evolutionary Biology, University of California, Los Angeles, California 90095, USA.
Email: ianghel@ucla.edu

Abstract

Premise: Understanding how plants successfully diversified in novel environments is a central question in evolutionary biology. *Linanthus* occurs in arid areas of western North America and exhibits extensive floral trait variation, multiple color polymorphisms, differences in blooming time, and variation in life history strategies. We reconstructed the evolutionary history of this genus.

Methods: We generated restriction-site associated (ddRAD) sequences for 180 individuals and target capture (TC) sequences for 63 individuals, with complete species sampling. Using maximum likelihood and pseudo-coalescent approaches, we inferred phylogenies of *Linanthus* and used them to model the evolution of phenotypic traits and investigate the genus's geographic speciation history.

Results: Relationships are consistent and well supported with both ddRAD and TC data. Most species are monophyletic despite extensive local sympatry and range overlap, suggesting strong isolating barriers. The non-monophyly of the night-blooming and perennial species may be due to rapid speciation or cryptic diversity. Perenniality likely evolved from annuality, a rare shift in angiosperms. Night-blooming evolved three times independently. Flower color polymorphism is an evolutionarily labile trait that is likely ancestral. No single geographic mode of speciation characterizes this diversification, but most species overlap in range, which suggests that they evolved in parapatry.

Conclusions: Our results illustrate the complexity of phylogenetic inference for recent radiations, even with multiple sources of genomic data and extensive sampling. This analysis provides a foundation for understanding aridity adaptations, such as evolution of flower color polymorphisms, night-blooming, and perenniality, as well as speciation mechanisms.

KEYWORDS

California, desert plants, diversification, *Linanthus*, night-blooming, perennial, phylogeny, polymorphism, RADseq, target capture

Desert biomes make up 17% of Earth's land surface (Millennium Ecosystem Assessment, 2005) and appear lifeless for most of the year, when lack of precipitation and high temperatures hamper plant survival and growth. However, the desert comes alive in an explosion of color after it rains, when dormant seeds germinate and flowers form thick brushstrokes of pigment. This uncommon and brief outburst of life is made possible by the annual angiosperms, which await suitable environmental cues to germinate in mass. In California, deserts make up 38% of the total land area (Mooney and Zavaleta, 2016), and 52% of the plant species in these deserts are annuals (Calflora, 2023). This overrepresentation of annuals in the desert and their varied adaptations make them ideal

organisms to study the patterns and processes of diversification in this harsh environment.

The deserts of California and their diversity of annual species seem to be young on a geologic time scale, emerging <2 mya (Thorne, 1986). Meta-analyses of California plant diversity indicate that plants have recently diversified in this geographic area. For instance, Kraft et al. (2010) showed that the California desert regions have a high concentration of young species with restricted geographic ranges, estimating that Mojave Desert plants originated since the late Miocene. Thornhill et al. (2017) found a significant concentration of short phylogenetic branches (i.e., recent diversification events) restricted to the California deserts. However, how deserts facilitate the rapid diversification of plant species remains poorly understood. Habitat heterogeneity, large ranges with isolated populations, and a broad diversity of adaptations to xeric landscapes have all been proposed as potential drivers of high rates of speciation in desert species (Stebbins, 1952). The evolution of the annual habit seems to be associated with unstable environments with dry conditions and unpredictable rainfall (Friedman, 2020), which is typical in deserts. Annual plants have a fast rate of evolution, which is correlated with their short life cycle, isolated populations, and variable environment (Smith and Donoghue, 2008; Smith and Beaulieu, 2009). Deserts may promote the evolution of new annual plant species (Stebbins, 1952) and provide a system to study the largely unexplored correlation between harsh environments and plant diversification (Stebbins, 1952; but see Hernández-Hernández et al., 2014; Singhal et al., 2021; Lichter-Marck and Baldwin, 2023).

There are few phylogenetic studies of California desert annuals. Most published studies do not include all members of the focal clades and almost never include multiple individuals per species, limiting understanding of the patterns and processes of intraspecific and interspecific diversification. Additionally, most studies have used only a handful of loci, largely resulting in poorly resolved phylogenies (e.g., Spencer and Porter, 1997; Moore and Jansen, 2006; Evans et al., 2009; Porter et al., 2010; Cacho et al., 2014; Walden et al., 2014; Azani et al., 2019; Vasile et al., 2020). More recently, a limited number of studies have used genomic approaches and more extensive taxon sampling, shedding light on species-level relationships, patterns of within-species genetic variation, and potential drivers of diversification (Simpson et al., 2017; Mabry and Simpson, 2018; Lichter-Marck et al., 2020; Pearman et al., 2021; Rose et al., 2021; Singhal et al., 2021). Given that California deserts are young hotspots of biodiversity (Kraft et al., 2010), well-sampled groups are needed to assess the monophyly of lineages, elucidate species relationships, and untangle complex evolutionary patterns, typical of recent radiations. Thorough phylogenetic reconstructions are especially worthwhile in geographic regions where species overlap extensively in their geographic ranges and have ample opportunities for gene flow. Such robust phylogenetic studies could inform our understanding of the evolution of traits, adaptations to extreme environments, and whether aridity can promote evolution (Stebbins, 1952; Axelrod, 1972).

The genus *Linanthus* Benth. (Polemoniaceae) is an ideal system to study species diversification in the heterogeneous arid environments of southwestern North America. Half of the currently recognized species co-occur in a geologic transition zone characterized by exceptional plant endemism and environmental variation in southern California (Kraft et al., 2010). Nineteen of 25 species overlap in geographic range, and at least fourteen species pairs co-flower and co-occur at a local scale (I. G. Anghel, personal observation). Reproducing in a span of a few weeks in the spring, sympatric *Linanthus* species likely have strong reproductive isolation mechanisms to maintain their genetic and phenotypic integrity. *Linanthus* species also display extensive interspecific diversity in habit, blooming time, flower color, and floral scent. These traits attract a diverse suite of pollinators, including beetles, moths, butterflies, hoverflies, long-tongued flies, and bees (Chess et al., 2008; Rose and Sytsma, 2021; I. G. Anghel, personal observation), which likely facilitate the reproductive differentiation of species (Figure 1). In addition, seven *Linanthus* species have extremely restricted geographic ranges and are ranked as rare, threatened, or endangered in California, including *L. bellus* (A.Gray) Greene, *L. bernardinus* N.S. Fraga & D. Bell, *L. concinnus* Milliken, *L. jaegeri* (Munz) J.M. Porter & L.A. Johnson, *L. killipii* H. Mason, *L. maculatus* (Parish) Milliken, and *L. orcuttii* (Parry) Jeps. (California Native Plant Society). The extensive sympatry of species, their diversity of pollinator attraction strategies, and their asymmetric geographic ranges point to a complex speciation history that may include speciation with gene flow, micro-allopatry, parapatric speciation, and ecological isolation. However, these phenomena have not been examined in detail.

Linanthus species also show extensive intraspecific variability in petal color and nectary markings, with individual plants taking on different phenotypes in the same population or in disparate areas of the species range. Eleven of the 26 species are petal color polymorphic, with both white and colored individuals across their ranges (Figure 1: I and J, O and P, R and S, T and U, V and X). Seven of these color polymorphic species exhibit within-population color variation, with both colors present in the same population (I. G. Anghel, personal observation). One notable example is *L. parryae* (A. Gray) Greene, the purple and white polymorphic species at the center of the classic evolutionary debate investigating whether natural selection or genetic drift maintains intraspecific polymorphisms. Different color morphs seem to fare better in wet or dry years, a potential adaptation to the variable desert environment (Epling and Dobzhansky, 1942; Wright, 1943; Schemske and Bierzychudek, 2007). Another potential strategy of several *Linanthus* species for survival in arid environments is night anthesis, whereby flowers are open only at night and closed during the day during the flowering season. Heat can have a suite of negative fitness repercussions on flowers,

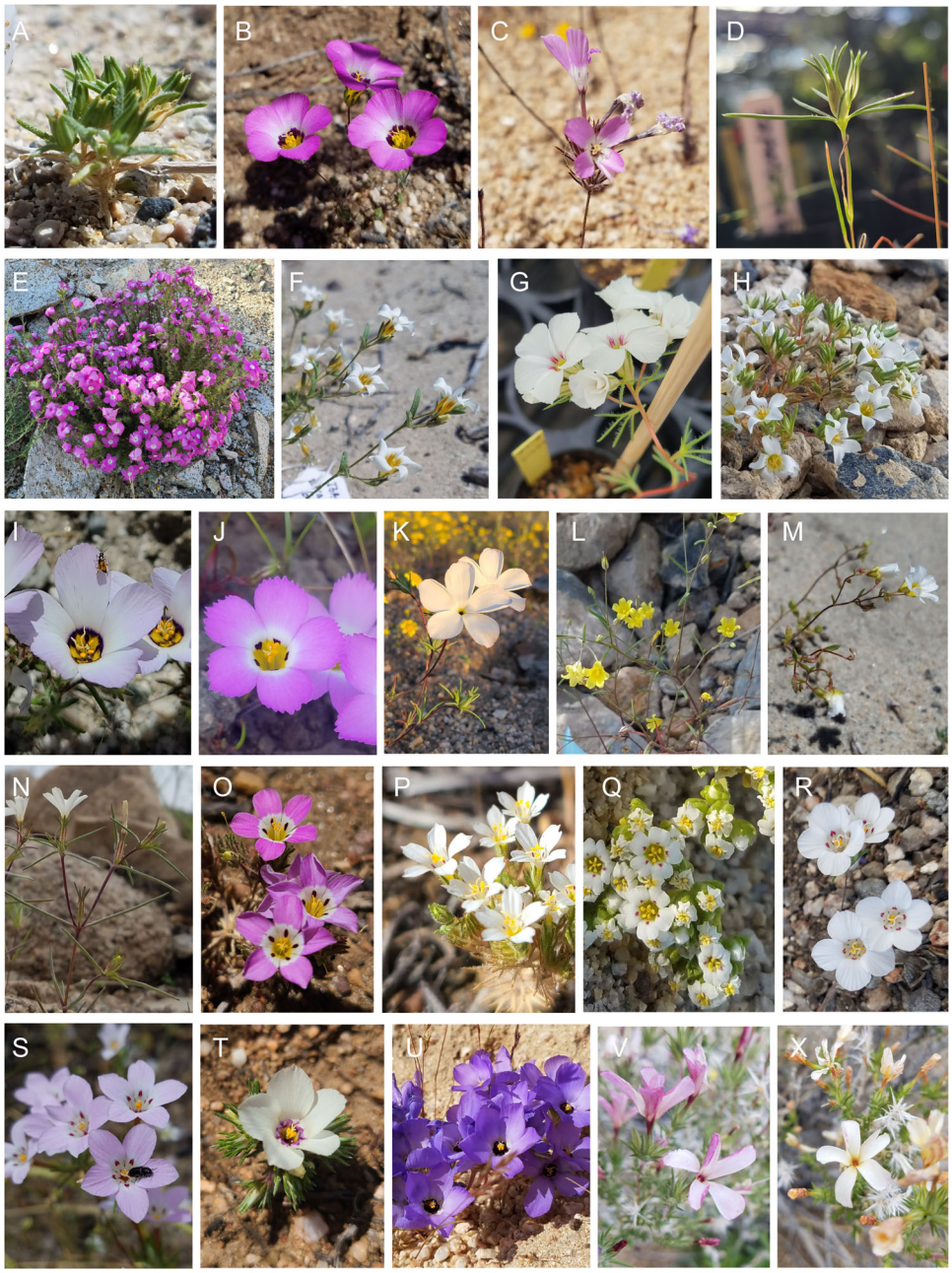

FIGURE 1 *Linanthus* encompasses extensive floral diversity, with many species that are polymorphic in color and floral markings. Species differ greatly in corolla tube depth and in pollinators observed to visit flowers. (A) *L. arenicola*. (B) *L. bellus*. (C) *L. bernardinus*. (D) *L. bigelovii*. (E) *L. californicus*. (F) *L. campanulatus*. (G) *L. concinnus*. (H) *L. demissus*. (I, J) *L. dianthiflorus*. (K) *L. dichotomus*. (L) *L. filiformis*. (M) *L. inyoensis*. (N) *L. jonesii*. (O, P) *L. killipii*. (Q) *L. maculatus*. (R, S) *L. orcuttii*. (T, U) *L. parryae*. (V, X) *L. pungens*. Photographs by I. G. Anghel.

including reduced pollen fertility and nectar evaporation (Borghi et al., 2019). Flowers opening during the cooler nights may be a strategy to avoid the adverse effects of higher temperatures. It is not known whether these desert adaptations have facilitated the diversification of *Linanthus*.

Linanthus, as currently recognized, includes 26 species (Porter and Johnson, 2000; Porter and Patterson, 2015) that have likely diversified in the Miocene, and is sister to a clade that includes the genera *Leptosiphon* Benth. and *Phlox* L. (Bell and Patterson, 2000). Previous phylogenetic work for 15 species using the nuclear ribosomal internal transcribed spacer (ITS) recovered low support for most relationships across the group (Bell et al., 1999). Another ITS and matK phylogeny that sampled 17 species did not provide a better resolution of the phylogenetic relationships between species (Bell and Patterson, 2000). A study using 14 nuclear loci for 22 species (with two samples per species) inferred a well-resolved phylogeny and found that eight species were not monophyletic (Landis, 2016). However, this study lacked complete taxonomic and broad geographic sampling per

species. A robust phylogeny is needed to provide a backbone for tackling interspecific and intraspecific evolutionary questions in this group.

Here, we present a phylogeny of *Linanthus* with complete species sampling and broad intraspecific sampling, using two types of genomic data: double-digest restriction-site associated DNA (ddRAD) sequencing (Peterson et al., 2012) and target capture (TC) of 353 nuclear angiosperm specific genes (Johnson et al., 2019). Our ddRAD data further include multiple samples from 17 species that co-occur locally with a congener to examine the potential for gene flow between sympatric individuals from different species. We used the resulting phylogenies to investigate the monophyly of species, evolutionary relationships, and patterns of floral evolution and life history shifts. We also used the ddRAD data to study population structure within selected clades to better understand existing species delimitations for several species. Lastly, we explored the geography of speciation across *Linanthus* to determine the most likely speciation mode in this remarkable desert radiation.

