PNAS RESEARCH ARTICLE | EVOLUTION

Chromosomal evolution, environmental heterogeneity, and migration drive spatial patterns of species richness in *Calochortus* (Liliaceae)

Nisa Karimi[a,b], Christopher P. Krieg[b], Daniel Spalink[c], Alan R. Lemmon[d], Emily Moriarty Lemmon[e], Evan Eifler[b], Adriana I. Hernández[f,g], Patricia W. Chan[b], Aarón Rodríguez[h], Jacob B. Landis[f,h,i], Susan R. Strickler[i], Chelsea D. Specht[f,g], and Thomas J. Givnish[b,1]

Edited by Douglas Schemske, Michigan State University, East Lansing, MI; received March 31, 2023; accepted December 20, 2023

We used nuclear genomic data and statistical models to evaluate the ecological and evolutionary processes shaping spatial variation in species richness in *Calochortus* (Liliaceae, 74 spp.). *Calochortus* occupies diverse habitats in the western United States and Mexico and has a center of diversity in the California Floristic Province, marked by multiple orogenies, winter rainfall, and highly divergent climates and substrates (including serpentine). We used sequences of 294 low-copy nuclear loci to produce a time-calibrated phylogeny, estimate historical biogeography, and test hypotheses regarding drivers of present-day spatial patterns in species number. Speciation and species coexistence require reproductive isolation and ecological divergence, so we examined the roles of chromosome number, environmental heterogeneity, and migration in shaping local species richness. Six major clades—inhabiting different geographic/climatic areas, and often marked by different base chromosome numbers (n = 6 to 10)—began diverging from each other ~10.3 Mya. As predicted, local species number increased significantly with local heterogeneity in chromosome number, elevation, soil characteristics, and serpentine presence. Species richness is greatest in the Transverse/Peninsular Ranges where clades with different chromosome numbers overlap, topographic complexity provides diverse conditions over short distances, and several physiographic provinces meet allowing immigration by several clades. Recently diverged sister-species pairs generally have peri-patric distributions, and maximum geographic overlap between species increases over the first million years since divergence, suggesting that chromosomal evolution, genetic divergence leading to gametic isolation or hybrid inviability/sterility, and/or ecological divergence over small spatial scales may permit species co-occurrence.

California Floristic Province | diversification | environmental heterogeneity | historical biogeography | phylogenomics

Significance

Teasing apart the roles of evolution (speciation, reproductive isolation) and ecology (resource diversity, resource partitioning, geographic spread) is fundamental to understanding how spatial variation in diversity arises within species-rich lineages. We inferred relationships in the bulbous genus *Calochortus* using phylogenomic data and analyzed spatial occurrence data to determine how chromosomal evolution, environmental heterogeneity, dispersal, and historical biogeography help explain *Calochortus'* spatial biodiversity patterns and center of diversity in the California Floristic Province, a global biodiversity hotspot. Local species richness appears driven by environmental heterogeneity (including serpentine), evolution of chromosomal number diversity, and proximity to multiple physiographic regions. Divergence in chromosome number appears to play a key role in allowing several mountain-associated lineages to co-occur following secondary contact.

For closely related species to coexist and contribute to local species richness within a lineage, they must be both ecologically and reproductively isolated (1, 2). That is, they must be able to partition a range of resources (e.g., regions, habitats, microsites, successional stages) or otherwise avoid competitively excluding each other, and must possess extrinsic or intrinsic mating barriers (e.g., differences in distribution, phenology, mating compatibility, karyotype). Competition between close relatives can select for ecological isolation via character displacement and adaptive radiation (3–5). Selection can also increase reproductive isolation via adaptation to different conditions—leading to ecological speciation (6)—and via reinforcement based on lower hybrid fitness (7). Ecological and reproductive isolation can occur simultaneously via a special form of ecological speciation, in which characters crucial for ecological divergence are also favored by mates or pollinators (6, 8). Mating barriers usually considered non-ecological (e.g., differences in chromosome number and rearrangements) may be favored initially only if tied to ecological advantages, based on the edge of more common ancestral forms in avoiding hybridization (9). However, that edge might be reversed if new karyotypic mutants arise in allopatric populations, giving them an initial numerical advantage, especially in selfing or inbred taxa. Non-ecological speciation and non-adaptive radiation driven by poor dispersal and local genetic differentiation can multiply species within lineages with little or no ecological divergence (3, 10) but is unlikely to result in local coexistence of close relatives. Overall, local species richness within a lineage should increase with resource heterogeneity available for partitioning, frequency of extrinsic or intrinsic mating barriers, rate of local speciation, and immigration of species or clades from other areas (3, 11–16).

Recent orogenies—leading to high local topographic, climatic, and soil heterogeneity, and topographically dissected, often extensive landscapes—are frequently tied to high

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: givnish@wisc.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2305228121/-/DCSupplemental.

Published February 23, 2024.

local species richness and rapid species diversification. Examples in plants include the Hawaiian lobelioids (13), South American bromeliads, bellflowers, orchids, and other groups (17–22), New Guinea orchids (19), western North American *Lupinus* (23), several lineages in the Hengduan Mountains (24), and ferns worldwide (25). Comparable cases in animals include Andean birds, glass frogs, and lizards (26, 27), Hengduan voles (28), and amphibians, birds, and mammals worldwide (29, 30). Mountain ranges are unusually diverse for their area, with 87% of all higher vertebrates and a large share of plant diversity in 25% of the land area outside Antarctica (19, 31, 32). Mountains are likely to foster higher speciation rates due to environmental heterogeneity and abundant habitat barriers, lower extinction rates due to short migration distances needed to accommodate climate change, and greater coexistence over ecological timeframes due to local habitat diversity (13, 17–25). Lineage immigration and accumulation of species can also contribute to spatial diversity patterns, as seen in several conifer hotspots (16), high Neotropical orchid diversity (19), and co-occurrence of elements of multiple oak lineages along several ecological gradients (33). Drivers of spatial variation in diversification and species richness within clades have been studied less frequently and only in a few instances at fine spatial scales within and among individual mountain ranges and basins (13, 17–19, 25, 34). For example, in mountainous western North America, several recent studies address the phylogeny and time of origin of species-rich plant lineages (e.g., refs. 35–40), but few calculate rates of species diversification, and none evaluate multiple potential drivers of diversification rate or spatial variation in species richness. Integrative analyses are needed to relate the diversity of coexisting species to local environmental heterogeneity, mating barriers, biogeography, and evolutionary history.

Here, we investigate the ecological and evolutionary drivers shaping spatial variation in species richness in *Calochortus* (74 spp., Liliaceae) from western North America. *Calochortus* is an excellent system for studying the determinants of spatial variation in species richness, given its many species with narrow geographic distributions and their occurrence across a wide range of climates, soils, elevations, and mountain ranges of different elevations and ages (41–44). *Calochortus* is an iconic genus of bulbous geophytes, marked by kaleidoscopic variation in floral form and color. It ranges across the western United States and Mexico in a wide variety of habitats, most of them open and at lower to middle altitudes, with a center of diversity in the California Floristic Province (CFP) (Fig. 1). Here, up to 10 species can co-occur within 0.25° × 0.25° cells in the Transverse Ranges and Bay Area and 4 to 7 species across the Coast Ranges and Sierra Nevada, vs. only 1 to 2 species over the western Great Plains. The CFP is marked by rich native flora, multiple orogenies, and highly divergent climates and substrates (including several areas of serpentine) (45). Thirty *Calochortus* species (41%) occur on or are endemic to serpentine soils (43–47); 22 (49%) of the 45 species native to California are considered state or globally rare (G1-G3 or S1-S3) (48).

In a previous study of this system, Patterson & Givnish (42) inferred a phylogeny for 64 species of *Calochortus* based on three plastid genes and identified seven major clades associated with different geographic areas; those clades were generally well supported, but relationships within them were frequently unresolved or poorly supported. They suggested that the large number of *Calochortus* species, their often-narrow distributions, and recurrent evolution of serpentine tolerance reflected short-distance seed dispersal, leading to differentiation and speciation at small spatial scales. Consistent with this, Henns et al. (49) found spatial genetic structure within *Calochortus albus* at small scales, with a mean inferred distance of overall gene flow 0 = 5 to 43 m.

Patterson and Givnish proposed that topographic complexity—and resulting barriers to gene flow—in the mountainous areas occupied by *Calochortus* was a major impetus to geographic speciation, as was edaphic speciation involved with recurrent invasion of serpentine substrates. The latter have high concentrations of heavy metals, high Mg:Ca ratios, and large fractions of bare ground, all hostile to most plants (46, 50). Patterson and Givnish also proposed that *Calochortus* species diversity partly reflected shifts in chromosome number, arguing that such shifts provided reproductive isolation and allowed three pairs of major clades to "double up" and diversify in the same geographic areas (Coast Ranges/Sierra Nevada, northern Rocky Mountains, Transverse/Peninsular Ranges) without hybridizing. They suggested that adaptive radiation into different habitats across small spatial scales also contributed to diversification in *Calochortus*.

Based on the above considerations, here we extend the analyses of Patterson & Givnish (42) using greatly improved phylogenetic methods, almost complete species sampling, and adding spatially structured data on ecological factors that might affect species diversity to test the following predictions regarding the genesis and maintenance of species richness in *Calochortus*:

1. Recently divergent sister species should have peripatric ranges, reflecting limited dispersal and geographic speciation;
2. The likelihood of species overlap should increase with time since divergence from a common ancestor, reflecting the time needed for reproductive and ecological isolating mechanisms to evolve; and
3. Local species richness should increase with a) local environmental heterogeneity in elevation, climate, and soils, including serpentine; b) local diversity of chromosome numbers; c) immigration of clades from adjoining regions, d) lower soil moisture and greater seasonality in potential evapotranspiration, because moist, less seasonal habitats result in dense canopies that adversely affect understory bulbous herbs (51), and e) greater concentration of precipitation in winter, given the abundance and adaptive advantage of bulbous/cormous herbs in winter-rainfall climates — where they comprise 5% of vascular plant species in the CFP (52) and 17% in southwest South Africa (53)—and where storage organs allow rapid deployment of leaf tissue to photosynthesize during the brief window of favorable conditions in spring.

To address these hypotheses—founded on the principle that speciation and species coexistence require both reproductive isolation and ecological divergence—here we present the first nuclear molecular phylogeny for *Calochortus* based on putative single-copy loci. We calibrate this phylogeny against time and use it to estimate the evolution of chromosome number, serpentine tolerance, and patterns of geographic spread and spatial overlap. We then evaluate the roles of environmental heterogeneity, variation in chromosomal number, immigration, and rate and duration of speciation in generating spatial and phylogenetic patterns in *Calochortus* diversity.