MATERIALS AND METHODS

Data collection and processing

We included representative samples from all currently recognized species of *Linanthus* (Moran, 1977; Porter and Johnson, 2000; Porter and Patterson, 2015) and outgroups from *Leptosiphon*, *L. breviculus* (A. Gray) J.M. Porter & L.A. Johnson, *L. chrysanthus* J.M. Porter & R. Patt, *L. lemmonnii* (A. Gray) J.M. Porter & L.A. Johnson, and *L. nuttallii* J.M. Porter & L.A. Johnson, as well as *Phlox stansburyi* (Torr.) A. Heller. To minimize potential bias introduced by misidentifications, we included only names at the species level, because many of the infraspecific taxa can be reliably identified only by experts with the aid of high-power magnification. We collected 83 individuals in the field, across California and Nevada, and sampled 125 individuals from eight herbaria across the western United States (Appendix S1). We stored the field-collected tissue in silica gel until storage at −20°C or in liquid nitrogen until storage at −80°C in the laboratory at UCLA (Los Angeles, California, USA). Some *Linanthus* species are minute; therefore, we carefully dissected all collections to ensure that we selected tissue from only one individual. In total, we included samples from 50 individuals that co-occurred with a congeneric species, as observed by us in the field or when herbarium specimens came from the exact same location. We included sympatric individuals for most sympatric species pairs. Seventeen of the species sampled co-occurred with a congener, with a total of 24 combinations of species pairs. See Appendix S1 for complete sample metadata.

We collected genomic data using two approaches, ddRAD sequencing and TC. RAD sequencing generates data that has been used to successfully address questions in both population genomics and phylogenetic relationships in closely related species (Rubin et al., 2012; Eaton et al., 2016; Jacobs et al., 2021). We used a double-digest RAD approach, which uses two enzymes to better recover the same fragments of DNA across samples and reduce sequencing costs (Peterson et al., 2012). In total, we sampled 192 individuals across the 26 currently recognized species and four outgroups in *Phlox* and *Leptosiphon*, at an average of seven individuals sampled per species, to represent the species' breadth of geographic range and phenotypic diversity. The number of individuals sampled per species ranged from 1 to 17, proportional to the species' geographic range size (Appendices S2 and S3).

For TC, we use the Angiosperms353 bait set (Johnson et al., 2019). This approach can sequence both exons and their flanking regions, providing informative phylogenetic data at various phylogenetic scales (Larridon et al., 2020; Slimp et al., 2021). We sampled 63 individuals across all species and four *Leptosiphon* outgroups. We dropped one species, *Linanthus uncialis* (Brandegee) Moran, because DNA extraction did not generate enough DNA. We sampled a range of one to five individuals per species, with an average of 2.5 individuals sampled per species of *Linanthus*.

DNA extraction, library preparation, and sequencing

We extracted genomic DNA using a modified CTAB technique (Doyle and Doyle, 1987; Cullings, 1992) that includes an additional incubation step to remove pectin. We prepared sequencing libraries at the Evolutionary Genomics Laboratory of the Museum of Vertebrate Zoology at UC Berkeley (Berkeley, California, USA).

ddRAD

For the RAD-seq libraries, we followed a modified version of the 3RAD protocol (Bayona-Vásquez et al., 2019) with adapters and indexing oligos provided by the Glenn lab at the University of Georgia (https://baddna.uga.edu/; Athens, Georgia, USA). We normalized a total of 192 samples to 125 ng of DNA in 10 μL volume and used this as input for a combined digestion and ligation reaction. We used XbaI andEcoRI-HF restriction enzymes to digest the genomic DNA, and the third enzyme (NheI-HF) served to digest adapter dimer produced during the ligation stages of the reaction. Adapter-ligated samples were purified using Solid-phase reversible immobilization (SPRI) beads (Rohland and Reich, 2012; Jolivet and Foley, 2015). We then amplified the adapter-ligated libraries with indexing oligos (Glenn lab) and the KAPA HiFi PCR Kit (Roche, Indianapolis, Indiana, USA) using 16 cycles of PCR, following this with another SPRI bead cleaning. We quantified the libraries, pooled them in equimolar amounts, and then size selected fragments at a length of 375–525 base pairs (bp) using a Pippin

Prep (Sage Science, Beverly, Massachusetts, USA) at the Functional Genomics Laboratory at UC Berkeley. We quantified the final size-selected library pool with the Qubit Fluorometer with dsDNA High Sensitivity Assay Kit (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and checked for quality using a Bioanalyzer DNA1000 kit (Agilent Technologies, Santa Clara, California, USA). We sequenced libraries on one lane of Illumina NovaSeq SP 150PE at Vincent J. Coates Genomics Sequencing Lab (GSL) at UC Berkeley for a total of 115 Gb of data at 20× coverage.

Target capture

We fragmented DNA from 63 samples, ranging from 175–1100 ng, to a target size of 350 bp using a qSonica sonicator (Newton, Connecticut, USA). We sonicated samples with high-quality DNA for 9 min at 40% amplitude with a 15 s on/15 s off pulse. Samples that had a variety of fragment sizes were sonicated for 3 min, while those that were already highly fragmented were not sonicated at all. We cleaned the sonicated DNA and size selected with a double-sided SPRI-bead cleaning using 0.525x for right-side selection and 0.675x for left-side selection. We prepared uniquely dual-indexed libraries for Illumina sequencing using the KAPA HyperPrep Kit (Roche) using one-quarter reagent volumes, with a custom-designed dual indexing oligo set developed at the Functional Genomics Laboratory at UC Berkeley. For capture hybridization we enriched 4 µg pools containing equal mass of seven to nine samples each at 62°C for 40 h using the Angiosperms353 gene set (Johnson et al., 2019) ordered from Arbor Biosciences ("Angiosperms 353 v1," catalog no. 308108.v5) and their myBaits Target Capture version 5 reagents and protocol (http://www.arborbiosci.com/mybaits-manual; Ann Arbor, Michigan, USA). We amplified enriched products with KAPA HiFi 2X HotStart ReadyMix PCR Kit for 10–12 cycles. We checked the resulting libraries for quality with an Agilent Bioanalyzer DNA1000 assay (Agilent Technologies, Santa Clara, California, USA) and quantified them with the Qubit Fluorometer with dsDNA High Sensitivity Assay Kit. We sequenced 150 paired-end reads at QB3 Genomics on a partial lane of an Illumina NovaSeq S4 for a total of 80 Gb of data.

Assembly

ddRAD

The sequencing facility demultiplexed the raw sequence reads and quality checked them with FASTQC v0.11.9 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). We used ipyrad to assemble the ddRAD reads using the reference genome for *L. parryae* (Eaton and Overcast, 2020; Anghel et al., 2022). We trimmed the variable length adapter sequences using cutadapt (Martin, 2011). We filtered low-quality base calls with a phred Qscore of 33 and allowed up to five low-quality bases per read. We set the minimum length of the reads after trimming to 35 bp, and a stricter filter for adapters at two. We set the clustering threshold within samples to 94% and between samples to 90%. We set the minimum number of samples per locus to four (see full parameters in Appendix S4). After filtering, we used 180 samples in the final assembly.

Target capture

Raw demultiplexed sequence reads were trimmed and quality filtered using Trimmomatic version 0.39 (Bolger et al., 2014). Specifically, we used a sliding window of five base pairs, cutting when the average base quality score was <20, removed trailing and leading low-quality base pairs when their quality was <20, and dropped reads whose length was <36. We assembled sequences for using HybPiper version 1.3 (Johnson et al., 2016) with the mega353 target file (McLay et al., 2021) using the "-bwa" option to match reads to loc. We assembled flanking regions using "intronerate.py" in HybPiper and identified paralogs using the "paralog_investigator.py" script. The output of HybPiper was exons, supercontigs, paralogs, and introns. Assemblies of multiple contigs containing >75% of the length of the reference protein were identified as paralogs. We removed sequences found to be paralogous and used supercontigs for further analysis because they contain both exons and flanking regions, which can provide information on both the conserved gene regions and the more variable non-coding regions. We used MAFFT version 7.505 to align supercontigs (Katoh et al., 2009). Trimal was used to clean alignments using a gap threshold of 0.9 (Capella-Gutierrez et al., 2009). We assessed the occupancy of samples in supercontig alignments with the "phyluce_align_get_only_loci_with_min_taxa" function in PHYLUCE (Faircloth, 2016) and kept only supercontigs found in all samples for further analyses.

Phylogenetic inference and species trees

ddRAD

We inferred phylogenetic trees using an alignment of concatenated filtered ddRAD loci with 98,538 variant sites. We used a maximum likelihood analysis approach in IQ-TREE with the GTR model accounting for ascertainment bias (Minh et al., 2020). The ascertainment bias model is used for single-nucleotide polymorphisms (SNPs) and other types of data that do not contain constant sites, to prevent overestimation of branch lengths (Lewis, 2001). We assessed support with 1000 ultrafast bootstrap replicates. We reconstructed phylogenetic trees with minimums of four, nine, 18, and 36 samples per locus to assess the role of missing data in recovering topologies. Within- and between-species relationships were consistent between these

data sets. Results of analyses with the data set using a minimum of four samples per locus are shown here. We inferred species trees with SVDQuartets (Chifman and Kubatko, 2014) in PAUP* (Swofford, 2002). To obtain branch lengths under the multispecies coalescent, we used the "qAge" command (Peng et al., 2022). As with the concatenated analyses, we estimated species trees with matrices including minimums of four, nine, 18, and 36 samples per locus. The species tree recovered with the matrix using a minimum of four samples per locus is presented below.

We also reassembled the data to include one sample per species to generate a species tree to use for downstream analyses where only one tip represents the species. To do this, we chose the sample within a species with the highest number of recovered loci. If the species was not monophyletic (see below), we chose the sample with the highest number of loci that belonged to the clade including most samples for such species. We built this tree in IQ-TREE with the GTR model accounting for ascertainment bias and assessed support with 1000 ultrafast bootstrap replicates.

Target capture

We inferred gene trees using the most conservative data set of 219 supercontigs with 100% occupancy using IQ-TREE with 1000 bootstrap replicates and the GTR + F default model of substitution (Minh et al., 2020). Then we used Newick utilities to collapse poorly supported gene tree nodes (ML bootstrap support < 0.2) into polytomies (Junier and Zdobnov, 2010). We used the gene trees as input for ASTRAL-III version 5.7.8 (Zhang et al., 2018) to infer a species tree for all 63 taxa as well as for the 24 species by assigning taxa to species a priori. We also estimated a phylogenetic tree with a concatenated matrix of the loci with 100% occupancy for the 63 taxa in IQ-TREE.

Population structure

To assess population structure within certain clades where the taxonomic assignment pointed to non-monophyletic species (see below), we used rmaverick version 1.0.5 (Verity and Nichols, 2016). We used only the RAD data for these analyses, with an average of seven samples per species. To prepare files for input into rmaverick, we processed the ipyrad Structure file following methods outlined at https://github.com/zapata-lab/ms_rhizophora/blob/main/analysis_code (Aburto-Oropeza et al., 2021). We processed the vcf file output from ipyrad using vcftools (Danecek et al., 2011). We removed non-biallelic sites and filtered genotypes called below a certain threshold across all individuals (--max-missing 0.25, 0.50, 0.75) for all the population structure analyses. We did this to compare the effect of missing data on the genetic structure output. We also kept only the center SNP from each locus to avoid effects of linked loci.

In rmaverick, we ran the MCMC sampling every 100 steps, with 10% burn-in, 1000 sampling iterations, and 20 rungs, and chose a K range from 1 to $n + 1$ where n is the number of taxonomic assignments in the focal clade. We ran these analyses for 0.25, 0.50, and 0.75 missing data sets, and with and without the admixture model. Comparing the data sets using 0.25, 0.50, and 0.75 missing data, we found little difference in the structure cluster assignments or values of K, so we present the 0.50 missing data sets here. Because rmaverick is more accurate than other population clustering programs at estimating the number of subpopulations, we chose to report only the value of K with the highest evidence shown by the largest posterior probability (Verity and Nichols, 2016). In all cases, the model without admixture was a better fit to the data.

Ancestral state reconstruction

To model the evolution of phenotypic traits, we coded all species for annuality/perenniality, day/night blooming, and lack/presence of corolla lobe anthocyanin pigment polymorphism (Appendix S1). We gathered this phenotypic data using the Jepson eflora, the Flora of North America, the monograph of *Linanthus*, and personal observations (Danforth, 1945; Patterson and Porter, 2021; R. W. Patterson and J. M. Porter, unpublished data). While corolla lobe polymorphic species have white, yellow, pink, lavender, purple, and/or peach forms, we focused only on the anthocyanin pigments. Therefore, we coded species with white and yellow corolla lobes as lacking anthocyanin pigments, and the species with pink, lavender, and purple corolla lobes as having anthocyanin pigments (Tanaka et al., 2008). All species with anthocyanin pigments were polymorphic, with populations or individuals within populations with white corollas.