Results

Nuclear DNA Phylogeny. Maximum likelihood (ML) analysis of concatenated nuclear loci—and ASTRAL analysis using an approximation of the multi-species coalescent model—yielded six major clades, each corresponding to a geographic region associated with or largely restricted to one or more mountain ranges (Fig. 2, *SI Appendix*, Figs. S1 and S2; and see *SI Appendix*, Table S1 for data on gene trees). The major clades are fully supported based on ASTRAL posterior probabilities (PP = 1; *SI Appendix*, Fig. S1) as

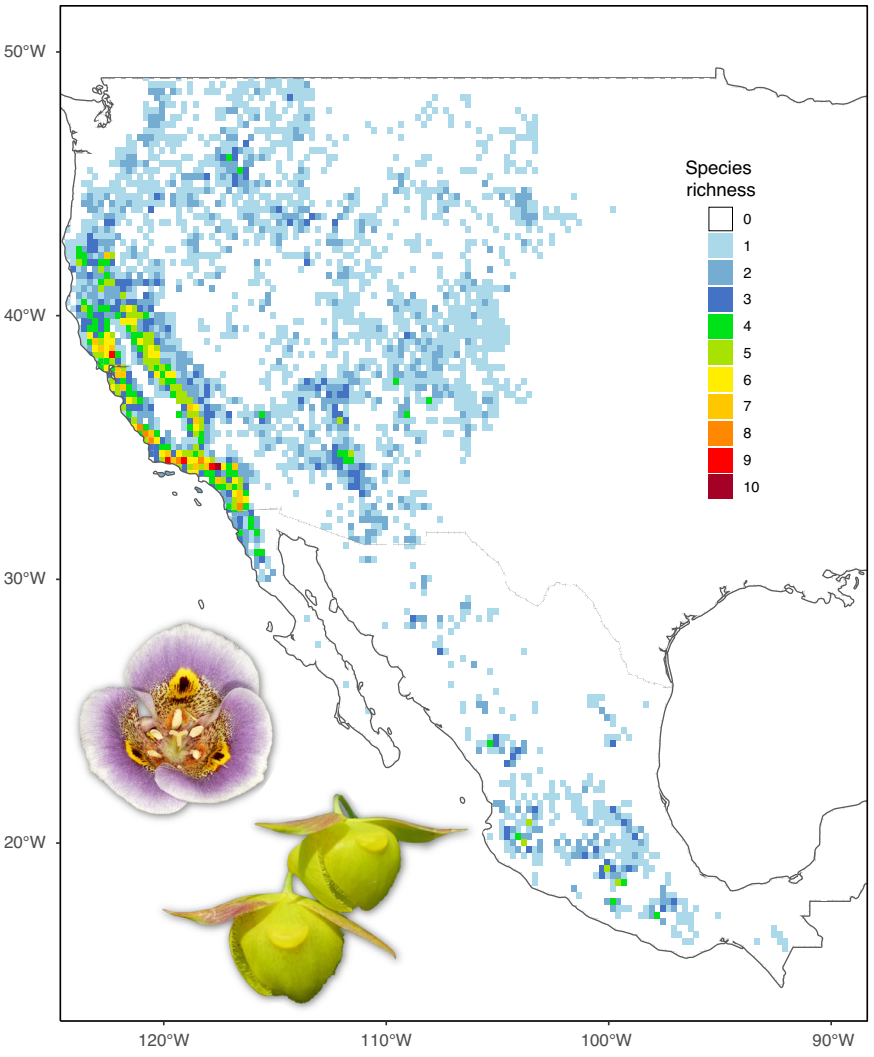

Fig. 1. Map of *Calochortus* species richness in 0.25° × 0.25° cells. (*Inset*) Flowers of *C. superbus* (*Upper*) and *C. pulchellus*.

well as bootstrap support values given the ML nuclear concatenated tree (BS = 100; *SI Appendix*, Fig. S2). These clades include the **Bay Area clade** (centered on the Bay Area, with some species in other parts of the Coast Ranges, as well as the Transverse and Peninsular Ranges, Sierra Nevada, Siskiyous, and Cascades); the **Pacific Northwest clade** (Siskiyous and Cascades, with some species in drier habitats to the east); the **Transverse-Peninsular clade** (with some species in the Southern Coast Ranges); the **California clade** (distributed around the Central Valley, in the Sierra Nevada, Coast Ranges, Transverse Ranges, and Peninsular Ranges); the **Rocky Mountain-Great Basin (RMGB) clade** (ranging from the South Coast Ranges to the Dakotas), and the **Mexico clade** (inhabiting parts of the Sierra Madre Occidental, Sierra Madre Oriental, Trans-Mexican Volcanic Belt, Sierra Madre del Sur, Chihuahuan Desert, Chiapas Highlands, and Pacific lowlands) (Fig. 2). The Bay Area and Pacific Northwest clades are sister to each other, and jointly sister to the Transverse-Peninsular clade; the Mexico and California clades are sister to each other, and jointly sister to the RMGB clade (Fig. 2). These clades and their relationships to each other are consistent with those inferred from plastid data by Patterson & Givnish (36), except that our California clade commingles species from their Southwestern California and Coast Ranges–Sierra Nevada clades. ML analysis of complete plastomes also yields these primary geographic clades but with a slightly different position for the RMGB clade (*SI Appendix*, Fig. S3).

Gene concordance factors (gCFs) and nucleotide site concordance factors (sCFs) estimated by IQ-TREE for the nuclear tree are generally high with a few exceptions (gCF = 0.38 to 0.97; sCF = 0.50 to 0.98; *SI Appendix*, Fig. S1), suggesting that sampling variance is low but combined with some gene tree conflict. The latter could reflect stochastic error from limited data or incomplete lineage sorting (ILS) early in clade diversification, although hybridization/introgression (HI) cannot be excluded at this time (*Discussion*). Nonetheless, the current tree is a reasonable hypothesis for species relationships and suitable for downstream analyses. Most species represented by multiple samples are resolved as monophyletic; the few cases of paraphyly and polyphyly include taxa with peripatric ranges, consistent with ancestor-descendent relationships or recent HI (*SI Appendix*, Table S2).

Evolutionary Timeline. The nuclear chronogram places the *Calochortus* stem at 44.0 Mya and its crown at 10.3 Mya (Fig. 2 and *SI Appendix*, Table S3). These dates are consistent in the plastome and nuclear topologies, with overlapping CI (*SI Appendix*, Fig. S4 and *Methods*). The California and Mexico clades diverged 7.8 Mya, with crowns at 4.5 and 2.6 Mya, respectively. Their common ancestor diverged from that of the RMGB clade 9.5 Mya. The Bay Area and Pacific Northwest clades split 5.4 Mya, with crowns at 2.4 and 2.6 Mya, respectively. The Transverse-Peninsular clade diverged from the Bay Area + Pacific Northwest clades at 5.4 Mya.

Historical Biogeography. Patterson & Givnish (42) used parsimony to infer that *Calochortus* arose in the California Coast Ranges and dispersed from there to most other areas, but our

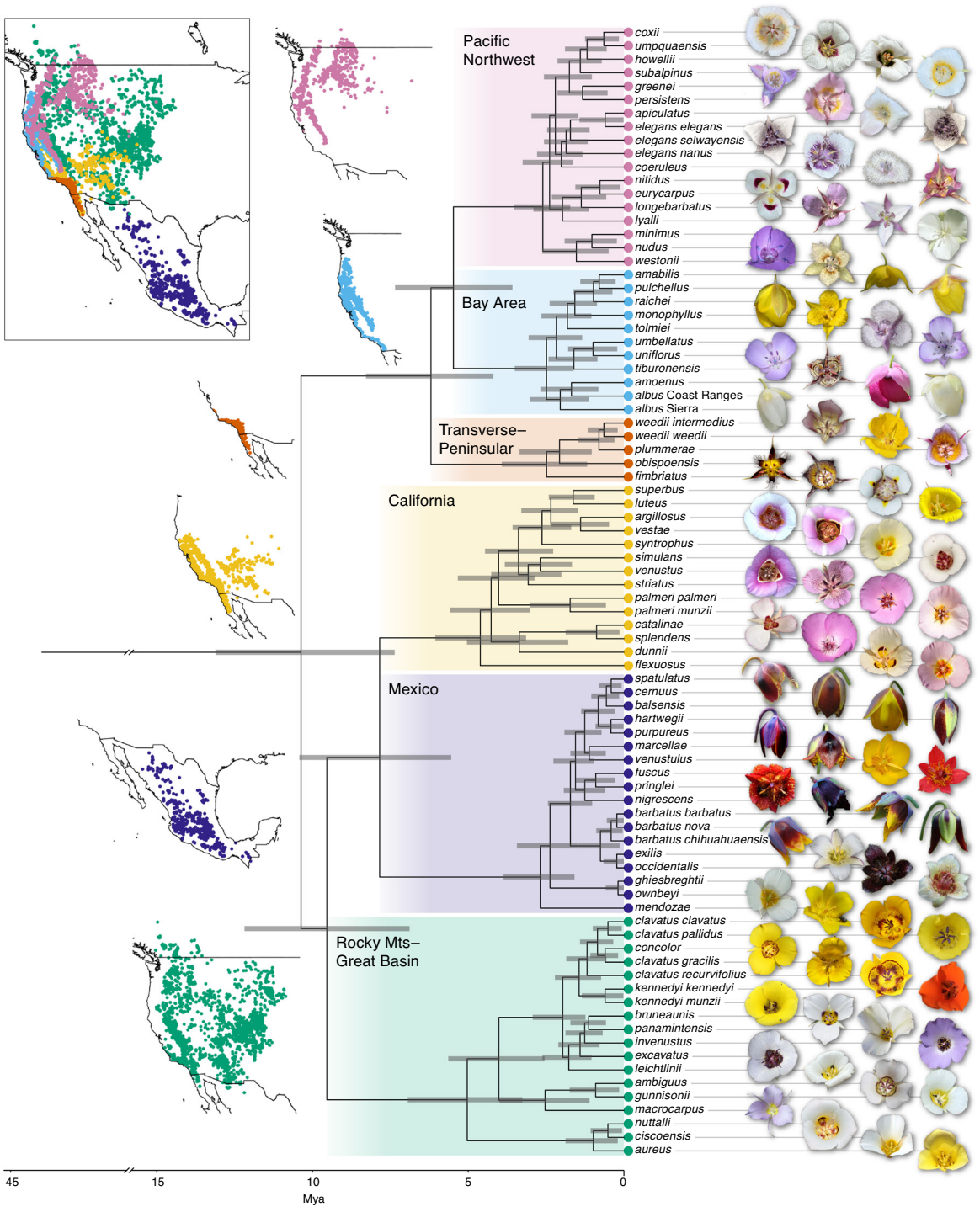

Fig. 2. Time-calibrated ASTRAL phylogeny of *Calochortus*, with each major clade color-coded (*Middle*), the spatial distribution of each major clade (*Left*), and flowers of each species (*Right*, not to scale). Gray bars indicate 95% CI for node ages. Note extensive evolution in floral form and color in several lineages.