Prior to inferring ancestral states, we calibrated the phylogeny built with ddRAD data and one sample per species to relative time using Penalized Likelihood with the "chronos" function in the R package phytools version 1.9-23 (Revell, 2012). Given uncertainty in the timing of the radiation, we applied an arbitrary age of 1 to the root and estimated relative divergence times under a discrete clock model with 10 distinct rate categories. We then fitted a Markov-k (Mk) model for discrete character evolution (Lewis, 2001). To do this, we used phytools and the function "fitMk" (Revell, 2012). We took an agnostic approach in choosing the model of evolution for the perenniality and night-blooming traits. For these two traits, using the Akaike Information Criterion (AIC), we selected the best model of evolution between equal rates, all rates different, and irreversible rates where loss of a trait is possible while its gain is not (Appendix S5). We used the irreversible model for perenniality and night-blooming because it had the lowest AIC scores (Appendix S5). For anthocyanin absence or polymorphism, we chose the "all rates different" model because of the likely different evolutionary rates between

loss and gain of flower pigmentation (Rausher, 2008), though we report stochastic character density maps with alternate models of evolution (Appendix S6). We simulated stochastic character maps onto the phylogenetic tree and 100 simulations (Huelsenbeck et al., 2003). Using the "densityMap" function, we visualized the stochastic mapping posterior probability density as a color gradient along the branches of the tree. Then, using the "density" function, we calculated the relative distribution of state changes from the stochastic mapping across the tree. To account for potential non-identifiability of rates of character evolution and diversification, we additionally inferred ancestral states for each character using a BiSSE model in the R package diversitree (Maddison et al., 2007; Fitzjohn, 2012) and tested for state-dependent speciation using analysis of variance.

Geography of speciation

To explore the geographic speciation history of *Linanthus*, we used data on species' geographic ranges and phylogenetic divergences (Barraclough and Vogler, 2000). Specifically, we estimated the relationship between phylogenetic distance and geographic range overlap between species pairs to evaluate evidence for a predominant geographic (allopatric vs. sympatric) mode of speciation in *Linanthus*. Under this approach, if allopatric speciation is the dominant process, geographic range overlap between young species pairs should increase from ~0% to random association as species pairs become more divergent over time. By contrast, if sympatric speciation is the dominant process, geographic range overlap should be ~100% between young species pairs but decrease over time among older pairs due to post-speciation geographic range changes (Losos and Glor, 2003; Fitzpatrick and Turelli, 2006; Skeels and Cardillo, 2019). To determine species' geographic ranges, we downloaded species location data from the Southwestern Environmental Information Network (SEINet, 2022) and the California Consortium of Herbaria (CCH, 2022), using only records backed by herbarium collections. We filtered the data to exclude records without precise GPS coordinates (fewer than two decimals precision) and clear outliers in terms of known species distributions. We used these records to create range maps for each species and estimate geographic range overlaps between species pairs using the R package hypervolume (Blonder et al., 2014) and the approach implemented at https://github.com/eliotmiller/ebirdr/blob/master/R/hypervolumeOverlaps.R (Miller et al., 2019). Because hypervolumes are geometric shapes with complex geometries, including the presence of holes (Blonder, 2016), they can describe species' geographic ranges more faithfully beyond simple ellipsoids or convex hulls. Areas of the species range perimeter with no occurrence points can be excluded, and ranges can have disjunctions (Blonder et al., 2014). Consequently, this method enables estimates of range overlap in heterogeneous environments with potential for small-scale allopatry. We calculated the Sørenson similarity index for each species pair and recorded the values in a similarity matrix, with a value of 0 representing no range overlap and a value of 1 representing complete overlap. We used the phylogeny built with ddRAD data and one sample per species to calculate the phylogenetic distance between all species pairs using the "cophenetic" function in the R package ape (Paradis et al., 2004). We coded species pairs belonging to the same clade as "within clade comparisons" and to different clades as "between clade comparisons." To test whether phylogenetic distance is a predictor of range overlap, we fitted a zero inflated beta regression model using the R package gamlss (Rigby and Stasinopoulos, 2005; Cribari-Neto and Zeileis, 2010). Beta regressions are well suited for our response variable, range overlap, which is bounded between 0 and 1 (Ferrari and Cribari-Neto, 2004). We ran three regressions using all of the data, comparisons of species in the same clade, and of species belonging to different clades.

RESULTS

Data processing

The ddRAD sequencing yielded an average of 2,440,318 reads per sample. The assembly with 180 samples in a minimum of four individuals yielded 36,861,279 bp, 3,131,821 SNPs, and a total of 165,943 loci with 95% missing data. An average of 7646 loci per sample were retained. The assembly using one sample per species yielded 4,047,394 bp, 459,834 SNPs, 17,751 loci in a minimum of four individuals, and an average of 4053 loci per sample. For the TC data, we selected a total of 219 contigs with 100% taxon occupancy for downstream analyses.

Phylogenetic inference

The ddRAD and the TC data sets both resulted in well-resolved phylogenies. In the concatenated ddRAD maximum likelihood phylogeny, most nodes had a bootstrap value >95, with only two nodes at 73 and 76 (Figure 2; Appendix S7). With an average of seven samples per species, 17 species resolved as monophyletic, while eight were nonmonophyletic (Figure 2). The TC phylogeny inferred using IQtree (Figure 3A) recovered nearly all of the same species relationships as in the phylogeny using ddRAD (Figure 2). There were two exceptions in the congruence of the ddRAD and TC phylogenies. One is in a clade recovered with ddRAD data that included *L. bellus*, *L. concinnus*, *L. dianthiflorus* (Benth.) Greene, *L. orcuttii*, and *L. uncialis*. The TC data resolves *L. dianthiflorus* and *L. concinnus*, and *L. bellus* and *L. orcuttii* as a grade, but no samples of *L. uncialis* were included in the TC analysis. The other incongruence was in a clade recovered with ddRAD data that included *L. demissus* (A. Gray) Greene, *L. bernardinus*, *L. killipii*, and *L. parryae*. Target capture data support *L. bernardinus*, *L. killipii*, and *L. parryae* as more

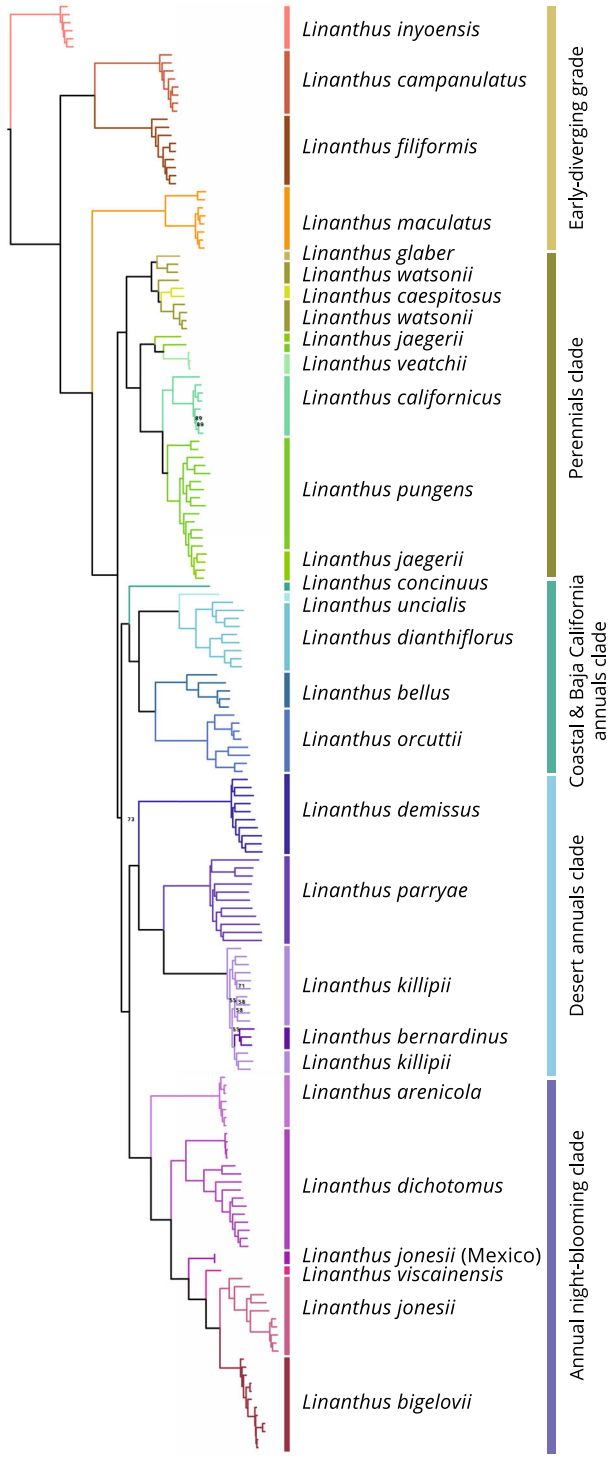

FIGURE 2 The phylogeny of *Linanthus* inferred using ddRAD data is well resolved and highly supported. A maximum likelihood phylogenetic tree was inferred in IQtree, including 180 samples for all species of *Linanthus*, using a concatenated matrix of ddRAD data and a minimum of four samples per locus. Values at nodes represent bootstrap support <90. The average of seven samples per species included shows that most of the species are monophyletic. The clades recovered share common morphological features, habitats, or habits.

closely related to the annual night-blooming clade than to *L. demissus*. The ddRAD and TC data represent two independent sources of data, and some incongruence between the evolutionary history of these two types of genetic data might be expected due to incomplete lineage sorting (Appendix S8). Given that our results were generally consistent across data sets, we are informally referring to several common clades and one grade as follows.

Early-diverging grade

This grade includes the following species previously described as *Gilia* section Giliastrum: *L. inyoensis* (I.M. Johnst.) J.M. Porter & L.A. Johnson, *L. campanulatus* (A. Gray) J.M. Porter & L.A. Johnson, *L. filiformis* (Parry Ex. A. Gray) J.M. Porter & L.A. Johnson, and *L. maculatus* (Parish) Milliken (Grant, 1959; Bell et al., 1999). While these species did not form a monophyletic group, the grade relationships are well supported and identical in the ddRAD and TC phylogenies (Figures 2 and 3). The relationships were also consistent with previous studies with limited sampling (Bell et al., 1999; Bell and Patterson, 2000). All five species in this grade were recovered as monophyletic (Figures 2 and 3).

Perennial clade

Both the ddRAD and TC analyses support the monophyly of a group including the perennial species *L. caespitosus* (Nutt.) J.M. Porter & L.A. Johnson, *L. californicus* (Hook. & Arn.) J.M. Porter & L.A. Johnson, *L. glaber* (R. Patt & Yoder-Will) J.M. Porter & L.A. Johnson, *L. jaegeri*, *L. pungens* (Torr.) J.M. Porter & L.A. Johnson, *L. veatchii* (Parry Ex. Greene) J.M. Porter & L.A. Johnson, and *L. watsonii* (A.Gray) Wherry (Figures 2 and 3). These species were previously recognized as the genus *Leptodactylon* Hook. & Arn. (Rydberg, 1906). Although a single origin of the perennial clade was also supported in a recent study (Landis, 2016), it was inconclusive in earlier phylogenetic analyses (Bell and Patterson, 2000). The perennial clade included two nested subclades. One subclade included *L. caespitosus*, *L. glaber*, and *L. watsonii*, all of which occur outside of California in the western United States (in Arizona, Colorado, Idaho, Montana, Nebraska, Nevada, Utah, Wyoming). The other subclade included *L. californicus*, *L. jaegeri*, *L. pungens*, and *L. veatchii*, all of which are restricted to California or Baja California, with the exception of the more widespread *L. pungens*, which is found throughout western North America.

Most species in the perennial clade were not recovered as monophyletic (Figures 2 and 3). Although *L. caespitosus* and *L. glaber* were recovered as monophyletic (however, note the small sample size), both

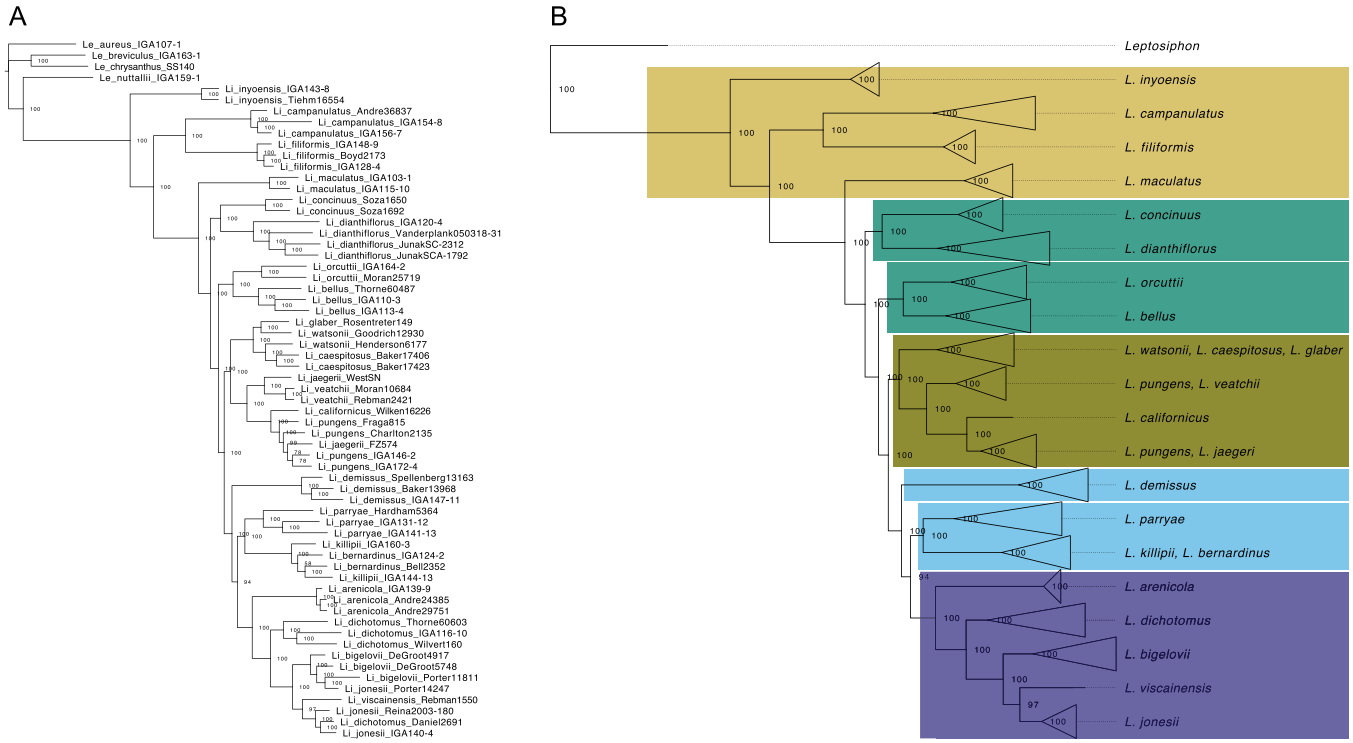

FIGURE 3 The phylogeny of *Linanthus* inferred using target capture data is well resolved and highly supported. This phylogeny is highly consistent with the phylogeny inferred using ddRAD data (see Figure 2). (A) Maximum likelihood phylogenetic tree inferred in IQtree, including 63 samples across 22 species of *Linanthus* and four outgroups, using a matrix of 219 concatenated target capture (Angiosperms353) loci with 100% occupancy. Values at nodes represent bootstrap support. Most species are recovered as monophyletic. (B) The same phylogenetic tree as in A, with collapsed species.

species were nested within a paraphyletic *L. watsonii*. In the other subclade, both *L. veatchii* and *L. californicus* were recovered as monophyletic, but they were nested within a clade that included the paraphyletic species *L. pungens* and *L. jaegeri*.