ML analysis reveals a more complex history. The DEC (Fig. 3 and *SI Appendix*, Fig. S5) and DEC+J (*SI Appendix*, Fig. S6) ancestral-area models were the best-fit, with nearly identical log-likelihoods (lnL = 137.12 vs. −137.11), followed by the DIVALIKE and DIVALIKE+J models (lnL = −161.68, −153.97). Given *Calochortus*' limited dispersal capacity (43), the controversy around +J models (54, 55), and the nearly identical likelihoods of the two leading models (Fig. 3 and *SI Appendix*, Figs. S5 and S6), here we present the DEC results.

Cascades/Siskiyous + Sierra Nevada is the most likely ancestral range for the crown group of the Pacific Northwest clade (62% probability); Cascades/Siskiyous + Coast Ranges + Sierra Nevada for the Bay Area clade (42%); Transverse/Peninsular Ranges for the Transverse–Peninsular clade (72%); Mexican mountain ranges for the Mexico clade (96%); and the Rocky Mountains/Great Basin for the RMBG clade (60%) (Fig. 3 and *SI Appendix*, Fig. S5). The most likely ancestral area for the California clade was Rocky Mountains/Great Basin *or* Rocky Mountains/Great Basin + Cascades/Siskiyous + Transverse/Peninsular Ranges (equiprobable at 27%). Given the disparate ancestral areas of the major clades, a wide range of scenarios—each with low likelihood—emerged for the ancestral area of the *Calochortus* crown:

Rocky Mountains/Great Basin + Sierra Nevada had the highest probability (30%), followed by Rocky Mountains/Great Basin + Sierra Nevada + Mexico (29%).

Evolution of Chromosome Number and Serpentine Tolerance. Consistent with Patterson & Givnish (42), most major clades differ in base chromosome number (*SI Appendix*, Fig. S7). The optimal ancestral-state estimation used a constant-rates model (*SI Appendix*, Table S4) and resulted in an inferred ancestral chromosome number n = 9 (*SI Appendix*, Fig. S7). This model implies one shift from n = 9 to n = 10 in the Bay Area + Pacific Northwest clades; a second from n = 9 to n =7 in the California clade; a third from n = 9 to n =7 in the stem of the RMBG clade; and a fourth, fifth, and sixth within that clade, from n = 8 to n = 7 and n = 9 (*SI Appendix*, Fig. S7). Karyotypic shifts occur in 13.9% of 72 speciation events, including six shifts in base chromosome number and four tetraploidization events.

Serpentine tolerance—including both **serpentine endemics** (e.g., *Calochortus tiburonensis*) and **serpentine tolerators** (e.g., *C. albus*) that occur on both serpentine and non-serpentine substrates—appears to have arisen and been lost multiple times (*SI Appendix*, Fig. S7), in a pattern consistent with that found by Patterson and Givnish (42).

Shifts in Species Geographic and Environmental Overlap with Time Since Divergence. Members of five major clades are native to parts of California, with maximum areas of geographic overlap among clades

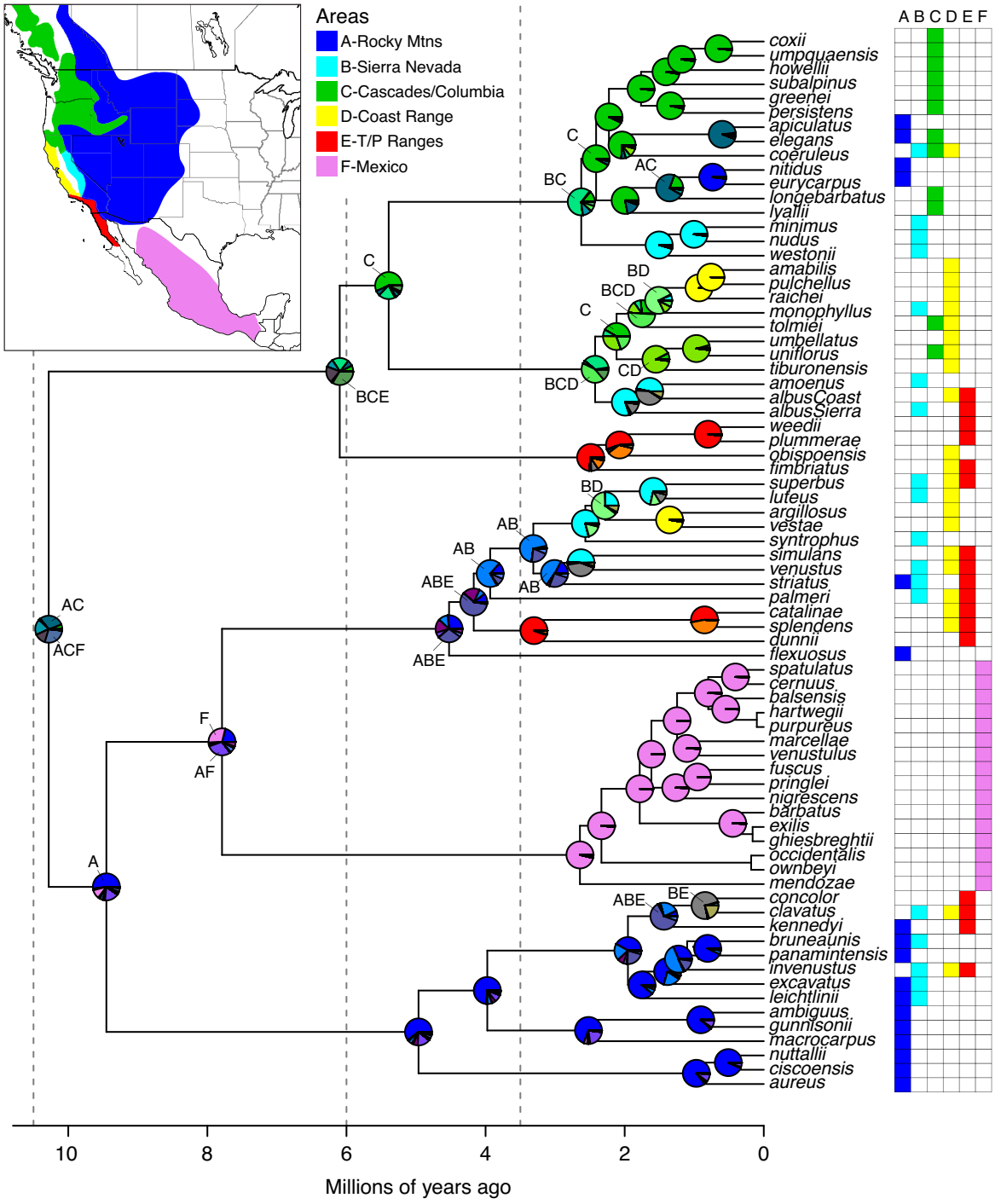

Fig. 3. BioGeoBears DEC model for ancestral-area estimation; current species distributions are indicated in columns to *Right*.

in the Coast Ranges, Transverse Ranges, Sierra Nevada, and Peninsular Ranges (Fig. 2). Pianka's index of geographic overlap is frequently greater between different species in the same clade (that is, among close relatives) than between more distantly related species drawn at random from different clades (*SI Appendix*, Figs. S8 and S9). Note high overlaps between the Transverse-Peninsular clade (n = 9) and California (n = 7) and RMBG (n = 8) clades; between the Bay Area (n = 10) and California (n = 7) clades; and between the California (n = 7) and RMBG clades (n = 8)—all involving pairs of clades with different base chromosome numbers. Permutation tests show that species within the Bay Area, Mexico, California, and RMGB clades overlap significantly more frequently with each other with than species in other clades (*SI Appendix*, Fig. S9).

As predicted, most (18 of 23) pairs of sister species have peripatric ranges (*SI Appendix*, Table S5 and Fig. S10). Peripatric species pairs diverged more recently (0.77 ± 0.41 Mya) than non-peripatric pairs (1.44 ± 0.72 Mya) ($P < 0.007$ for 1-tailed *t*-test assuming homoscedasticity; $P < 0.055$ for the 1-tailed *t*-test with heteroscedasticity). Maximum geographic overlap increases with time since divergence between species over the first million years ($P = 0.011$ to 0.065, depending on binning interval and mapping resolution; *SI Appendix*, Fig. S11 and Table S6). Across all species pairs, geographic overlap declined slightly with time since divergence, but that pattern explains only 1 to 2% of the variance and is driven by many distantly related species having zero overlap. The average distance between the centers of species ranges increases significantly ($r^2 = 0.17$, $P < 0.001$) with time since divergence (*SI Appendix*, Fig. S12), reflecting the spread of *Calochortus* through geographic space with time.

Drivers of Spatial Variation in Species Richness and Phylogenetic Diversity (PD). *Calochortus* species richness increases strongly with elevational range ($r^2 = 0.51$, $P < 0.001$), diversity of chromosome numbers (count of different n values among species) ($r^2 = 0.67$, $P < 0.001$), and winter rainfall index (% of annual precipitation falling from November–March) ($r^2 = 0.21$, $P < 0.001$) in 0.25° × 0.25° grid cells analyzed using independent univariate regressions. When we included all hypothesized non-collinear predictors of local species richness in a spatially unstructured GLM (Generalized linear model), ten variables had significant effects (Table 1). Ranked by significance, diversity of chromosome numbers, range of % soil silt content, serpentine index, range of soil N content, elevational range, range of soil bulk density, mean soil water content, and % winter rainfall had significant positive effects on species richness, while seasonality of potential evapotranspiration and mean soil water content had significant negative effects.