The rmaverick population structure analysis for the perennial clade showed the highest posterior probability for three distinct genetic clusters (Verity and Nichols, 2016; Figure 4). One cluster included all the samples of the species *L. watsonii* and *L. caespitosus*, which occur outside of California, another cluster included all the samples of the Baja California endemic *L. veatchii* and two samples of *L. pungens* from southern California and Baja California, and a third cluster included all samples of the species *L. pungens*, *L. californicus*, and *L. jaegeri* (Figure 4). The population structure results are consistent with the phylogenetic results showing a lack of genetic cohesiveness of the currently recognized taxonomic species in this clade.

Coastal California and Baja California annuals clade

The ddRAD phylogeny recovered a monophyletic group including *L. bellus*, *L. concinnus*, *L. dianthiflorus*, *L. orcuttii*, and *L. uncialis*, all of which occur in the coastal or higher-elevation areas of southern California and Baja California (Figures 2 and 3). A recent phylogenetic study also recovered this clade (Landis, 2016). While the monophyly of this group was not fully supported in the TC phylogeny, all species within this clade were closely related and formed a grade (Figure 3). The ddRAD analysis with multiple samples per species recovered all species as monophyletic, though we could include only one sample for *L. concinnus* and *L. uncialis* (Figure 2).

Desert annuals clade

In the ddRAD phylogeny, *L. bernardinus*, *L. demissus*, *L. killipii*, and *L. parryae* formed a clade sister to the annual night-blooming clade (see below; Figure 2). The TC species tree did not support the monophyly of the desert annual clade (Figure 3). Instead, this tree recovered *L. demissus* as sister to a clade with two subclades, one including the remaining species in the desert annual clade (*L. bernardinus*, *L. killipii*, and *L. parryae*) and a subclade including the annual night-blooming species (see below; Figure 3). The broad sampling in the ddRAD phylogeny showed that most species in the desert annual clade were recovered as monophyletic, with the exception of *L. bernardinus* and *L. killipii*, both of which formed a clade with intermixed samples.

Annual night-blooming clade

The annual night-blooming clade included the species *L. arenicola* (M.E. Jones) Jeps. & L.H. Bailey, *L. bigelovii*

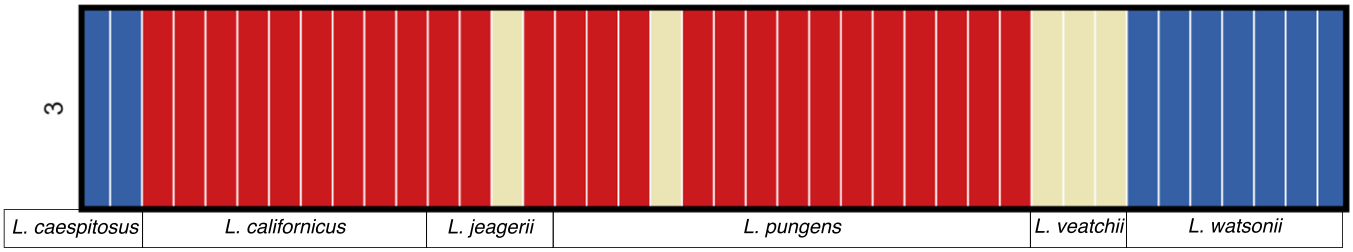

FIGURE 4 Genetic clusters did not match species taxonomy in the perennial *Linanthus* clade. $K = 3$ had the highest posterior probability, meaning that three genetic clusters were recovered for six taxonomic species. The red cluster included all samples identified as *L. californicus* and most samples identified as *L. pungens* and as *L. jaegeri*. The blue clade included all samples identified as *L. caespitosus* and *L. watsonii*. The tan group included all samples identified as *L. veatchii*, one sample identified as *L. jaegeri*, and one sample identified as *L. pungens*.

(A. Gray) Greene, *L. dichotomus* Benth., *L. jonesii* (A. Gray) Greene, *L. maricopensis* J.M. Porter & R. Patt, and *L. viscainensis* Moran. All species in this clade are night-bloomers, with the exception of *L. dichotomus*, in which populations of *L. dichotomus* subsp. *meridianus* flower during the day. This group was monophyletic in both the ddRAD and TC phylogenetic analyses (Figures 2 and 3). This corroborates the section *Linanthus* described by Grant, which included *L. bigelovii*, *L. dichotomus*, and *L. jonesii* (Grant, 1959). *Linanthus viscainensis* was later added to the section *Linanthus* on the basis of its morphological similarities to *L. arenicola* (Moran, 1977). This clade was also recovered in previous analyses with the addition of one accession of *L. filiformis* (Landis, 2016).

Most species in this clade were highly paraphyletic or polyphyletic (Figures 2 and 3). *Linanthus arenicola* was the only monophyletic species. Six of 11 samples of *L. jonesii* formed a clade, but it was nested within a more inclusive clade with several intermixed samples of *L. dichotomus* and *L. bigelovii*. Eight of 11 *L. bigelovii* samples formed a clade, but it was nested in a more inclusive clade with several samples of *L. dichotomus* and *L. jonesii*. The only sample of *L. viscainensis* included here was nested within a more inclusive clade including samples of *L. bigelovii*, *L. dichotomus*, and *L. jonesii*. Thirteen of 17 *L. dichotomus* samples and two samples of *L. bigelovii* formed a clade.

The population structure analysis in the annual night-blooming clade showed the highest posterior probability for four distinct genetic clusters (Figure 5). One cluster included the majority of samples identified as *L. bigelovii* (8/11), some samples identified as *L. jonesii* (5/11), and one sample identified as *L. dichotomus* (1/17). A second cluster included the majority of samples identified as *L. jonesii* (6/11), some samples identified as *L. dichotomus* (3/17), and one sample identified as *L. bigelovii* (1/11). A third cluster included only samples identified as *L. dichotomus* (11/17). A fourth cluster included some samples identified as *L. bigelovi* (2/11) and some samples identified as *L. dichotomus* (2/17) (Figure 5). We did not detect any geographic signal characterizing these genetic clusters.

Trait evolution and ancestral states

Our study traits do not seem to affect diversification rates. Perenniality, night-blooming, and petal color polymorphism did not have an effect on diversification rates ($P = 0.52$, $P = 0.51$, $P = 1$, respectively), with AIC values supporting the state independent diversification model in each trait. Mapping the distribution of phenotypic traits onto the ddRAD phylogeny shows that perenniality is clustered in one clade, night-blooming appears in the perennial clade and in the annual night-blooming clade, and petal color polymorphisms are dispersed across the phylogeny in three of the major clades (Figure 6). Stochastic character mapping showed that perenniality evolved once, with no reversals to annuality (Figure 7A). Night-blooming evolved three times: once in the annual night-blooming clade and twice in the perennial clade, all but two species of which have night-blooming populations (Figure 7B; Appendix S9). Corolla lobe anthocyanin pigment polymorphism may be ancestral in *Linanthus*, with several reversals to unpigmented corolla lobes (Figure 7C). The posterior probability distribution for the number of changes shows that gain of anthocyanin polymorphism may have occurred twice and loss of anthocyanins ten times (Appendix S9).

Geography of speciation

The majority of species pairs showed <50% overlap in geographic ranges regardless of phylogenetic distance (Figure 8). Nonetheless, species range overlap varied from 0 to 80%. Between species pairs, there was no effect of phylogenetic distance on geographic range overlap ($P = 0.93$; Figure 8). Young species pairs often had non-overlapping geographic ranges; however, species comparisons within clades did not show a relationship between phylogenetic distance and range overlap ($P = 0.90$). Between species belonging to different clades, phylogenetic distance was also not a good predictor of geographic range overlap ($P = 0.29$). Together, these results suggest that a single geographic mode of speciation (allopatric or sympatric) does not predominate among *Linanthus* species.

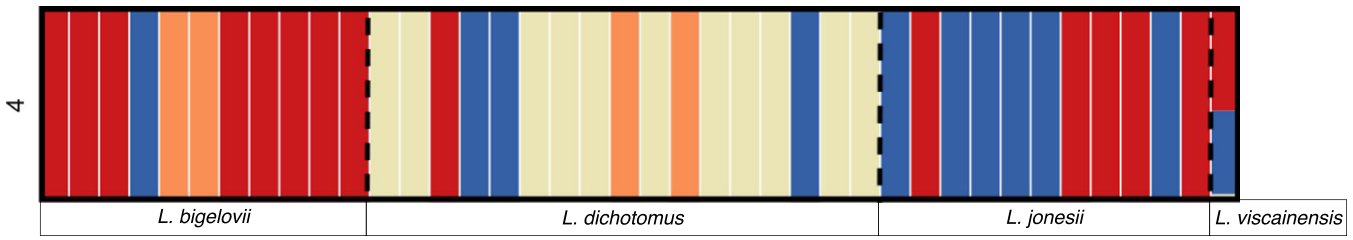

FIGURE 5 Genetic and taxonomic groups in the annual night-bloomers within *Linanthus* showed little congruence. $K = 4$ had the highest posterior probability, meaning that four genetic clusters were recovered for four taxonomic species. The red cluster included a majority of samples identified as *L. bigelovii*. The blue cluster included a majority of samples identified as *L. jonesii*. The tan cluster included only samples identified as *L. dichotomus*. The orange cluster included two samples identified as *L. bigelovi* and two samples identified as *L. dichotomus*. We excluded *L. arenicola* from this analysis because it was monophyletic.

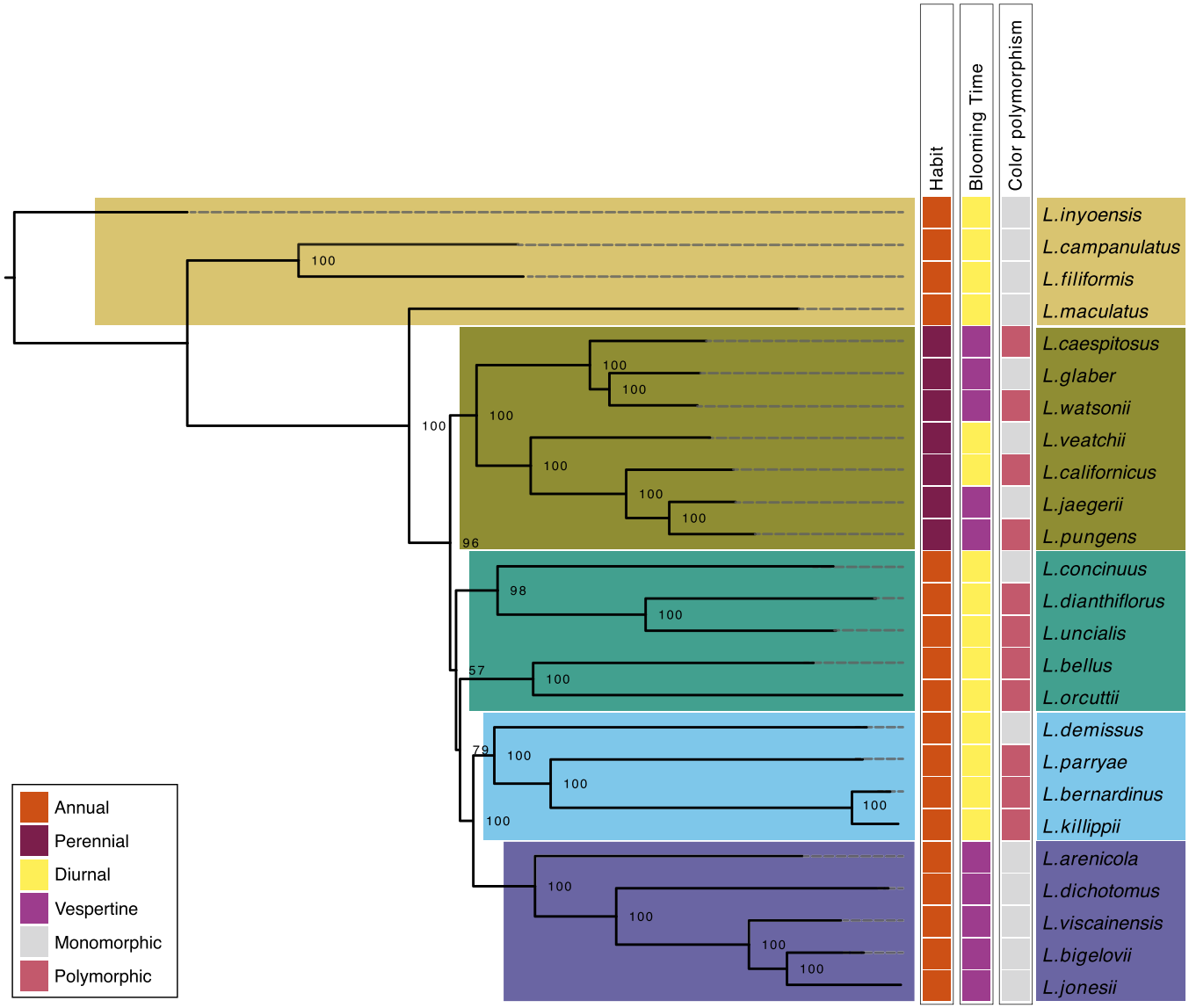

FIGURE 6 Phylogeny of *Linanthus* showing the distribution of three phenotypic traits. Left, *Linanthus* phylogeny using ddRAD data and one individual per species. Values represent bootstrap support. Right, diagram showing the character states distribution across species. Colors grouping taxa represent the clades defined in Figure 2.