Inclusion of lagged values improved the model based on likelihood ratio tests, with the spatial Durbin model as the best-fit SRM ($\chi^2 = 1{,}459.8$, $P < 10^{-6}$ for 14 d.f.); its AIC was lower than those for SAC, lag, and error models (4,138.6 vs. 4,202.0, 4,287.2, and 4,282.6, respectively). For the optimal SRM, greater diversity of chromosome numbers, elevational range, range of % soil silt content, range of soil bulk density, range of soil nitrogen content, range of % soil clay content, leaf area index, and moisture seasonality index had the most significant positive effects on species richness in rank order, while soil pH range had a significant negative effect (Table 1). That is, a greater diversity of chromosome numbers and greater within-cell ranges in elevation and correlated soil and vegetation characteristics had the most significant positive effects on *Calochortus* species richness, while moister areas and those with greater PET seasonality had negative effects.

The standardized effect size of PD_{ses} shows a geographic pattern similar to that of species richness (*SI Appendix*, Fig. S13), and similar environmental variables and karyotypic variance are main predictors of high PD (*SI Appendix*, Table S7).

No Variation among Clades in Net Species Diversification. Net species diversification rates increased toward the present by ~17% since the *Calochortus* crown but showed no differences among major clades (*SI Appendix*, Fig. S14 and Table S3). Diversification rate showed no relationship to clade species richness ($r^2 = 0.001$, $P > 0.94$). No significant differences emerged in diversification rates among clades and the regions they occupy.

Discussion

Our nuclear data concur with past plastid data (42) in delimiting the same major lineages (i.e., geographic clades) and nearly identical relationships among those lineages, despite large differences in **methodology** (ML and ASTRAL vs. parsimony for phylogenetic inference), **datasets** (nuclear sequences vs. cpDNA restriction sites), and **species coverage**. Our analyses of 294 putatively single-copy nuclear genes fully resolve—mostly with strong support—relationships across *Calochortus*, including many unresolved, poorly supported, or contrary relationships obtained by Patterson & Givnish (42) and the positions of the nine species they did not study. Relationships of species within clades often depart substantially from those obtained earlier, however. For example, not one of the interspecific relationships in the Bay Area clade derived earlier is supported by our analyses. In some cases, this incongruence appears to reflect sampling error inherent to the earlier, far smaller plastid dataset, with relationships in some clades in our plastome phylogeny (*SI Appendix*, Figs. S3 and S4) more closely approaching those in the nuclear phylogeny than those in the earlier plastid phylogeny (e.g., see the RMBG clade); in others, hybridization/ introgression or ILS appears likely, based on conflict between the nuclear and plastome phylogenies and non-monophyly of some species in the plastome phylogeny (e.g., see the Bay Area clade) (Fig. 2 and *SI Appendix*, Figs. S3 and S4). Detailed analyses of discordance between the nuclear and plastome datasets, and among nuclear loci, will be provided elsewhere.

Phylogeny and Historical Biogeography. Our time-calibrated nuclear phylogeny indicates that *Calochortus* diverged from other lineages ~44.0 Mya and began diverging into six major clades between ~7-13 Mya (Fig. 2 and *SI Appendix*, Fig. S5). Most *Calochortus* speciation has been recent: the mean ± SD age of cladogenetic events is 2.19 ± 1.95 Mya, with the youngest third of such events occurring between 0.14 and 1.23 Mya. Estimated crown ages for the major clades are almost all younger than their inferred areas of origin, given the apparent uplift of the Coast Ranges < 3.5 Mya, Transverse Ranges < 5.3 Mya, Peninsular Ranges < 6 Mya, Cascades < 7 Mya, Basin and Range physiographic province < 17 Mya, Sierra Nevada > 30 Mya, and Rocky Mountains > 40 Mya (*SI Appendix*, Table S8).

Among the source areas identified by BioGeoBears, the Bay Area clade most likely arose in the Cascades/Siskiyous and Sierra Nevada, given the younger age of the Coast Ranges (Fig. 3 and *SI Appendix*, Fig. S5). Four species are restricted to the Coast Ranges in the Bay Area and three more occur there. The Pacific Northwest clade arose in the Cascades/Siskiyous and Sierra Nevada (Fig. 3) and then spread into the Coast Ranges and Rocky Mountains. The Transverse/Peninsular clade arose in situ and spread into the South Coast Ranges. Dispersal among the Transverse Ranges, Coast Ranges, and Sierra Nevada is a plausible scenario for the distribution of the combined (TP, (BA, PNW)) clade (Fig. 3). Migration around the Central Valley is also seen in *C. venustus,* which originated in the South Coast Ranges and expanded north along the Coast Ranges and east into the Transverse Ranges and Sierra Nevada (56).

The Rocky Mountains/Great Basin is the most likely ancestral region for the lineage formed by the remaining three clades;

Table 1. GLM and spatial regression model (SRM) results for *Calochortus* species richness at 0.25° × 0.25° grid cells

Predictor	Estimate	SE	z	Pr(>|z|)
GLM • McKelvery's pseudo-R^2 = 0.31, AIC = 6,506.27 • Cragg–Uhler pseudo-R^2 = 0.45; McFadden pseudo-R^2 = 0.18				
# chromosome numbers	3.34E-01	2.61E-02	12.77	2.00E-16
Elevational range	2.13E-04	7.38E-05	2.89	0.004
% silt (range)	2.64E-03	6.41E-04	4.12	3.77E-05
% clay (range)	1.20E-03	6.12E-04	1.96	0.050
% nitrogen (range)	1.41E-03	4.01E-04	3.52	0.0004
pH (range)	7.15E-04	8.14E-03	0.09	0.93
bulk density (bdod) range	1.04E-02	3.65E-03	2.86	0.0042
Thornthwaite's aridity index	−2.45E-03	3.93E-03	−0.62	0.534
P/PET	1.32E-04	2.58E-04	0.51	0.61
leaf area index (LAI) range	4.55E-04	7.62E-04	0.60	0.55
soil water content (SWC)	−8.33E-03	4.11E-03	−2.03	0.043
PET seasonality	−6.59E-05	1.45E-05	−4.55	5.50E-06
Winter rainfall index	2.54E-03	1.26E-03	2.01	0.045
Moisture seasonality index	−1.88E-03	5.70E-03	−0.33	0.742
Proportion of serpentine tolerators	2.36E-01	5.82E-02	4.05	5.13E-05
SRM • Nagelkerke pseudo-R^2 = 0.81, AIC = 4,350.3 • Rho = 0.42893, LR test value: 324.05, $P < 0.001$; Wald statistic = 406.44, $P < 0.001$				
# chromosome numbers	8.27E-01	2.97E-02	27.838	2.20E-16
Elevational range	7.35E-04	7.91E-05	9.287	2.20E-16
% silt (range)	5.23E-03	6.48E-04	8.074	6.66E-16
% clay (range)	2.65E-03	5.91E-04	4.490	7.14E-06
% nitrogen (range)	2.46E-03	4.05E-04	6.072	1.27E-09
pH (range)	−4.96E-02	7.81E-03	−6.358	2.04E-10
bulk density (bdod) range	2.76E-02	3.60E-03	7.651	2.00E-14
Thornthwaite's aridity index	1.02E-03	2.44E-03	0.416	0.677
P/PET	−4.25E-05	2.60E-04	−0.164	0.870
PET seasonality	2.07E-05	3.18E-05	0.651	0.515
Winter rainfall index	−3.71E-03	2.85E-03	−1.301	0.193
Moisture seasonality index	1.66E-02	7.85E-03	2.119	0.034
Proportion of serpentine tolerators	1.66E-02	7.35E-02	0.226	0.821

Predictor variables were selected after removing those correlated at <0.7 threshold based on Pearson correlation coefficient and selected by step AIC.

given the age of the Rockies, it is unsurprising that it is also a common element in the origin of the *Calochortus* crown (Fig. 3). The rise of the RMGB clade in the Great Basin is suggested by the presence there of two early divergent lineages (*nuttallii-ciscoensis-aureus, ambiguus-gunnisonii-macrocarpus*). Occurrence of the *Calochortus* crown group in the southern Great Basin [including the Sonoran Desert starting about 8 Mya (57)] would have allowed dispersal under earlier, more mesic conditions across small gaps now covered by deserts (in southwest CA, southern AZ, northern Sonora) into the Transverse Ranges and northern Mexican mountains, providing access to the ancestral areas for other clades. Mountain ranges and their foothills probably helped shape the current distribution and past dispersal of *Calochortus* because they are associated with greater precipitation and lower evaporation, and thus provided more mesic habitats and corridors in semi-arid regions of the western US and Mexico exposed to past cycles of aridification.

Evolution of Chromosome Number and Serpentine Tolerance. From an ancestral n = 9, we infer six shifts in base chromosome number, in the Pacific Northwest + Bay Area clades (n = 10), the California clade (n = 7), and the Rocky Mountain clade (n = 8, then n = 9 and twice for n = 6) (*SI Appendix*, Fig. S7). Species of these clades with divergent chromosome numbers overlap in distributions in the Coast Ranges, Sierra Nevada, Transverse Ranges, Peninsular Ranges, and northern Rocky Mountains, contributing to local and global diversity in *Calochortus*.

We estimate five origins and 13 losses of serpentine tolerance in the Pacific Northwest, Bay Area, Transverse-Peninsular, and RMGB clades, with gains all associated with serpentine outcrops in California, Oregon, and Washington (*SI Appendix*, Fig. S7). Reversals of serpentine endemism are rare in the California flora, but rates of reversal of serpentine tolerance are high (58).

The origins of serpentine tolerance in *Calochortus* are obscured by the lack of common-garden studies to determine whether tolerance is present in species found only in regions without serpentine (e.g., *C. amoenus, C. concolor*) or in non-serpentine populations of species that grow on serpentine elsewhere (e.g., *C. simulans, C. monophyllus*). Plants on serpentinite-derived soils must tolerate unusual soil chemistry (high heavy metal concentrations, low Ca:Mg ratios, low N, P, and Ca concentrations), thin droughty soils, sparse coverage, exposure to drying sun/wind, and high apparency of individual plants to

herbivores (46, 47, 50, 59–61). Serpentine-tolerant plants are often assumed to be competitively inferior on richer, moister, or less toxic soils, but this tradeoff has been demonstrated infrequently under field conditions (46, 59, 60), and is not true in all cases (59, 62). Adaptation to sparse competition on bare soils (46, 63) or herbivore attack (61) have been implicated in serpentine tolerance in some groups. Serpentine endemics are found on barer and chemically more extreme soils than serpentine tolerators (63), and adaptation to bare soil has been proposed as a precursor for the evolution of serpentine tolerance (50, 61).