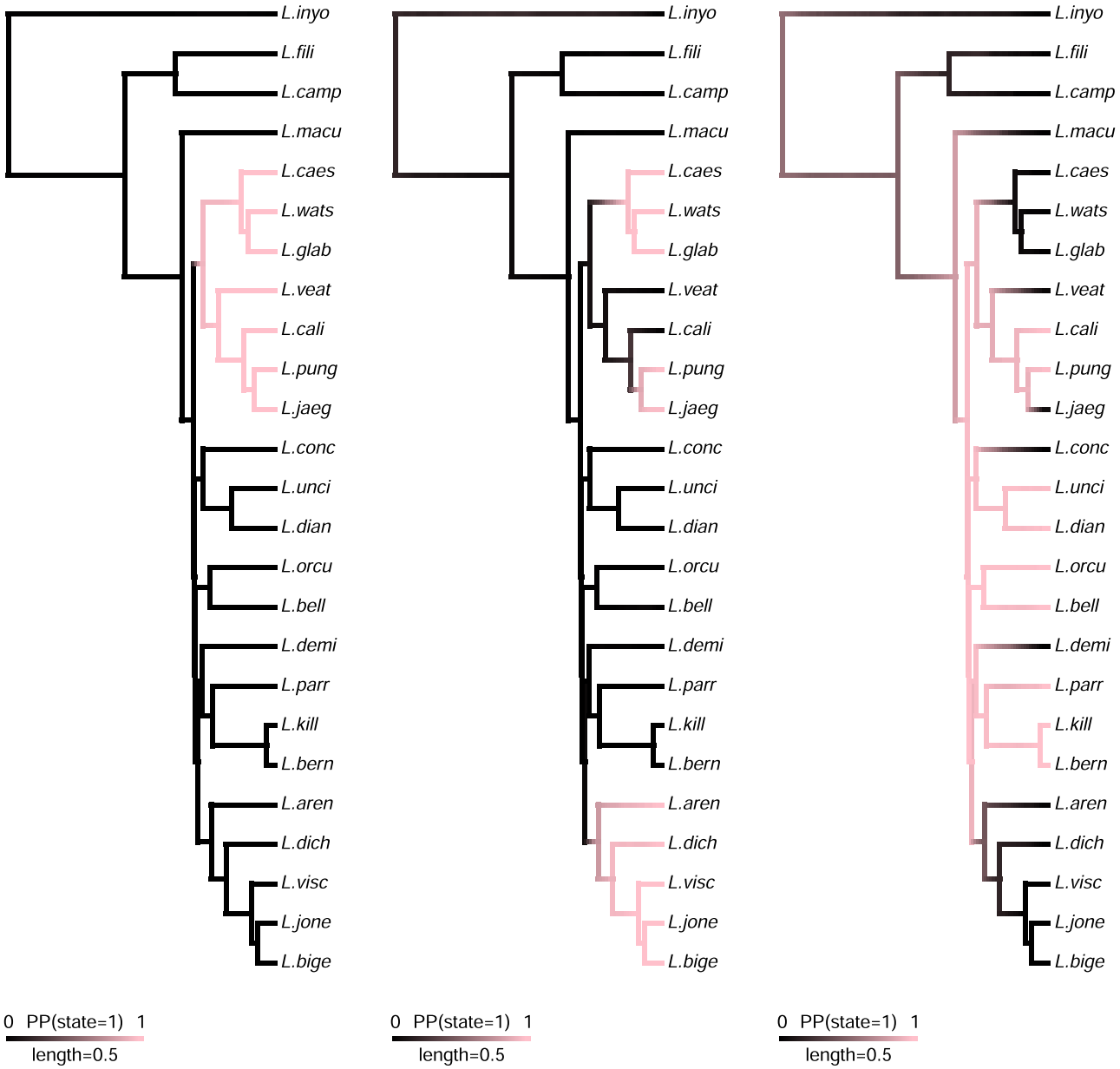

FIGURE 7 Across the *Linanthus* radiation, some phenotypic traits show stability while others are highly labile. Each panel shows stochastic character density maps for different traits. (A) Perenniality evolved once (1 = perenniality). (B) Night-blooming evolved three times: once in an annual clade and twice in the perennial clade (1 = nightblooming). (C) Corolla lobe anthocyanin pigment polymorphisms may be ancestral in *Linanthus*, with several reversals to unpigmented corolla lobes. The most dominant state across *Linanthus*'s evolutionary history is polymorphic (1 = polymorphism).

DISCUSSION

Phylogenetic relationships are well resolved and most species are monophyletic

Our phylogenetic analyses generated well-resolved phylogenies that included all of the currently recognized species in *Linanthus*. The TC phylogeny (Figure 3A) recovered nearly all of the same species relationships as in our concatenated ddRAD phylogeny (Figure 2), and they both were congruent with results from previous analyses with limited taxon and genetic sampling (Bell et al., 1999; Bell and Patterson, 2000; Landis, 2016). The species tree inferred with SVDQuartets from ddRAD data (Appendix S10) showed inconclusive species relationships due to very short branches, but this method is sensitive to large amounts of missing data common in RAD datasets (Nute et al., 2018). The species tree inferred with ASTRAL from TC data also recovered monophyly of most species, and most recent species relationships were congruent with the ddRAD data species trees (Appendices S10 and S11). Notably, our results showed that most species are

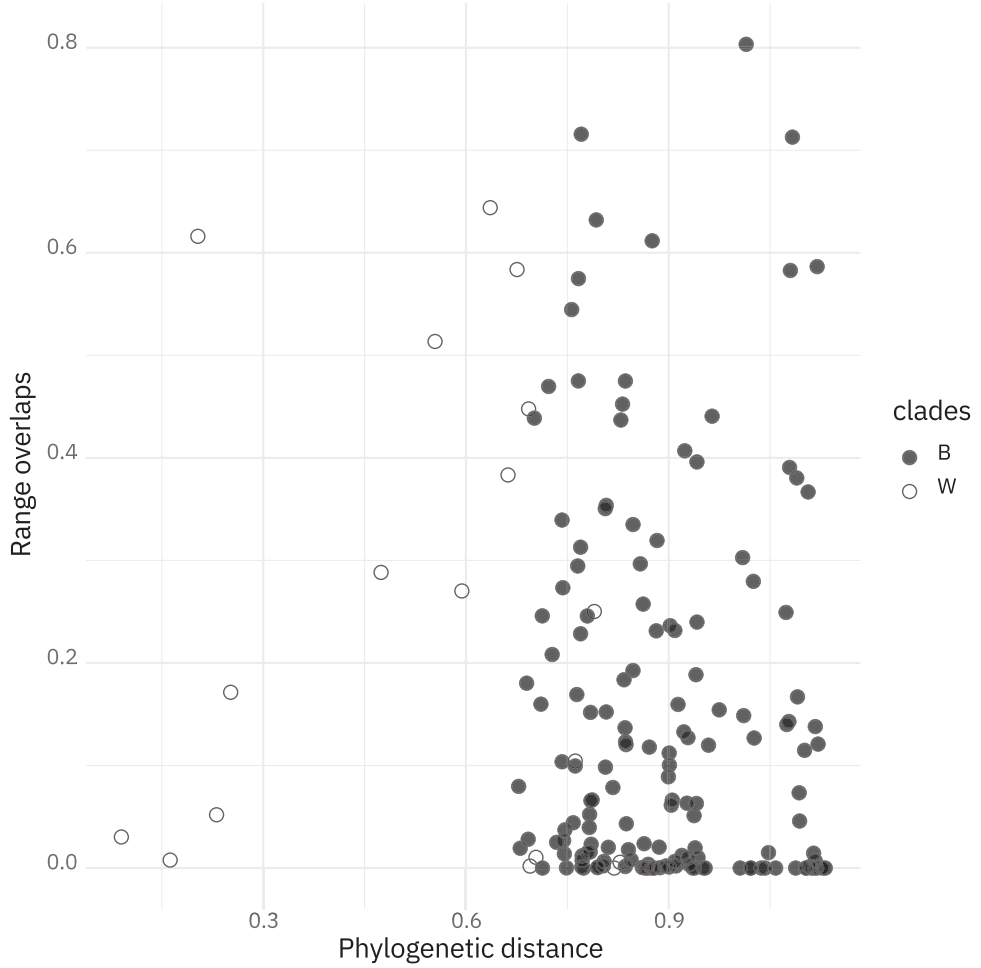

FIGURE 8 Phylogenetic distance does not predict range overlap of pairs of *Linanthus* species, suggesting that there is no predominant geographic mode of speciation. There is no effect of phylogenetic distance on range overlap between species either within clades (W) or between clades (B). Pairwise phylogenetic distance is based on the ddRAD phylogeny using one individual per species, and geographic range overlap is estimated as the overlap in hypervolumes corresponding to geographic ranges for each species.

monophyletic even when we included multiple samples from different species co-occurring in sympatry with ample opportunities for interspecific gene flow. This suggests that reproductive isolating barriers have likely evolved for most species in this recent desert radiation (see calibrated tree in Appendix S12). The lack of monophyly of some perennial and annual night-blooming species may be due to interspecific gene flow or to incomplete lineage sorting, or may indicate that current species limits are problematic. Porter and Patterson (2015) proposed a suite of infraspecific taxa within the perennial and night-blooming clades, reflecting the ongoing discovery of cryptic diversity in these groups. We did not explicitly test these taxonomic hypotheses, because of the difficulty of obtaining reliable identifications at subspecific rank for all herbarium specimens included in our study, but we suspect that the non-monophyly of species witnessed in our study may support the elevation of some of these taxa to specific rank. Further work, with more detailed study of herbarium material coupled with deeper taxon and genomic sampling and field experiments, could be useful for rigorously testing these various hypotheses.

Perennials evolved from annuals

The species in the perennial clade all share several synapomorphies, including the subshrub habit with a woody base, leaves that grow tightly in fascicles with sharp-tipped and filiform lobes or entire filiform leaves, and salverform corolla. These species occur in mountainous regions, consistent with the finding that alpine environments favor long-lived species that can persist in the colder winters (Billings and Mooney, 1968; Ricklefs and Renner, 1994).

Shifts from annual to perennial habit are considered rare across angiosperms, with perenniality thought to be the ancestral state in flowering plants and in most families and genera (Friedman, 2020). However, recent phylogenetic comparative work has shown examples of transitions from annual to perennial habit (Tank and Olmstead, 2008;

Soltis et al., 2013). In Polemoniaceae, transitions between annuality and perenniality in both directions are common, including in *Phlox* and *Leptosiphon*, which are two closely related genera of *Linanthus* (Barrett et al., 1996). Both the ddRAD and the TC phylogenies presented here support the evolution of the perennial species from an annual ancestor, and no reversals to annuality (Figure 7A). Thus, the transition to perenniality is unidirectional from annual ancestors. The evolution of perenniality may have implications for diversification and evolution in *Linanthus*. Perennial plants generally show a slower rate of molecular evolution than annual plants, which has been attributed to longer generation times (Andreasen and Baldwin, 2001) and larger population sizes (Bousquet et al., 1992). Our BiSSE analysis did not support an association between perenniality and diversification, but the BiSSE model requires a larger data set than was used here to produce statistically robust results (Davis et al., 2013).

Night-blooming evolved three times

We found that night-blooming has evolved three times in *Linanthus* (Figure 7B; Appendix S9), with one reversal to day blooming in *L. dichotomus*. Night-blooming has evolved multiple times across angiosperms but is relatively rare (Silberbauer-Gottsberger and Gottsberger, 1975; Grant, 1983). In Polemoniaceae, hawkmoth pollination seems to have evolved in three genera—*Linanthus*, *Ipomopsis* Michx., and *Phlox*—but only *Linanthus* has flowers that open exclusively at night (Grant, 1983). This suggests further specialization for hawkmoth pollination in *Linanthus* because there are no other pollinators available to visit the flowers at night.

Night-blooming is rare and may be a mechanism of allochronic speciation

Night-blooming species often occur in low densities and are spread across the landscape. This demographic pattern of night-bloomers is associated with dry habitats, where daytime anthesis can lead to excessive water loss through transpiration (Stebbins, 1970). Populations of the annual night-blooming *Linanthus* species show these demographic patterns. If night-blooming is indeed advantageous from a hydraulic perspective in *Linanthus*, the question is why more species have not evolved this trait given their widespread distribution in the hot, dry deserts. Detailed physiological and genomic studies could shed light on the biological mechanisms underpinning this phenotype. Alternatively, it is plausible that shifts in flowering time have evolved as prezygotic isolating barriers and are linked to a more complex phenotype involved in pollination and possibly allochronic speciation (Taylor and Friesen, 2017). Night-blooming flowers are usually white, pale yellow, or pale pink, with strong scents emitted at night, and with long corolla tubes (Grant, 1983; Knudsen and Tollsten, 1993). These traits are present in all the annual night-blooming *Linanthus*, with variation within and between species, making it a great system to study the genetic basis of this complex phenotype likely involved in speciation. For instance, this system can be used to study which suite of traits is specialized for moth pollination and which attracts a wider variety of pollinators. Future work can also investigate how the loss of petal pigment and specialization on night pollinators may be leading the night-bloomers to a potential evolutionary dead end (Tripp and Manos, 2008).

Our results show that populations of species in the night-blooming clade co-occur in sympatry with populations of species in other clades. Shift in blooming time could work as a temporal isolating barrier when no geographic barriers limit interspecific gene flow. Because night-blooming species co-occur with species that have diverged at different times (Figures 2 and 8), it is unclear whether night-blooming is the cause of reproductive isolation or evolved later in secondary sympatry as a reinforcement mechanism (Taylor and Friesen, 2017).

Selfing may be a mechanism of reproductive isolation between closely related species

Outside of night anthesis, the *Linanthus* species in the annual night-blooming clade are united morphologically by their white or yellow corolla color, short or nonexistent pedicel, and a cylindrical or urn-shaped calyx, with the membrane separating sepals that are wider than the lobes. The stamens are included in the corolla tube, with the pistil at or below the stamens. This combination of traits is often found in flowers that can self-pollinate (Ushimaru and Nakata, 2002). Indeed, most *Linanthus* species in the annual night-blooming clade can set seed without the corolla emerging out of the calyx (I. G. Anghel et al., personal observation). Self-compatibility is common in other hawkmoth-pollinated flowers (Grant, 1983). For instance, *L. arenicola* and some populations of *L. jonesii* and *L. bigelovii* have very short floral tubes, making hawkmoth pollination less likely. Therefore, it is possible that some populations of these species are autogamous. Further study is needed to determine whether these populations are facultative selfers, growing smaller autogamous flowers in years when conditions are not favorable. The gradient from selfing to outcrossing via hawkmoth pollination in this night-blooming clade provides the opportunity to investigate different strategies for reproductive isolation in closely related species. Similar trends in selfing features are present in *Leptosiphon*, a genus closely related to *Linanthus* (Goodwillie, 1999; Goodwillie and Stiller, 2001; Goodwillie and Ness, 2005). Selfing in *Leptosiphon* may have evolved to ensure reproductive success in environments with inconsistent pollinator visitation and with range expansion to drier habitats that decouple pollinator emergence from flowering time

(Goodwillie, 1997). Whether similar mechanisms operate and evolve independently in *Linanthus* is unknown.

Scent may act as a reproductive isolating mechanism, especially in night-blooming species

Night-blooming species use scent to attract pollinators because visual cues are less effective at night (Raguso and Willis, 2002). Many night-blooming flowers emit a heavy, sweet, musky scent, which is an olfactory attractor of noctuid moths (Raguso et al., 2003). *Linanthus dichotomus*, the showiest species in the night-blooming clade, emits scents associated with moth attraction, and different chemical profiles between the day- and night-blooming subspecies (Chess et al., 2008). Analyzing the chemical profiles of *Linanthus* species will improve our understanding of the evolution of the hawkmoth pollination syndrome. Further chemical ecology studies are needed to explore how floral scent works as a potential isolating mechanism in *Linanthus*.

Color polymorphism may be ancestral

Flower anthocyanin polymorphisms in *Linanthus* may have evolved early in the history of the genus, in the clade that includes most species in the genus except for *L. inyoensis*, *L. filiformis*, and *L. campanulatus* (Figure 7C). Although these three early-diverging species do not exhibit color polymorphisms, several other genera in Polemoniaceae have color-polymorphic species (Schemske and Bierzychudek, 2007). Anthocyanin pigmentation likely evolved in the ancestor of *Linanthus* or even earlier in the history of Polemoniaceae (Landis et al., 2018), but the evolution of anthocyanin-based polymorphism has not been studied in this family. A noteworthy feature of the anthocyanin polymorphisms in *Linanthus* is that no species are pigmented and monomorphic (i.e., there are no species with only pink or purple petals), and all pigmented species also have a non-pigmented morph (i.e., a species with pink petals also has individuals with white petals). The pigmented monomorphic trait may be a hidden state in the evolutionary history of *Linanthus*, but it is not possible to infer from our analyses whether pigmented-only species existed in the genus. Despite this, the prevalence of polymorphisms in the genus indicates that polymorphisms are a shared derived trait in at least some of the *Linanthus* lineages. It is expected that polymorphisms are rarely retained across speciation events, because the genetic variation responsible for the polymorphism as well as the disruptive selective pressure maintaining the polymorphism must persist through time (Jamie and Meier, 2020). While polymorphisms might be a precursor to speciation, polymorphisms that persist through a speciation event must be a result of different forces than those driving speciation (Gray and McKinnon, 2007). Small-scale habitat differences or temporal fluctuations are possible disruptive forces that do not directly lead to reproductive isolation, and have been credited with maintaining the color polymorphism in *L. parryae* (Schemske and Bierzychudek, 2001, 2007). The prevalence of color polymorphisms across the *Linanthus* radiation and the potential for its mechanism to be maintained through speciation events make this genus an ideal subject for investigating the dynamics between speciation and polymorphisms.

Evolutionary transitions in flower pigmentation are common across the angiosperm phylogeny (Rausher, 2008; Smith and Goldberg, 2015) and in Polemoniaceae (Landis et al., 2018), but the macroevolution of color polymorphisms remains understudied. There are few documented examples of multiple closely related species sharing flower color polymorphisms, with a few exceptions in *Antirrhinum* L. (Lamiaceae) and *Protea* L. (Proteacea; Carlson and Holsinger, 2015; Ellis and Field, 2016). In California-occurring Polemoniaceae, 36% of the species exhibit flower color polymorphism (Schemske and Bierzychudek, 2007), in *Ipomopsis*, *Leptosiphon*, *Linanthus*, and *Phlox*. In *Linanthus*, 44% of the species are polymorphic across populations and 28% within populations (Patterson and Porter, 2021; I. G. Anghel, personal observation). The pink, lavender, and purple color variation in *Linanthus* are likely anthocyanin-based pigments (Tanaka et al., 2008). The inability to synthesize anthocyanin pigments has been suggested as a mechanism for producing the white-corolla individuals (Warren and Mackenzie, 2001), but this mechanism remains to be tested in *Linanthus*. The loss of the color polymorphism seems to have occurred several times in the history of *Linanthus* (Appendix S7), indicating that color is likely more easily lost than gained (Rausher, 2008).

Many of the species in *Linanthus* exhibit continuous color variation. For example, the iconic *L. parryae* has been coded as dimorphic in the classic papers investigating the evolution of such polymorphism (Epling and Dobzhansky, 1942; Wright, 1943; Schemske and Bierzychudek, 2007). However, observations in the field across the flowering season indicate that individuals range in intensity of flower color, both within and between populations. In addition, the color of the purple morphs appears to fade with heat or time since anthesis (I. G. Anghel et al., personal observation). These characteristics point to a potentially more complex mechanism for the genetics and plasticity of flower color than previously suspected in *L. parryae* and the other polymorphic species in *Linanthus*.

While anthocyanin-produced pigment influences pollinator attraction, it also has survival functions, including deterrence of herbivores and response to abiotic stressors such as UV exposure and drought (Strauss and Whittall, 2006). Across angiosperms, flower color polymorphisms are most common in heterogeneous environments, like the Mediterranean or high-elevation biomes (Sapir et al., 2021). Experimental studies have shown that in white and pink/purple flower polymorphic species, the pink/purple morphs show greater tolerance to drought and heat stress (Warren and Mackenzie, 2001; Coberly and Rausher, 2003; Vaidya et al., 2018; Dittmar and Schemske, 2023). *Linanthus parryae* pollinator visitation did not differ between two color morphs, but reproductive success was

higher in white-flowered individuals in wetter years, and in purple-flowered individuals in dry years (Schemske and Bierzychudek, 2001). Color-morph differences correlated to rainfall abundance point to a potential pleiotropic effect of flower color and environmental factors in *L. parryae* (Schemske and Bierzychudek, 2007) but can also point to potential linked genes producing these seemingly unrelated phenotypes (Rausher, 2008). The white and pink/purple polymorphism commonly found in *Linanthus* may be an important adaptation that allowed plants to tolerate the variable environmental conditions in xeric areas of western North America and might have been a precursor to its diversification in this geographic area with high levels of UV radiation, drought, and heat. The prevalence and the potential ancestral origin of color polymorphisms in *Linanthus* open promising future avenues for studying the mechanisms that maintain polymorphisms within species and across speciation events.

No predominant mode of speciation, but range overlap in sister species indicates parapatric speciation

We did not find a relationship between phylogenetic distance and species range overlap; thus, we did not detect a prevalent geographic mode of speciation in *Linanthus*. Although current species distributions are not necessarily reflective of the species range at the time of speciation (Losos and Glor, 2003), some signature of this pattern is likely present in the current geography of closely related species (Barraclough and Vogler, 2000). Allopatric speciation has been considered the dominant mode of speciation (Mayr, 1959; Coyne and Orr, 2004). If this is the case, we would expect geographic range overlap to increase over time, with sympatry arising secondarily after populations have accumulated enough differences and become reproductively isolated. However, we did not detect signals of such a pattern (Figure 8). Recent studies have suggested that plant sympatric speciation is more prevalent than other geographic modes of speciation (Skeels and Cardillo, 2019). Under this scenario, we would expect geographic range overlap to decrease as a result of post-speciation geographic range changes (Losos and Glor, 2003). Yet our findings for *Linanthus* were not consistent with this pattern (Figure 8). Taken together, our results indicate that multiple speciation mechanisms are at play in the *Linanthus* radiation and that no single mode predominates. For instance, we found that some young species pairs show considerable geographic range overlap, while others overlap minimally. This difference suggests that in some cases, speciation could have happened in parapatry, with opportunities for homogenizing gene flow; and that in other cases, speciation likely happened in isolation in allopatry. Further geographic sampling at the landscape scale, combined with detailed studies in regions of geographic overlap using simulations and population genomic approaches, will be essential to confidently discern speciation modes.

Recent studies of California plants suggest that the patterns of geographic range overlap, geographic range asymmetry, and time since divergence between species pairs were not consistent with allopatry as the dominant mode of geographic speciation (Anacker and Strauss, 2014; Grossenbacher et al., 2014; Christie and Strauss, 2019). However, these studies estimated range overlap using a different approach than the one employed here. We used overlap of hypervolumes that account for density of points and holes in the distribution of species occurrences, and holes in overlap from calculations of sympatry (Blonder et al., 2014). This method is sensitive to overlap at a finer scale, potentially excluding areas of overlap where the potential for gene flow between species pairs is low. Fine-scale geographic partitioning or "micro-allopatry" may be common in California native plant species (Anacker and Strauss, 2014; Grossenbacher et al., 2014). These previous studies used the overlap between polygons formed by occurrence points, and the difference in range overlap calculation methods could explain the overall discrepancies between our findings and the ones reported in those studies.

The expected pattern of increasing overlap with increasing phylogenetic distance under allopatric speciation may be apparent in a genus that diversified across a more homogeneous environment, with fewer barriers to range expansion and contraction. However, this pattern may not emerge in *Linanthus* even if allopatric speciation is common, because some species pairs never experience secondary sympatry (Figure 8). The lack of geographic overlap could result from specialization on certain habitats or pollinators, the inability to expand geographic ranges, or the existence of a heterogeneous environment with barriers to dispersal. These are patterns we commonly see in the ecology and geography of *Linanthus*. Testing these hypotheses will lay the groundwork for future studies exploring speciation mechanisms in *Linanthus* across the harsh North American deserts.

CONCLUSIONS

We conducted a comprehensive phylogenetic study of the genus *Linanthus*, a diverse radiation of mostly annual plants from the biodiverse deserts of southwestern North America. Our phylogenies are the first to include complete species sampling and extensive intraspecific sampling. This approach allowed us to explore the monophyly of species and taxon relationships with increased rigor. Most species resolve as monophyletic despite rampant local sympatry and range overlap, suggesting the presence of strong isolating ecological barriers. The species within the annual night-blooming clade and the perennial clade are not monophyletic, and closer taxonomic and population-level studies are needed to untangle the evolutionary patterns of those species. Although we do not detect a strong signal for a predominant geographic mode of speciation, most species show some overlap in geographic range regardless of time since divergence. This suggests that some species could have

evolved in parapatry, likely in the face of gene flow, while others likely evolved in isolation and never attained secondary sympatry. The strategies that *Linanthus* species have evolved to deal with desert living—including flower color polymorphisms, facultative selfing, night-blooming, and annuality—make it a rich system to study how plants can adapt to a drier world.

AUTHOR CONTRIBUTIONS

I.G.A. and F.Z. designed the study. I.G.A. conducted the research, analyzed the data, and wrote the manuscript. L.L.S. and I.G.A. prepared sequencing libraries. I.L.M. assembled the target capture sequence data and advised on analyses. All authors contributed to the editing and revision of the manuscript.

ACKNOWLEDGMENTS

The authors thank J. Corpus-Sapida and S. Payne for helping with DNA extractions. We are grateful to the curators of ASU, BRY, CAS, RSA, SBBG, SD, UCLA, and UNR herbaria for permission to sample from herbarium specimens, especially to M. Nazaire, who provided a bulk of the herbarium samples used in this study and E. Makings, who provided a key determination. We thank R. Patterson for inspiring discussions about *Linanthus* taxonomy and evolution, and M. Porter for taxonomic context and training on the pectinase DNA extraction protocols. We are grateful to C. Henriquez for DNA extraction and library preparation training and support. We thank S. Singhal for discussion about alternate analyses when we seemed to hit a dead end, and I. Overcast for help troubleshooting our ipyrad assemblies. We thank D. Bell, M. Borg, and S. Serkanic for providing supplemental samples for this project. We are very appreciative to K. Kay, V. Sork, and N. Kraft for providing comments that strengthened the arguments of this manuscript. We thank two anonymous reviewers for their valuable comments that improved the quality of the paper. Funding for this work was generously provided to I.G.A. by the Ecology and Evolutionary Biology Department at UCLA, the UCLA La Kretz Center and Stunt Ranch "Graduate Research Grant," the Botanical Society of America "Bill Dahl Graduate Student Research Award," the American Society of Plant Taxonomists "Graduate Student Research Grant," the California Native Plant Society "Hardman Native Plant Research Award," and F.Z.'s start-up funding. I.L.M. was funded by NSF DBI no. 2209393. This work is part of I.G.A.'s Ph.D. dissertation at UCLA.