Two other key traits adapted to bare microsites—**basal leaves** and **bulbs**—could preadapt *Calochortus* for serpentine tolerance. Leaves held near the ground are adapted to bare or nearly bare sites (64). Basal leaves occur across *Calochortus* and may partly account for it frequently occurring on open, sparsely covered sites—including serpentine—and being especially prominent after fires. Bulbs are adapted to open sites in seasonally dry habitats, especially in winter-rainfall regions (51, 52), like many areas inhabited by *Calochortus*. Serpentine sites—by being open, with soils with low organic content and water storage capacity, and thus subject to seasonal drying in winter-rainfall areas—may give *Calochortus* and other bulbous genera [e.g., *Allium*, *Erythronium*, *Fritillaria*, *Lilium* (46, 47)] an advantage in the CFP.

Peripatry in Early Stages of Cladogenesis. Eighteen of 23 sister-species pairs have peri-patric distributions, supporting a key model of geographic speciation in lineages marked by limited dispersal ability. Consistent with that interpretation, non-peripatric species among these pairs diverged almost twice as long ago (1.44 ± 0.72 Mya) as peripatric species (0.77 ± 0.41 Mya). Our findings contrast with those of Anacker & Strauss (65), who found 80% of 71 plant sister-species pairs they examined in the CFP overlapped; 93% of sister species did, however, exhibit shifts in habitat or soils, implying peripatry at fine geographic scales. We found that only maximum overlap, not overlap across all species, increased with time since divergence over the first million years (*SI Appendix*, Fig. S11). Such a pattern might reflect the amount of time required for species to diverge sufficiently in ecology and mating barriers to permit overlap, and the tendency for more distantly related species, in different clades, to have dispersed to different regions and exhibit no spatial overlap at all. Harrison (66) found a pattern in animals like that in *Calochortus*, with allopatry/peripatry appearing to dominate in sister species endemic to the CFP, although no formal analysis was presented. Paralleling our findings, Drovetski et al. (67) found that—in the Asian bird family Prunellidae—species that diverged within the last 1.5 million years are almost always allopatric, while all older lineages are at least partly sympatric.

Drivers of Spatial Variation in *Calochortus* Species Richness. In accord with predictions, our GLM and SRM showed that the number of *Calochortus* species (and PD) in 0.25° × 0.25° cells increased significantly with local heterogeneity in elevation, soil characteristics, and chromosome base numbers. In the GLM *Calochortus* species richness also increased significantly with % winter rainfall and % serpentine-tolerant species and decreased with mean soil water content.

Topographic relief might be seen as a master driver of local species richness in *Calochortus*, given its impacts on elevational, climatic, and soil heterogeneity, creation of extrinsic habitat barriers to gene flow, and buffering of populations against extinction during climatic fluctuations. Qi & Yang (68) found that elevational heterogeneity was the most important determinant of local species richness of vascular plants in California landscapes. During glacial and interglacial periods, species in mountainous areas need only migrate a few hundred meters down- and upslope (69), a key consideration given *Calochortus'* limited dispersal ability. Such buffering and vertical rather than horizontal migration in mountainous areas may account for why many sister species in *Calochortus* remained peripatric despite large temperature fluctuations during glacial cycles over the past million years.

Perhaps our most surprising result is that variation in base chromosome number (providing reproductive isolation) plays as important a role in driving local species richness in *Calochortus* as environmental heterogeneity (providing ecological isolation and resource partitioning). Previous studies have shown the importance of reproductive isolation in determining whether close relatives can coexist. For example, differences in mast-fruiting years or flowering time within years allow various groups of oaks to coexist (70, 71) and permit the evolution of niche differentiation by largely preventing gene flow between closely related species (39). Closely related, interfertile *Eucalyptus* species rarely co-occur, and reproductive isolation appears more important than ecological isolation (e.g., habitat partitioning) in determining species co-occurrence (72). In *Calochortus,* the karyotypic signal associated with coexistence is strong, and persists when combined in analyses involving ecological predictors of species richness. Showing that the effects of chromosome number on local diversity are only direct, via the reproductive isolation provided by divergence in chromosome number—or partly the result of putative but inevident ecological differences that might have arisen after clades became reproductively isolated by different chromosome numbers—requires more study. On general principles, reproductive isolation is likely to be the key mechanism allowing geographic overlap between lineages soon after they diverge in chromosome number. Sobel et al. (9) argue that karyotypic mutations must be accompanied by ecological shifts if they are to spread, but this may not be true if such mutations arise in allopatry and become fixed, especially in selfing, inbred, or poorly dispersing taxa (described in the introductory paragraphs).

Increases in *Calochortus* species richness (and PD) with local range in soil characteristics can be interpreted simply as resource partitioning of a greater environmental space. But areas with a greater range in % silt and % clay content also often have higher mean % silt and clay contents as well. Such soils provide greater nutrient and moisture supplies than sandier soils and help ensure high photosynthetic rates per leaf mass—and thus, to repaying the costs of bulb and leaf construction during the short photosynthetic season of bulbous plants.

Serpentine tolerance arose multiple times in *Calochortus* and added substantially to its diversity, with 10 species endemic to serpentine substrates. Serpentine adaptation is a notable part of *Calochortus'* ecogeographic radiation in habitat and geographic distribution, adding 16% to the set of species not restricted to serpentine, and 23% to the species never found on serpentine; in all, 41% of *Calochortus* species have invaded ultramafic substrates. The abrupt shift in soil chemistry and plant coverage at the edge of serpentine outcrops means that speciation associated with serpentine can occur at fine spatial scales. This can be seen in *C. tiburonensis* (restricted to an outcrop <1 km across at Ring Mountain, Marin County), *C. raichei* (restricted to an outcrop <13 km long at The Cedars, Sonoma County [R. Raiche, pers. comm.]), and several species in the Siskiyous (each restricted to outcrops < a few tens of km across). While serpentine adds to *Calochortus*' regional diversity, speciation rates on serpentine across several California clades are significantly lower than those on other substrates (58).

Dispersal appears to have played an important role in *Calochortus* diversification, with at least four major clades (Bay Area, Pacific Northwest, California, Transverse/Peninsular) having evolved after dispersal from the ancestral Rocky Mountains/Great Basin + Sierra

Nevada (possibly joined by the Mexican mountains). Immigration also appears to have influenced the location of hotspots of species richness in the Transverse Ranges, Coastal Ranges, and Sierra Nevada of the CFP. One subclade of eight species (*clavatus, concolor, kennedyi, panamintensis, invenustus, excavatus, bruneaunis*) of the RMGB clade immigrated to the CFP after that clade arose in the Rocky Mountain/Great Basin, and the 13 species of the California clade also invaded or persisted in the CFP after that clade arose in Rocky Mountains/Great Basins *or* Rocky Mountains/Great Basin + Cascades/Siskiyous + Transverse/Peninsular Ranges. Both clades evolved chromosome numbers (n = 8 and n = 7, respectively) that differed from those of the Bay Area and Pacific Northwest clades (n = 10) and the Transverse/Peninsular Ranges (n = 9) that arose in the CFP, allowing these clades to overlap and contribute nearly half the 45 species now extant in California. Immigration of the California clade contributed seven species to the Transverse Ranges, seven to the Coast Ranges, five to the Sierra Nevada, three to the Peninsular Ranges, and two to the Siskiyous. Immigration by the RMGB clade contributed six species to the Sierra Nevada, three to the Transverse Ranges, two to the Coast Ranges, and one to the Peninsular Ranges. In all, immigration added ten species to the Transverse Ranges, nine to the Coast Ranges, and eleven to the Sierra Nevada, the three hotspots of *Calochortus* species richness. The multiplication of major clades at the beginning of the Pleistocene 2.4 to 2.6 Mya may reflect enhanced dispersal, then isolation by drier areas as glacial-interglacial climatic differences became far more marked than in the Pliocene. The ring of mountain chains surrounding the Central Valley likely contributed to high *Calochortus* diversity in California via initial differentiation within topographically complex cordilleras and subsequent dispersal to and through other, relatively mesic montane areas, facilitated by ecological heterogeneity and intrinsic mating barriers in *Calochortus* created by divergence in chromosome number.

Our analyses of *Calochortus* evolution integrate phylogenomics, environmental heterogeneity, mating barriers based on karyotypic differences, geological history, and historical biogeography to account for spatial variation in species richness. The drivers of spatial variation in *Calochortus* species richness we identified may also provide a model for aspects of diversification of the entire California flora. The three hotspots for *Calochortus* species richness – the Transverse Ranges, Coast Ranges, and Sierra Nevada (Fig. 1)—correspond to those for all vascular plants native to California (45) and the CFP (73). This concordance reflects peaks in elevational, climatic, and edaphic heterogeneity in these areas, and perhaps immigration by lineages into these areas from other areas, paralleling those in *Calochortus*. Indeed, the Transverse Ranges hotspot (Fig. 1) lies astride the greatest concentration of major physiographic regions in the CFP, including the Peninsular Ranges, Coast Ranges, Central Valley, Sierra Nevada, and Great Basin—presumably facilitating immigration of species from several other regions. The greatest concentrations of *Calochortus* species within the three hotspots are associated with areas of extensive serpentine outcrops in the Bay Area, North and South Coast Ranges, Sierra Nevada, and westernmost Transverse Ranges, reflecting adaptation of *Calochortus* and other bulbous plants to open/partly open vegetation on such substrates (51, 52). Peaks of *Calochortus* diversity are absent from the Siskiyous, which include the largest serpentine outcrops in western North America, but the extensive, dense, evergreen forest canopies of the high-rainfall Siskiyous militate against *Calochortus*. Immigration appears to be the most likely explanation for high numbers of *Calochortus* in the central and eastern Transverse Range. Overall, our findings support conclusions by Baldwin (45) and Thorne et al. (73) that topographic and edaphic complexity (including serpentine) are major drivers of plant diversification and species richness in California. Our study adds chromosomal number evolution and migration as being of equal importance in *Calochortus* diversification.

Diversification in *Calochortus* may have direct parallels in other lineages. *Erythranthe* (part of *Mimulus* s. l.) also has a center of diversity in the CFP, occupies steep environmental gradients in mountainous areas, exhibits para- and peripatric speciation, displays serpentine tolerance and endemism, includes clades with a dysploid series of chromosome numbers, and has a rich history of ecological and genetic studies of speciation (74, 75). It would be exciting to study the drivers of *Erythranthe* diversification and local diversity to see whether patterns like those seen here in *Calochortus* emerge.