DATA AVAILABILITY STATEMENT

All sequences generated for this work are available on NCBI Sequence Read Archive (SRA) under the BioProject PRJNA1142084: https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA1142084. Appendices, alignments, trees, character matrix and scripts will be available on github.com/ioanaanghel/Linanthus_phylogeny.

ORCID

Ioana G. Anghel http://orcid.org/0000-0002-1454-8718
Lydia L. Smith http://orcid.org/0009-0005-8556-0246
Isaac H. Lichter-Marck http://orcid.org/0000-0003-3575-6003
Felipe Zapata http://orcid.org/0000-0002-9386-0573

REFERENCES

Aburto-Oropeza, O., C. M. Burelo-Ramos, E. Ezcurra, P. Ezcurra, C. L. Henriquez, S. E. Vanderplank, and F. Zapata. 2021. Relict inland mangrove ecosystem reveals last interglacial sea levels. *Proceedings of the National Academy of Sciences* 118: 7–11.

Anacker, B. L., and S. Y. Strauss. 2014. The geography and ecology of plant speciation: range overlap and niche divergence in sister species. *Proceedings of the Royal Society B: Biological Sciences* 281: 1–9.

Andreasen, K., and B. G. Baldwin. 2001. Unequal evolutionary rates between annual and perennial lineages of checker mallows (*Sidalcea*, Malvaceae): Evidence from 18S-26S rDNA internal and external transcribed spacers. *Molecular Biology and Evolution* 18: 936–944.

Anghel, I. G., S. J. Jacobs, M. Escalona, M. P. A. Marimuthu, C. W. Fairbairn, E. Beraut, O. Nguyen, et al. 2022. Reference genome of the color polymorphic desert annual plant sandblossoms, *Linanthus parryae*. *Journal of Heredity* 113: 712–721.

Axelrod, D. I. 1972. Edaphic aridity as a factor in angiosperm evolution. *American Naturalist* 106: 311–320.

Azani, N., A. Bruneau, M. F. Wojciechowski, and S. Zarre. 2019. Miocene climate change as a driving force for multiple origins of annual species in *Astragalus* (Fabaceae, Papilionoideae). *Molecular Phylogenetics and Evolution* 137: 210–221.

Barraclough T. G., and A. P. Vogler. 2000. Detecting the geographical pattern of speciation from species-level phylogenies. *American Naturalist* 155: 419–434.

Barrett, S. C. H., L. D. Harder, and A. C. Worley. 1996. The comparative biology of pollination and mating in flowering plants. *Philosophical Transactions: Biological Sciences* 351: 1271–1280.

Bayona-Vásquez, N. J., T. C. Glenn, T. J. Kieran, T. W. Pierson, S. L. Hoffberg, P. A. Scott, K. E. Bentley, et al. 2019. Adapterama III: Quadruple-indexed, double/triple-enzyme RADseq libraries (2RAD/3RAD). *PeerJ* 7: e7724.

Bell, C. D., R. Patterson, and L. A. Hamilton. 1999. Sectional integrity in *Linanthus* (Polemoniaceae): A molecular phylogeny of section Dianthoides. *Systematic Botany* 24: 632–644.

Bell, C. D., and R. Patterson. 2000. Molecular phylogeny and biogeography of *Linanthus* (Polemoniaceae). *American Journal of Botany* 87: 1857–1870.

Billings, W. D., and H. A. Mooney. 1968. The ecology of arctic and alpine plants. *Biological Reviews* 43: 481–529.

Blonder, B. 2016. Do hypervolumes have holes? *American Naturalist* 187: E93–E105.

Blonder, B., C. Lamanna, C. Violle, and B. J. Enquist. 2014. The n-dimensional hypervolume. *Global Ecology and Biogeography* 23: 595–609.

Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120.

Borghi, M., L. Perez de Souza, T. Yoshida, and A. R. Fernie. 2019. Flowers and climate change: a metabolic perspective. *New Phytologist* 224: 1425–1441.

Bousquet J., S. H. Strauss, A. H. Doerksen, and R.A. Price. 1992. Extensive variation in evolutionary rate of rbcL gene sequences. *Proceedings of the National Academy of Sciences* 89: 7844–7848.

Cacho, I. N., A. Millie Burrell, A. E. Pepper, and S. Y. Strauss. 2014. Novel nuclear markers inform the systematics and the evolution of serpentine use in *Streptanthus* and allies (Thelypodieae, Brassicaceae). *Molecular Phylogenetics and Evolution* 72: 71–81.

California Native Plant Society, Rare Plant Program. 2023. Rare Plant Inventory (online edition, v9.5). Website: https://www.rareplants.cnps.org [accessed 26 April 2023].

Capella-Gutierrez, S., J. M. Silla-Martinez, and T. Gabaldon. 2009. trimAl: a tool for automated trimming in large-scale phylogenetic analyses. *Bioinformatics* 25: 1972–1973.

Carlson, J. E., and K. E. Holsinger. 2015. Extrapolating from local ecological processes to genus-wide patterns in colour polymorphism in South African *Protea*. *Proceedings of the Royal Society B: Biological Sciences* 282: 20150583.

Chess, S. K. R., R. A. Raguso, and G. LeBuhn. 2008. Geographic divergence in floral morphology and scent in *Linanthus dichotomus* (Polemoniaceae). *American Journal of Botany* 95: 1652–1659.

Chifman, J., and L. Kubatko. 2014. Quartet inference from SNP data under the coalescent model. *Bioinformatics* 30: 3317–3324.

Christie, K., and S. Y. Strauss. 2019. Reproductive isolation and the maintenance of species boundaries in two serpentine endemic Jewelflowers. *Evolution* 73: 1375–1391.

Coberly, L. C., and M. D. Rausher. 2003. Analysis of a chalcone synthase mutant in *Ipomoea purpurea* reveals a novel function for flavonoids: Amelioration of heat stress. *Molecular Ecology* 12: 1113–1124.

Consortium of California Herbaria (CCH). 2022. Data provided by the participants of the Consortium of California Herbaria. Website: ucjeps.berkeley.edu/consortium [accessed 12 October 2022].

Coyne, J. A., and H. A Orr. 2004. Speciation. Sinauer, Sunderland.

Cribari-Neto, F., and A. Zeileis. 2010. Beta Regression in R. *Journal of Statistical Software* 34: 1–24.

Cullings, K. W. 1992. Design and testing of a plant-specific PCR primer for ecological and evolutionary studies. *Molecular Ecology* 1: 233–240.

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, et al. 2011. The variant call format and VCFtools. *Bioinformatics* 27: 2156–2158.

Danforth, F. W. 1945. The genus *Linanthus*: a concept of the genus and a taxonomic study of its species. Stanford University, Masters Thesis.

Davis, M. P., P. E. Midford, and W. Maddison. 2013. Exploring power and parameter estimation of the BiSSE method for analyzing species diversification. *BMC Evolutionary Biology* 13: 38.

Dittmar, E. L., and D. W. Schemske. 2023. Temporal variation in selection influences microgeographic local adaptation. *American Naturalist* 202: 471–485.

Doyle, J. J., and J. L. Doyle. 1987. A rapid DNA Isolation procedure from small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11–15.

Eaton, D. A. R., and I. Overcast. 2020. ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics* 36: 2592–2594.

Eaton, D. A. R., E. L. Spriggs, B. Park, and M. J. Donoghue. 2016. Misconceptions on missing data in RAD-seq phylogenetics with a deep-scale example from flowering plants. *Systematic Biology* 66: 399–412.

Ellis, T. J., and D. L. Field. 2016. Repeated gains in yellow and anthocyanin pigmentation in flower colour transitions in the *Antirrhineae*. *Annals of Botany* 117: 1133–1140.

Epling, C., and T. Dobzhansky. 1942. Genetics of natural populations. VI. Microgeographic races in *Linanthus parryae*. *Genetics* 27: 317-322.

Evans, M. E. K., S. A. Smith, R. S. Flynn, and M.J. Donoghue. 2009. Climate, niche evolution, and diversification of the "bird-cage" evening primroses (*Oenothera*, sections Anogra and Kleinia). *American Naturalist* 173: 225–240.

Faircloth, B. C. 2016. PHYLUCE is a software package for the analysis of conserved genomic loci. *Bioinformatics* 32: 786–788.

Ferrari, S. L. P, and F. Cribari-Neto. 2004. Beta Regression for Modelling Rates and Proportions. *Journal of Applied Statistics* 31: 799–815.

FitzJohn, R. G. 2012. Diversitree: comparative phylogenetic analyses of diversification in R. *Methods in Ecology and Evolution* 3.6: 1084–1092.

Fitzpatrick, B. M., and M. Turelli. 2006. The geography of mammalian speciation: mixed signals from phylogenies and range maps. *Evolution* 60: 601–615.

Friedman, J. 2020. The evolution of annual and perennial plant life histories: Ecological correlates and genetic mechanisms. *Annual Review of Ecology, Evolution, and Systematics* 51: 461–481.

Goodwillie, C. 1997. Self-incompatibility and mating system evolution in *Linanthus* section Leptosiphon. Ph.D. dissertation, University of Washington, Seattle, WA, USA.

Goodwillie, C. 1999. Multiple origins of self-compatibility in *Linanthus* section Leptosiphon (Polemoniaceae): Phylogenetic evidence from internal transcribed sequence data. *Evolution* 53: 1387–1395.

Goodwillie, C., and J. M. Ness. 2005. Correlated evolution in floral morphology and the timing of self-compatibility in *Leptosiphon jepsonii* (Polemoniaceae). *International Journal of Plant Sciences* 166: 741–751.

Goodwillie, C., and J. W. Stiller. 2001. Evidence for polyphyly in a species of *Linanthus* (Polemoniaceae): Convergent evolution in self-fertilizing taxa. *Systematic Botany* 26: 273–282.

Grant, V. 1959. Natural history of the Phlox family. Vol. 1. Systematic botany. Martinus Nijhoff, The Hague, The Netherlands.

Grant, V. 1983. The systematic and geographical distribution of hawkmoth flowers in the temperate North American flora. *Botanical Gazette* 144: 439–449.

Gray, S. M., and J. S. McKinnon. 2007. Linking color polymorphism maintenance and speciation. *Trends in Ecology and Evolution* 22: 71–79.

Grossenbacher, D. L., S. D. Veloz, and J. P. Sexton. 2014. Niche and range size patterns suggest that speciation begins in small, ecologically diverged populations in North American Monkeyflowers (*Mimulus* spp.). *Evolution* 68: 1270–1280.

Hernández-Hernández, T., J. W. Brown, B. O. Schlumpberger, L. E. Eguiarte, and S. Magallón. 2014. Beyond aridification: Multiple explanations for the elevated diversification of cacti in the New World Succulent Biome. *New Phytologist* 202: 1382–1397.

Huelsenbeck, J. P., R. Nielsen, and J. P. Bollback. 2003. Stochastic mapping of morphological characters. *Systematic Biology* 52: 131–158.

Jacobs, S.J., M.C. Grundler, C.L. Henriquez, and Zapata F. 2021. An integrative genomic and phenomic analysis to investigate the nature of plant species in *Escallonia* (Escalloniaceae). *Scientific Reports* 11: 24013.

Jamie, G. A., and J. I. Meier. 2020. The persistence of polymorphisms across species radiations. *Trends in Ecology and Evolution* 35: 795–808.

Johnson, M. G., E. M. Gardner, Y. Liu, R. Medina, B. Goffinet, A. J. Shaw, N. J. C. Zerega, and N. J. Wickett. 2016. HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *Applications in Plant Science* 4: 1–7.

Johnson, M. G., L. Pokorny, S. Dodsworth, L. R. Botigué, R. S. Cowan, A. Devault, W. L. Eiserhardt et al. 2019. A universal probe set for targeted sequencing of 353 nuclear genes from any flowering plant designed using k-medoids clustering. *Systematic Biology* 68: 594–606.

Jolivet, P., and J. W. Foley. 2015. Solutions for purifying nucleic acids by solid-phase reversible immobilization (SPRI). Ludmer Centre for Neuroinformatics and Mental Health, Montreal, QC.

Junier, T., and E. M. Zdobnov. 2010. The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. *Bioinformatics* 26: 1669–1670.

Katoh, K., G. Asimenos, and H. Toh. 2009. *In* D. Posada [ed.], Multiple Alignment of DNA Sequences with MAFFT. Bioinformatics for DNA Sequence Analysis, 39–64. Totowa, NJ: Humana Press.

Knudsen, J. T., and L. Tollsten. 1993. Trends in floral scent chemistry in pollination syndromes: floral scent composition in moth-pollinated taxa. *Botanical Journal of the Linnean Society* 113: 263–284.

Kraft, N. J. B., B. G. Baldwin, and D. D. Ackerly. 2010. Range size, taxon age and hotspots of neoendemism in the California flora. *Diversity and Distributions* 16: 403–413.

Landis, J. B. 2016. Evolutionary development of flower variation in Polemoniaceae: Linking cellular phenotypes, genetics, and pollinator shifts. Ph.D. Dissertation, University of Florida, Gainesville, FL, USA.

Landis, J. B., C. D. Bell, M. Hernandez, R. Zenil-Ferguson, E. W. McCarthy, D. E. Soltis, and P. S Soltis. 2018. Evolution of floral traits and impact of reproductive mode on diversification in the phlox family (Polemoniaceae). *Molecular Phylogenetics and Evolution* 127: 878–890.

Larridon, I., T. Villaverde, A. R. Zuntini, L. Pokorny, G. E. Brewer, N. Epitawalage, I. Fairlie, et al. 2020. Tackling rapid radiations with targeted sequencing. *Frontiers in Plant Science* 10: 1–17.

Lewis, P. O. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology* 50: 913–925.

Lichter-Marck, I. H., and B. G. Baldwin. 2023. Edaphic specialization onto bare, rocky outcrops as a factor in the evolution of desert angiosperms. *Proceedings of the National Academy of Sciences* 120: e2214729120.