Methods

Sampling. We included 1 to 2 samples per species and subspecies; herbarium vouchers for new samples were deposited in the herbaria noted in *SI Appendix*, Table S9. Total genomic DNAs were extracted from silica-dried leaf or floral tissue using DNeasy plant kits (Qiagen, Valencia CA) following the manufacturer's instructions. We included all extant *Calochortus* species except extremely rare *C. rustvoldii* [known from only two sites/pixels in the western Transverse Ranges (76)]; including it would little affect our analyses.

Library Preparation. We prepared Illumina sequence libraries for 158 samples (156 *Calochortus* and 2 outgroups) following Lemmon et al. (77–79). We used a Covaris ultrasonicator (with reduced time for degraded samples) to fragment DNA to 140 to 400 bp; performed end-repair and A-tailing, ligated common Illumina adapters onto the template DNA ends using a Beckman Coulter FXp liquid-handling robot, and performed indexing PCR.

AHE Probe Design. Following Hamilton et al. (80) and Banker et al. (81), we developed hybrid enrichment probes for *Lobelia* and *Lilium*. We mapped sequences from two assembled transcriptomes–*Lobelia siphilitica* (from E. Carpenter) and *Lilium superbum* (from J. Leebens-Mack and C. dePamphilis)–to probe sequences from 27 references of the Angiosperm V1 AHE design (82, 83). Mapped transcriptome sequences were aligned to Angiosperm V1 reference sequences using MAFFT v7.023b (84). We used Geneious R9 (85) to visually inspect alignments, remove transcriptome sequences that were not clearly homologous, and trim transcriptome sequences to exons represented by the 27 reference sequences. Probes were tiled uniformly at 2.7× density. This probe set (Angiosperm V2 AHE) design contains 29 references (57,471 probes), including transcriptomes representing *Lobelia* and *Lilium*.

To improve phylogenetic resolution within *Calochortus*, we utilized data from two more species and expanded the targets into regions flanking the exons in Angiosperm V2 AHE. We first collected whole-genome sequence data (Illumina paired-end 200-bp protocol) for *Calochortus albus* (160M reads) and *C. flexuosus* (536M reads). We then used methods and scripts from Banker et al. (83) to identify loci and design probes. After merging overlapping reads following (86), we mapped merged reads to probe sequences from Angiosperm V2 AHE using Liliaceae as a reference. After extending consensus sequences into flanking regions using iterative mapping (80, 81), we aligned by locus the resulting consensus sequences to the reference, using Geneious to visually inspect alignments and trim poorly aligned regions from alignment ends. We removed 27 of the 517 alignments to ensure target loci did not overlap. After masking repetitive regions (80), final alignments contained 375,527 sites. Tiling 120 bp probes at 4.5× density for both references produced 16,600 probes. We used this probe set (AHE Cal1) to produce an Agilent Technologies Custom SureSelect XT kit for hybrid DNA enrichment.

Library Enrichment and Sequencing. We pooled indexed libraries in groups of 16 before enriching with the probe kit just described. Before sequencing, we pooled enriched libraries and assessed quality by Kapa qPCR (Roche). We sequenced samples on an Illumina NovaSeq6000 at Florida State with a PE-150 bp protocol and 8-bp (dual) indexing. After filtering poor-quality reads with the Illumina CASAVA v1.8 high-chastity filter and demultiplexing, we obtained an average of 5.4M read pairs per sample (~1.6 Gb).

Nuclear Assemblies and Alignment. We processed reads, corrected sequencing errors, and trimmed adaptors following (80, 81, 86). Assembly used a quasi-de novo approach, mapping reads to probe-region sequences for both *Calochortus* species in the Cal1 design. We used the resulting consensus sequences with ≥98× coverage to determine orthology (80) and then formed orthologous clusters and removed clusters containing <50% of the individuals, resulting in 294 loci analyzed. For each locus, we aligned sequences in Mafft v7.023 and then trimmed/masked alignments with an automated procedure (80) using default parameters (i.e., MINGOODSITES=14, PROPGOOD=0.5). We verified alignment quality by inspecting the alignments in Geneious. Upon manuscript acceptance, accession codes will be provided for raw reads in the Sequence Read Archive and sequences and alignments in GenBank.

Plastome Assemblies. We used getOrganelle (87) for de novo assembly of a *Calochortus venustus* plastome from Illumina short reads and Oxford Nanopore long reads (88). We used this plastome for subsequent reference-based assemblies. Raw reads were trimmed using Trimmomatic ver. 0.40 (89) with the simple clip threshold (2:30:10) for adaptors and quality-trimmed based on a 5-bp sliding window with a minimum Phred score of 20. Reads were mapped against the reference plastome using BWA with the *bwa mem* algorithm (90). The resulting bam files were sorted, PCR duplicates removed, SNPs phased, and merged using SAMtools ver. 1.3 (91). Resulting plastomes were aligned using MAFFT.

Phylogenetic Inference. We inferred maximum-likelihood phylogenies for plastomes, each nuclear locus, and concatenated nuclear genes using IQ-TREE (92) with 100 bootstrap replicates, implemented with ModelFinder (93) to select appropriate models of nucleotide substitution. Gene trees were used to infer a species tree with ASTRAL-III (94) after collapsing nodes in gene trees with <20% bootstrap support. We used IQ-TREE to estimate gCFs, representing the percentage of gene trees that support each clade, and sCFs representing the percentage of decisive sites supporting each clade (95). We used *Prosartes lanuginosa* and *Scoliopus bigelovii* as outgroups (see ref. 96).

Time-Calibrated Phylogeny. We used BEAST 2.6.6 (97) to calibrate the ASTRAL-III phylogeny against time employing five plastid genes (*atpB*, *psaA*, *psbD*, *rbcL*, *rps4*s) extracted from assembled plastomes using custom BLAST (98). Given the lack of fossil calibration points in Liliales, we spliced our *Calochortus* tree into the across-monocots chronogram (99) pruned to Liliales using *ape* v5.6 (100). The branching topology of *Calochortus* and closely related Liliaceae s. s., *Tricyrtis*, and *Streptopus* was modified to that of Lu et al. (96) based on complete plastomes. We enforced prior calibrations at crown Liliales (110.53 My) and Liliaceae (42.74 My), each with a SD of 2 My and a normal distribution based on the 95% CI of (99). The birth–death tree prior was applied with an uncorrelated relaxed lognormal clock and GTR model. Site model parameters included four gamma categories, a gamma shape distribution of 1.0, estimated proportion of invariant sites, and empirical frequencies. We conducted two independent runs, each with MCMC chain length of 100,000,000 iterations, logged every 10,000 generations. A maximum clade credibility tree was inferred after a burn-in of 50% using TreeAnnotator in BEAST 2.6.6. We repeated this process using the nuclear phylogeny as a constraint and reran the dating analysis with additional secondary priors assigned to the *Calochortus* stem and crown from the plastome chronogram (*SI Appendix*, Fig. S4). The resulting CI for the primary geographic clades between the nuclear and plastome chronograms were overlapping, so subsequent analyses used the nuclear tree only.

Diversification Rates. We applied BAMM (101) to our nuclear chronogram pruned to one exemplar per species, excluding subspecies, to infer shifts in diversification rates. Priors were generated using BAMMtools (102) for four independent MCMC chains run for 100M generations and logged every 10,000 generations. Convergence was assessed using CODA (103). We assessed 1 to 3 possible diversification shifts and compared resulting log likelihoods for the best-fit model.

Biogeographic Analyses. Ancestral biogeographic ranges were inferred using BioGeoBears (104). Discrete geographic areas were defined based on floristic affinities and geographic features (i.e., mountainous regions) to which species of *Calochortus* are generally restricted. These included A) Rocky Mountains, Great Basin, Columbia Plateau, Colorado Plateau, and Sonoran Desert; B) Sierra Nevada; C) Cascades, Klamath-Siskiyou Mountains, and Columbian Plateau; D) North and South Coast Ranges; E) Transverse and Peninsula Ranges, and F) Mexican mountains (Sierra Madre Occidental, Sierra Madre Oriental, Sierra Madre del Sur, Trans-Mexican Volcano Belt), and Chihuahuan Desert. We applied DEC (dispersal-extinction-cladogenesis), DEC+J [DEC plus jump-speciation (52)], and DIVALIKE [dispersal-vicariance analysis (105)] models in BioGeoBears using a time-stratified analysis with time periods at 3.5 My, 6.0 My and 10.5 My and adjusted dispersal probabilities (*SI Appendix*, Table S10) based on mountain-building history (*SI Appendix*, Table S8) and ages of faults causing the Peninsular Ranges to split and slide northward from the Sierra Madre Occidental (106) and of emergence of Sonoran-Desert climate in the southwest Great Basin (57); delay in the latter facilitated dispersal among the Peninsular/Transverse Ranges, Mexican mountains, and Great Basin. Ancestral taxa were allowed to occupy no more than three areas, and combinations of non-adjacent areas (e.g., Pacific Northwest and Mexico) were excluded. Log likelihoods and AIC scores were used for final model selection.

Ancestral Trait Estimation. Ancestral states for serpentine tolerance were inferred using the maximum-likelihood Markov-k model implemented in *corHMM* (107). We used equal rates, symmetric rates, and all rates different models of trait evolution. The best model for each trait was selected using Akaike weights derived from the AIC score of each model, with and without the root state estimated (108). Serpentine tolerance was scored (*SI Appendix*, Table S11) based on published sources (41–47, 109; *SI Appendix*, Table S11), including *serpentine endemics* (specialists restricted to serpentinite-derived soils) and *serpentine tolerators* (generalists found on both serpentinite and non-serpentinite soils) (58, 59).

Ancestral chromosome numbers were inferred using ChromEvol v2.2 (110, 111) given published chromosome numbers (*SI Appendix*, Table S11). Chromosome counts were available for 40 of 73 (55%) taxa in our phylogeny. Two species with reported intraspecific variation were assigned all counts. We implemented ten likelihood-based models of chromosomal evolution (mainType = All_Models) using the BEAST maximum credibility tree with all parameters set at default; the best-fit model was determined by comparing AIC values.

Species Occurrences and Environmental Data. We extracted *Calochortus* species occurrence records from GBIF (GBIF.org), including only records representing preserved specimens using *rgbif* ver. 3.7 (112). Records were cleaned using CoordinateCleaner (113). Misidentified or invalid points were pruned using isolation forests with an anomaly score threshold of 0.75 based on their environmental occupancy (see below) using the R package *solitude* (https://cran.r-project.org/web/packages/solitude/index.html). The remaining points were manually curated in QGIS (https://docs.qgis.org/3.22/en/docs/user_manual/index.html).

For each occurrence coordinate we compiled 42 environmental variables derived from topographic, soil, and climate data or custom indices (*SI Appendix*, Table S12). Elevation was obtained from the NASA Shuttle Radar Topography Mission at 90-m resolution; soil data, from the ISRIC SoilGrids global soil database, including % sand, silt, and clay, nitrogen content, and pH at 5- to 15-cm depths; temperature and precipitation, based on the 19 BioClim variables from WorldClim (http://www.worldclim.org/bioclim); the CGAIR-CSI global aridity index; P/PET (the ratio of mean annual precipitation to mean annual potential evapotranspiration); a drought index based on monthly VPDs and soil water contents; and several additional variables from the ENVIREM dataset, including seasonal PET and Thornthwaite aridity index. We created a custom *winter rainfall index* (% annual precipitation falling from November to March) and a custom *moisture seasonality index* (SD of mean precipitation/PET in each quarter).

Species Co-Occurrence and PD. Geographic grid cells were created across *Calochortus'* range, using the United States and Mexico as the geographic extent, and species occurrence records overlaid using the R package *sf* (114). We tested sensitivity to spatial scale by using grids of 0.25° and 0.1° resolution. Results were qualitatively similar; therefore, in most cases, we present only results from the 0.25° analyses. We quantified spatial overlap per grid cell between species pairs using the Pianka index (115) implemented in *spaa* ver. 0.2.2 (116). We calculated the geographic distance between each pair of species as the haversine distance of median latitude and longitude for their geographic ranges. We used linear regression to relate overlap and geographic distance to time

since divergence. Using overlap matrices at the 0.1° and 0.25° resolutions, we sorted entries by time since divergence from 0 to 1 My; binned those entries at 0.1 and 0.25 My intervals, chose the points with maximum overlap in each interval; and used linear regression to estimate maximum overlap as a function of time since divergence.

PD of grid cells was calculated as PD_{ses} – the standardized effect size of Faith's PD, which incorporates species richness and the relationships of the lineages present while factoring out null expectations for lineage diversity at the observed species richness – using the R package *picante* (117).

Peripatry, allopatry, and partial/complete sympatry) of sister species in the ASTRAL phylogeny were scored (*SI Appendix*, Table S5) using species ranges given by (41, 44, 109; see *SI Appendix*, Fig. S10 for distribution maps based on GBIF data). Peripatry involves abutting or close to abutting species ranges (<5% overlap in ranges or ranges closer than 100 km), in contrast to allopatry (no overlap in geographic range) or partial or complete sympatry (>5% overlap in geographic range).

Predictors of Species Richness. For each cell, we quantified species richness (the number of *Calochortus* species present based on GBIF data) and predictors of species richness after removing all but one correlated environmental variable of interest (justification of predictors is provided in the introductory paragraphs) among those with $|r| > 0.70$ based on Pearson's correlation coefficient. If multiple accessions of the same species existed within a given grid, we utilized the mean value of predictors to calculate within-cell ranges (maximum–minimum) of elevations occupied, soil parameters, aridity index (P/PET, where P is annual precipitation and PET is annual potential evapotranspiration), PET seasonal diversity, winter rainfall index (percent of annual rainfall occurring within November–March), moisture seasonality index, counts of different chromosome numbers among species present, and serpentine index (% of serpentine-tolerant species).

We constructed GLM in *lme4* (118) to evaluate predictors of species richness. We fit a Poisson error distribution, checked for overdispersion by plotting standardized deviance residuals against predicted counts, and calculated the McKelvery–Zavoina pseudo-R^2 (119, 120). To account for spatial autocorrelation (variables in adjacent cells being more similar than expected by chance), we tested for spatial autocorrelation using Moran's I, extracted model residuals, and added them to our GLM dataset to adjust for spatial autocorrelation via a SRM using *spData* (121) zero policy given some cells have no neighbors. We tested for the best-fit spatial model applying a lag model, error model, sac model, and spatial Durbin model.

Data, Materials, and Software Availability. Raw reads have been deposited in the NCBI Sequence Read Archive under BioProject PRJNA1069663 (122). Alignments, trees, and analysis scripts have been deposited in the Dryad Data Repository at DOI: 10.5061/dryad.kwh70rzbw (123). All other study data are included in the article and/or *SI Appendix*.

ACKNOWLEDGMENTS. We are grateful to Ron Parsons for providing photographs for Fig. 2 and thank those who contributed tissue samples: Deborah Ayres, Peggy Fiedler, Kristin Forgrave, Ryan Fuller, Mary Gerritsen, Matt Guillians, Susan Holiday, Patricia Jacquinth, Mary Louis Knapp, Ben Ledger, Larry Lodwick, Virginia Meyer, Mit Mcglaughlin, Kate Monsted, Gareth Pearson, Megan Poulette, Matt Shinderman, Lisa Stratton, Morgan Stickrod, and A. Walker. Sarah Friedrich polished the figures. This work was funded by NSF grant DEB 1929296 to T.J.G. and C.D.S.

Author affiliations: [a]Science and Conservation Division, Missouri Botanical Garden, St. Louis, MO 63110; [b]Department of Botany, University of Wisconsin-Madison, Madison, WI 53706; [c]Department of Ecology and Conservation Biology, Texas A&M University, College Station, TX 77845; [d]Department of Scientific Computing, Florida State University, Tallahassee, FL 32306; [e]Department of Biological Science, Florida State University, Tallahassee, FL 32306; [f]School of Integrative Plant Science, Cornell University, Ithaca, NY 14853; [g]L. H. Bailey Hortorium, Cornell University, Ithaca, NY 14853; [h]Departamento de Botánica y Zoología, Universidad de la Guadalajara, Zapopan, Jalisco 45200, Mexico; and [i]Boyce Thompson Institute for Plant Research, Ithaca, NY 14853

Author contributions: N.K. and T.J.G. designed research; D.S., A.R.L., E.M.L., E.E., A.I.H., P.W.C., A.R., J.B.L., and T.J.G. performed research; N.K., C.P.K., D.S., A.R.L., E.M.L., and E.E. analyzed data; and N.K. and T.J.G. wrote the paper, with additional contributions by C.P.K., A.R.L., E.M.L., S.R.S., and C.D.S.

1. E. Mayr, *Systematics and the Origin of Species* (Harvard University Press, Cambridge, MA, 1942).
2. D. Lack, *Darwin's Finches* (Cambridge University Press, Cambridge, UK, 1947).
3. T. J. Givnish, "Adaptive radiation and molecular systematics: Issues and approaches" in *Molecular Evolution and Adaptive Radiation*, T. J. Givnish, K. J. Sytsma, Eds. (Cambridge University Press, 1997), pp. 1–54.
4. D. Schluter, *The Ecology of Adaptive Radiation* (Oxford University Press, 2000).
5. P. R. Grant, B. R. Grant, Evolution of character displacement in Darwin's finches. *Science* **313**, 224–226 (2006).
6. D. Schluter, Ecology and the origin of species. *Trends Ecol. Evol.* **18**, 372–380 (2001).
7. J. A. Coyne, H. A. Orr, *Speciation* (Sinauer, 2004).
8. T. J. Givnish, Ecology of plant speciation. *Taxon* **59**, 1326–1366 (2010).
9. J. M. Sobel, G. F. Chen, L. R. Watt, D. W. Schemske, The biology of speciation. *Evol.* **64**, 295–315 (2010).
10. R. J. Rundell, T. D. Price, Adaptive radiation, nonadaptive radiation, ecological speciation and nonecological speciation. *Trends Ecol. Evol.* **24**, 394–399 (2009).
11. A. Stein, K. Gerstner, H. Kreft, Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. *Ecol. Lett.* **17**, 866–880 (2014).
12. A. L. Angert, T. E. Huxman, P. Chesson, D. L. Venable, Functional tradeoffs determine species coexistence via the storage effect. *Proc. Nat. Acad. Sci. U.S.A.* **106**, 11641–11645 (2009).
13. T. J. Givnish *et al.*, Origin, adaptive radiation, and diversification of the Hawaiian lobeliads (Asterales: Campanulaceae). *Proc. R. Soc. B* **276**, 407–416 (2009).
14. D. Spalink *et al.*, Biogeography of the cosmopolitan sedges (Cyperaceae) and the area-richness correlation in plants. *J. Biogeogr.* **43**, 1893–1904 (2016).
15. R. E. Ricklefs, F. He, Region effects influence local tree species diversity. *Proc. Nat. Acad. Sci. U.S.A.* **113**, 674–679 (2016).
16. M. Sundaram *et al.*, Accumulation over evolutionary time as a major cause of biodiversity hotspots in conifers. *Proc. Roy. Soc. B* **286**, 20191887 (2019).
17. T. J. Givnish *et al.*, Adaptive radiation, correlated and contingent evolution, and determin- ants of net species diversification in Bromeliaceae. *Mol. Phylogen. Evol.* **71**, 55–78 (2014).
18. L. P. Lagomarsino, F. L. Condamine, A. Antonelli, A. Mulch, C. C. Davis, The abiotic and biotic drivers of rapid diversification in Andean bellflowers (Campanulaceae). *New Phytol.* **210**, 1430–1442 (2016).
19. T. J. Givnish *et al.*, Orchid historical biogeography, diversification, Antarctica, and the paradox of orchid dispersal. *J. Biogeogr.* **43**, 1905–1916 (2016).
20. O. A. Pérez-Escobar *et al.*, Recent origin and rapid speciation of Neotropical orchids in the world's richest plant biodiversity hotspot. *New Phytol.* **215**, 891–905 (2017).
21. W. L. Testo, E. Sessa, D. S. Barrington, D. S., The rise of the Andes promoted rapid diversification in Neotropical *Phlegmariurus* (Lycopodiaceae). *New Phytol.* **222**, 604–613 (2019).
22. O. A. Pérez-Escobar *et al.*, The Andes through time: Evolution and distribution of Andean floras. *Trends Plant Sci.* **27**, 364–378 (2022).
23. C. S. Drummond, R. J. Eastwood, S. T. S. Miotto, C. E. Hughes, Multiple continental radiations and correlates of diversification in *Lupinus* (Leguminosae): Test for key innovation with incomplete taxon sampling. *Syst. Biol.* **61**, 443–460 (2012).
24. Y. Xing, R. H. Ree, Uplift-driven diversification in the Hengduan Mountains, a temperate biodiversity hotspot. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E3444–E3451 (2017).
25. J. S. Suissa, M. A. Sundue, W. L. Testo, Mountains, climate and niche heterogeneity explain global patterns of fern diversity. *J. Biogeogr.* **48**, 1296–1308 (2021).
26. N. A. Hazzi, J. S. Moreno, C. Ortiz-Movliav, R. D. Palaci, Biogeographic regions and events of isolation and diversification of the endemic biota of the tropical Andes. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 7985–7990 (2018).
27. L. M. Boschman, F. L. Condamine, Mountain radiations are not only rapid and recent: Ancient diversification of South American frog and lizard families related to Paleogene Andean orogeny and Cenozoic climate variations. *Glob. Planet. Change* **208**, 103704 (2022).
28. X. Y. Wang *et al.*, Phylogenomics reveals the evolution, biogeography, and diversification history of voles in the Hengduan Mountains. *Commun. Biol.* **5**, 1124 (2022).
29. A. García-Rodríguez *et al.*, Amphibian speciation rates support a general role of mountains as biodiversity pumps. *Am. Nat.* **198**, E68–E79 (2021).
30. J. Igea, A. J. Tanentz, Global topographic uplift has elevated speciation in mammals and birds over the last 3 million years. *Nat. Ecol. Evol.* **5**, 1530–1535 (2021).
31. C. Rahbek *et al.*, Humboldt's enigma: What causes global patterns of mountain biodiversity? *Science* **365**, 1108–1113 (2019).
32. H. Kreft, W. Jetz, Global patterns and determinants of vascular plant diversity. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 5925–5930 (2007).
33. J. Cavender-Bares, Diversification, adaptation, and community assembly of the American oaks (*Quercus*), a model clade for integrating ecology and evolution. *New Phytol.* **221**, 669–692 (2019).
34. J. Smyčka *et al.*, Tempo and drivers of plant diversification in the European mountain system. *Nat. Comm.* **13**, 2750 (2022).
35. D. C. Tank, R. G. Olmstead, From annuals to perennials: Phylogeny of subtribe Castillejinae (Orobanchaceae). *Am. J. Bot.* **95**, 608–625 (2008).
36. D. O. Burge *et al.*, Diversification of *Ceanothus* (Rhamnaceae) in the California Floristic Province. *Int. J. Plant Sci.* **172**, 1137–1164 (2011).
37. V. T. Parker, C. Y. Rodriguez, G. Wechsler, M. C. Vasey, Allopatry, hybridization, and reproductive isolation in *Arctostaphylos*. *Am. J. Bot.* **107**, 1798–1814 (2020).
38. S. Singhal *et al.*, Diversification, disparification and hybridization in the desert shrubs *Encelia*. *New Phytol.* **230**, 1228–1241 (2021).
39. A. D. Wolfe, P. D. Blischak, L. S. Kubatko, Phylogenetics of a rapid, continental radiation: Diversification, biogeography, and circumscription of the beardtongues (*Penstemon*; Plantaginaceae). bioRxiv [Preprint] (2021). https://doi.org/10.1101/2021.04.20.440652 (Accessed 20 December 2023).
40. I. H. Lichter-Marck, B. G. Baldwin, Edaphic specialization onto bare, rocky outcrops as a factor in the evolution of desert angiosperms. *Proc. Nat. Acad. Sci. U.S.A.* **120**, e2214729120 (2023).
41. M. Ownbey, A monograph of the genus *Calochortus*. *Ann. Mo. Bot. Gard.* **27**, 371–560 (1940).