Lichter-Marck, I. H., W. A. Freyman, C. M. Siniscalchi, J. R. Mandel, A. Castro-Castro, G. Johnson, and B. G. Baldwin. 2020. Phylogenomics of *Perityleae* (Compositae) provides new insights into morphological and chromosomal evolution of the rock daisies. *Journal of Systematics and Evolution* 58: 853–880.

Losos, J. B., and R. E. Glor. 2003. Phylogenetic comparative methods and the geography of speciation. *Trends in Ecology and Evolution* 18: 220–227.

Mabry, M. E., and M. G. Simpson. 2018. Evaluating the Monophyly and Biogeography of *Cryptantha* (Boraginaceae). *Systematic Botany* 43: 53–76.

Maddison W. P., P. E. Midford, and S.P. Otto. 2007. Estimating a binary character's effect on speciation and extinction. *Systematic Biology* 56: 701–710.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17:10–12.

Mayr, E. 1959. Isolation as an evolutionary factor. *Proceedings of the American Philosophical Society* 103: 221–230.

McLay, T. G., J. L. Birch, B. F. Gunn, W. Ning, J. A. Tate, L. Nauheimer, E. M. Joyce, et al. 2021. New targets acquired: Improving locus recovery from the Angiosperms353 probe set. *Applications in Plant Sciences* 9: e11420.

Millennium Ecosystem Assessment. 2005. Ecosystems and human well-being: Desertification synthesis. Washington, DC.

Miller, E. T., G. M. Leighton, B. G. Freeman, A. C. Lees, and R. A. Ligon. 2019. Ecological and geographical overlap drive plumage evolution and mimicry in woodpeckers. *Nature Communications* 10: 1602.

Minh, B. Q., H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. von Haeseler, and R. Lanfear. 2020. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution* 37: 1530–1534.

Mooney, H., and E. Zavaleta. 2016. Ecosystems of California. University of California Press.

Moore, M. J., and R. K. Jansen. 2006. Molecular evidence for the age, origin, and evolutionary history of the American desert plant genus *Tiquilia* (Boraginaceae). *Molecular Phylogenetics and Evolution* 39: 668–687.

Moran, R. 1977. New or renovated Polemoniaceae from Baja California, Mexico (*Ipomopsis, Linanthus, Navarretia*). *Madroño* 24: 141–159.

Nute, M., J. Chou, E. K. Molloy, and T. Warnow. 2018. The performance of coalescent-based species tree estimation methods under models of missing data. *BMC Genomics* 19(Suppl 5): 1–22.

Paradis, E., J. Claude, and K. Strimmer. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20: 289–290.

Patterson, R. W., and J. M. Porter. 2021. In Jepson Flora Project (Eds.), *Linanthus*. Jepson eFlora, Revision 9. Website: https://ucjeps.berkeley.edu/eflora/eflora_display.php?tid=8871, [accessed 15 April 2023].

Pearman, P. B., T. S. Alioto, J. R. P. Trotta, and J. T. Columbus. 2021. Genotyping-by-sequencing resolves relationships in Polygonaceae tribe *Eriogoneae*. *Taxon* 70: 826–841.

Peng, J., D. L. Swofford, and L. Kubatko. 2022. Estimation of speciation times under the multispecies coalescent. *Bioinformatics* 38: 5182–5190.

Peterson, B. K., J. N. Weber, E. H. Kay, H. S. Fisher, and H. E. Hoekstra. 2012. Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One* 7: e37135.

Porter, J. M., L. A. Johnson, and D. Wilken. 2010. Phylogenetic systematics of *Ipomopsis* (Polemoniaceae): Relationships and divergence times estimated from chloroplast and nuclear DNA sequences. *Systematic Botany* 35: 181–200.

Porter, J. M., and L. A. Johnson. 2000. A phylogenetic classification of Polemoniaceae. *Aliso* 19: 55–91.

Porter, J. M., and R. W. Patterson. 2015. A fistful of Polemoniaceae: New names and combinations. *Aliso* 32: 55–88.

Raguso, R. A., R. A. Levin, S. E. Foose, M. W. Holmberg, and L. A. McDade. 2003. Fragrance chemistry, nocturnal rhythms and pollination "syndromes" in *Nicotiana*. *Phytochemistry* 63: 265–284.

Raguso, R. A., and M. A. Willis. 2002. Synergy between visual and olfactory cues in nectar feeding by naive hawkmoths, *Manduca sexta*. *Animal Behaviour* 64: 685–695.

Rausher, M. D. 2008. Evolutionary transitions in floral color. *International Journal of Plant Sciences* 169: 7–21.

Revell, L. J. 2012. phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* 3: 217–223.

Ricklefs, R.E., and S.S. Renner. 1994. Species richness within families of flowering plants. *Evolution* 48: 1619–1636.

Rigby, R. A., and D. M. Stasinopoulos. 2005. Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society Series C: Applied Statistics* 54: 507–554.

Rohland, N., and D. Reich. 2012. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Research* 22: 939-946.

Rose, J. P., R. Kriebel, L. Kahan, A. DiNicola, J. G. González-Gallegos, F. Celep, E. M. Lemmon, et al. 2021. Sage insights into the phylogeny of *Salvia*: Dealing with sources of discordance within and across genomes. *Frontiers in Plant Science* 12: 1–14.

Rose, J. P., and K. J. Sytsma. 2021. Complex interactions underlie the correlated evolution of floral traits and their association with pollinators in a clade with diverse pollination systems. *Evolution* 75: 1431–1449.

Rubin B. E. R., R. H. Ree, and C. S. Moreau. 2012. Inferring phylogenies from RAD sequence data. *PLoS One* 7: e33394.

Rydberg, P. 1906. Studies on the Rocky Mountain flora. XVI. *Bulletin of the Torrey Botanical Club* 33: 149.

Sapir, Y., M. K. Gallagher, and E. Senden. 2021. What maintains flower colour variation within populations? *Trends in Ecology and Evolution* 36: 507–519.

Schemske, D. W., and P. Bierzychudek. 2001. Perspective: Evolution of flower color in the desert annual *Linanthus parryae*: Wright revisited. *Evolution* 55: 1269–1282.

Schemske, D. W., and P. Bierzychudek. 2007. Spatial differentiation for flower color in the desert annual *Linanthus parryae*: Was Wright right? *Evolution* 61: 2528–2543.

SEINet. 2022. Southwest Environmental Information Network. Website: http//:swbiodiversity.org/seinet/index.php [accessed 12 October 2022].

Silberbauer-Gottsberger, I., and G. Gottsberger. 1975. Über sphingophile Angiospermen Brasiliens/About Sphingophilous Angiosperms in Brazil. *Plant Systematics and Evolution* 123: 157–184.

Simpson, M. G., C. M. Guilliams, K. E. Hasenstab-Lehman, M. E. Mabry, and L. Ripma. 2017. Phylogeny of the popcorn flowers: Use of genome skimming to evaluate monophyly and interrelationships in subtribe Amsinckiinae (Boraginaceae). *Taxon* 66: 1406–1420.

Singhal, S., A. B. Roddy, C. DiVittorio, A. Sanchez-Amaya, C. L. Henriquez, C. R. Brodersen, S. Fehlberg, and F. Zapata. 2021. Diversification, disparification, and hybridization in the desert shrubs Encelia. *New Phytologist* 230: 1228–1241.

Skeels, A., and M. Cardillo. 2019. Reconstructing the geography of speciation from contemporary biodiversity data. *American Naturalist* 193: 240–255.

Slimp, M., L. D. Williams, H. Hale, and M. G. Johnson. 2021. On the potential of Angiosperms353 for population genomic studies. *Applications in Plant Sciences* 9: 1–13.

Smith, S. A., and J. M. Beaulieu. 2009. Life history influences rates of climatic niche evolution in flowering plants. *Proceedings of the Royal Society B: Biological Sciences* 276: 4345–4352.

Smith, S. A., and M. J. Donoghue. 2008. Rates of molecular evolution are linked to life history in flowering plants. *Science* 322: 86–89.

Smith, S. D., and E. E. Goldberg. 2015. Tempo and mode of flower color evolution. *American Journal of Botany* 102: 1014–1025.

Soltis, D. E., M. E. Mort, M. Latvis, E. V. Mavrodiev, B. C. O'meara, P. S. Soltis, J. G. Burleigh, and R. R. De Casas. 2013. Phylogenetic

relationships and character evolution analysis of Saxifragales using a supermatrix approach. *American Journal of Botany* 100: 916–929.

Spencer, C., and J. M. Porter. 1997. Evolutionary Diversification and Adaptation to Novel Environments in *Navarretia*. *Systematic Botany* 22: 649–668.

Stebbins, G. L. 1952. Aridity as a stimulus to plant evolution. *American Naturalist* 86: 33–44.

Stebbins, G. L. 1970. Adaptive radiation of reproductive characteristics in Angiosperms, I: Pollination mechanisms. *Annual Reviews of Ecology and Systematics* 1: 307–326.

Strauss, S. Y., and J. B. Whittall. 2006. Non-pollinator agents of selection on floral traits. Pages 120–138 in L. D. Harder and S.C.H. Barrett, eds. Ecology and evolution of flowers. Oxford University Press, Oxford.

Swofford, D. L. 2002. PAUP. Phylogenetic analysis using parsimony (and other methods). Version 5. Sinauer Associates, Sunderland, Massachusetts.

Tanaka, Y., N. Sasaki, and A. Ohmiya. 2008. Biosynthesis of plant pigments: Anthocyanins, betalains and carotenoids. *Plant Journal* 54: 733–749.

Tank, D. C., and R. G. Olmstead. 2008. From annuals to perennials: Phylogeny of subtribe Castillejinae (Orobanchaceae). *American Journal of Botany* 95: 608–625.

Taylor R. S., and V. L. Friesen. 2017. The role of allochrony in speciation. *Molecular Ecology* 26: 3330-3342.

Thorne, R. F. 1986. A historical sketch of the vegetation of the Mojave and Colorado Deserts of the American Southwest. *Annals of the Missouri Botanical Garden* 73: 642–651.

Thornhill, A. H., B. G. Baldwin, W. A. Freyman, S. Nosratinia, M. M. Kling, N. Morueta-Holme, T. P. Madsen, et al. 2017. Spatial phylogenetics of the native California flora. *BMC Biology* 15: 1–18.

Tripp, E. A., and P. S. Manos. 2008. Is floral specialization an evolutionary dead-end? Pollination system transitions in *Ruellia* (Acanthaceae). *Evolution* 62: 1712–1737.

Ushimaru, A., and K. Nakata. 2002. The evolution of flower allometry in selfing species. *Evolutionary Ecology Research* 8: 1217–1227.

Vaidya, P., A. McDurmon, E. Mattoon, M. Keefe, L. Carley, C. R. Lee, R. Bingham, and J. T. Anderson. 2018. Ecological causes and consequences of flower color polymorphism in a self-pollinating plant (*Boechera stricta*). *New Phytologist* 218: 380–392.

Vasile, M. A., J. Jeiter, M. Weigend, and F. Luebert. 2020. Phylogeny and historical biogeography of Hydrophyllaceae and Namaceae, with a special reference to *Phacelia* and *Wigandia*. *Systematics and Biodiversity* 18: 757–770.

Verity, R., and R. A. Nichols. 2016. Estimating the number of subpopulations (K) in structured populations. *Genetics* 203: 1827–1839.

Walden, G. K., L. M. Garrison, G. S. Spicer, F. W. Cipriano, and R. W. Patterson. 2014. Phylogenies and chromosome evolution of *Phacelia* (Boraginaceae: Hydrophylloideae) Inferred from nuclear ribosomal and chloroplast sequence data. *Madroño* 61: 16–47.

Warren, J., and S. Mackenzie. 2001. Why are all colour combinations not equally represented as flower-colour polymorphisms? *New Phytologist* 151: 237–241.

Wright, S. 1943. An analysis of local variability of flower color in *Linanthus parryae*. *Genetics* 28: 139–156.

Zhang, C., M. Rabiee, E. Sayyari, and S. Mirarab. 2018. ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* 19: 15–30.

SUPPORTING INFORMATION

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

Appendix S1. Sample metadata with location and provenance.

Appendix S2. Characteristics of *Linanthus* species, including number of samples included in analyses, and phenotypic traits.

Appendix S3. Location of samples for all the samples used in the ddRAD phylogenetic and population structure analyses.

Appendix S4. Ipyrad parameters for assembly of RAD data.

Appendix S5. AIC scores for models of evolution to use in stochastic character mapping.

Appendix S6. Posterior probability density trees of anthocyanin polymorphism evolutionary history with two alternate models of evolution: equal rates and irreversible models.

Appendix S7. A maximum likelihood phylogenetic tree inferred in IQtree, including 180 samples for all species of *Linanthus* using a concatenated matrix of ddRAD data and a minimum of four samples per locus.

Appendix S8. Tanglegram of RAD and TC phylogenies with samples in common.

Appendix S9. A distribution of trait changes across 1000 stochastic character mapped trees for perenniality, night-blooming, and corolla anthocyanin polymorphism.

Appendix S10. ddRAD species trees.

Figure S1. Species tree with ddRAD data generated with SVDQuartets.

Figure S2. The same species tree with branch lengths showed inconclusive species relationships due to very short branches.

Appendix S11. TC species trees.

Figure S1. Phylogenetic tree built with 63 samples across 22 species of Linanthus and four outgroups with 219 loci recovered via target capture (Angiosperms353) with 100% occupancy generated in ASTRAL.

Figure S2. Species tree built with 219 loci recovered via target capture (Angiosperms353) with 100% occupancy generated in ASTRAL.

Appendix S12. The species tree, calibrated using the divergence time of 20.8–42.5 mya (Landis et al., 2018), for a range of lambda values ranging from zero, representing a free rate model, to one, representing a clock-like model.

How to cite this article: Anghel, I. G., L. L. Smith, I. H. Lichter-Marck, and F. Zapata. 2025. When the sand blossoms: Phylogeny, trait evolution, and geography of speciation in *Linanthus*. *American Journal of Botany* 112(3): e70005. https://doi.org/10.1002/ajb2.70005