42. T. B. Patterson, T. J. Givnish, Geographic cohesion, chromosomal evolution, parallel adaptive radiations, and consequent floral adaptations in *Calochortus* (Calochortaceae): Evidence from a cpDNA phylogeny. *New Phytol.* **161**, 253–264 (2004).
43. M. E. Gerritsen, R. Parsons, *Calochortus: Mariposa Lilies and Their Relatives* (Timber Press, Portland, OR, 2007).
44. P. L. Fiedler, R. K. Zebell, Calochortus. Flora North America. online. http://www.efloras. org/florataxon.aspx?flora_id=1&taxon_id=105173 (2012).
45. B. G. Baldwin, Origins of plant diversity in the California Floristic Province. *Ann. Rev. Ecol. Evol. Syst.* **45**, 347–369 (2015).
46. A. R. Kruckeberg, *California Serpentines: Flora, Vegetation, Geology, Soils, and Management Problems* (University of California Press, Berkeley, 1984).
47. H. Safford, J. E. D. Miller, An updated database of serpentine endemism in the California flora. *Madroño* **67**, 85–104 (2020).
48. California Native Plant Society, Rare plant inventory (v9-011.5, online edition, 2022), https://www.rareplants.cnps.org.
49. J. M. Henss, J. R. Moeller, T. J. Theim, T. J. Givnish, Spatial scales of genetic structure and gene flow in *Calochortus albus* (Liliaceae). *Ecol. Evol.* **3**, 1461–1470 (2013).
50. N. I. Cacho, S. Y. Strauss, Occupation of bare habitats, an evolutionary precursor to soil specialization in plants. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 15132–15137 (2014).
51. T. B. Patterson, T. J. Givnish, Phylogeny, concerted convergence, and phylogenetic niche conservatism in the core Liliales: Insights from rbcL and ndhF sequence data. *Evol.* **56** 233–252 (2002).
52. P. W. Rundel, Making sense of geophyte diversity. *Fremontia* **44**, 7–15 (2016).
53. P. Goldblatt, J. C. Manning, Plant diversity of the Cape Region of South Africa. *Ann. Mo. Bot. Gard.* **89**, 281–302 (2002).
54. R. H. Ree, I. Sanmartín, Conceptual and statistical problems with the DEC+J model of founder-event speciation and its comparison with DEC via model selection. *J. Biogeogr.* **45**, 741–749 (2018).
55. N. J. Matzke, Statistical comparison of DEC and DEC+J is identical to comparison of two CLaSSE submodels, and is therefore valid. *J. Biogeogr.* **49**, 1805–1824 (2022).
56. A. I. Hernández, J. B. Landis, C. D. Specht, Phylogeography and population genetics reveal ring species patterns in a highly polymorphic California lily. *J. Biogeogr.* **49**, 416–430 (2022).
57. I. Ferrusquía-Villafranca, L. I. Gonzáles-Guzmán, *Biodiversity, Ecosystems, and Conservation in Northern Mexico* (Oxford University Press, 2005).
58. B. L. Anacker, J. B. Whittall, E. E. Goldberg, S. P. Harrison, Origins and consequences of serpentine endemism in the California Flora. *Evol.* **65**, 365–376 (2011).
59. N. Rajakaruna, Lessons on evolution from the study of edaphic specialization. *Bot. Rev.* **84**, 39–78 (2018).
60. J. B. M. Sambatti, K. J. Rice, Local adaptation, patterns of selection, and gene flow in the California serpentine sunflower (*Helianthus exilis*). *Evol.* **60**, 696–710 (2006).
61. S. Y. Strauss, N. I. Cacho, Nowhere to run, nowhere to hide: The importance of enemies and apparency in adaptation to harsh soil environments. *Am. Nat.* **182**, E1–E14 (2013).
62. K. L. Powell, T. M. Knight, Effects of nutrient addition and competition on biomass of five *Cirsium* species (Asteraceae), including a serpentine endemic. *Int. J. Plant Sci.* **170**, 918–925 (2009).
63. S. A. Sianta, K. M. Kay, Adaptation and divergence in edaphic specialists and generalist: Serpentine soil endemics in the California flora occur in barer serpentine habitats with lower calcium levels than serpentine tolerators. *Am. J. Bot.* **106**, 690–703 (2019).
64. L. A. Stockmeier, T. J. Givnish, Plant distribution, stature, rarity, and diversity in a patterned calcareous fen: Tests of geochemical and leaf-height models. *Am. J. Bot.* **106**, 807–820 (2019).
65. B. L. Anacker, S. Y. Strauss, The geography and ecology of plant speciation: Range overlap and niche divergence in sister species. *Proc. R. Soc. B* **281**, 20132980 (2014).
66. S. Harrison, *Plant and Animal Endemism in California* (University of California Press, Berkeley, CA, 2013).
67. S. V. Drovetski *et al.*, Geographic mode of speciation in a mountain specialist avian family endemic to the Palearctic. *Ecol. Evol.* **3**, 1518–1528 (2013).
68. Y. Qi, Y. Yang, Topographic effect on spatial variation of plant diversity in California. *Ann. GIS* **5**, 39–46 (1999).
69. L. T. Lancaster, K. M. Kay, Origin and diversification of the California flora: Re-examining classic hypotheses with molecular phylogenies. *Evol.* **67**, 1041–1054 (2013).
70. C. L. Mohler, Co-occurrence of oak subgenera: Implications for niche differentiation. *Bull. Torrey Bot. Club* **117**, 247–255 (1990).
71. J. Cavender-Bares *et al.*, Phylogeny and biogeography of the American live oaks (*Quercus* subsection *Virentes*): A genomic and population genetics approach. *Mol. Ecol.* **24**, 3668–3687 (2015).
72. L. J. Pollock, M. J. Bayly, P. A. Vesk, The roles of ecological and evolutionary processes in plant community assembly: The environment, hybridization, and introgression influence co- occurrence of *Eucalyptus*. *Am. Nat.* **185**, 784–796 (2015).
73. J. H. Thorne, J. H. Viers, J. Price, D. M. Storms, Spatial patterns of endemic plants in California. *Nat. Areas J.* **29**, 344–366 (2009).
74. P. M. Beardsley, S. E. Schoenig, J. B. Whittall, R. G. Olmstead, Patterns of evolution in western North American *Mimulus* (Phyrmaceae). *Am. J. Bot.* **91**, 474–489 (2004).
75. J. M. Sobel, Ecogeographic isolation and speciation in the genus *Mimulus*. *Am. Nat.* **184**, 565–579 (2014).
76. F. Callahan, *Calochortus rustvoldii* Callahan (Liliaceae), a new species from Los Angeles and Ventura counties, California. *Phytologia* **97**, 282–285 (2015).
77. A. R. Lemmon, S. A. Emme, E. M. Lemmon, Anchored hybrid enrichment for massively high-throughput phylogenetics. *Syst. Biol.* **61**, 721–744 (2012).
78. R. O. Prum *et al.*, A fully resolved, comprehensive phylogeny of birds (Aves) using targeted next generation DNA sequencing. *Nature* **526**, 569–573 (2015).
79. M. Meyer, M. Kircher, Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protoc.* **2010**, pdb.prot5448 (2010), 10.1101/pdb.prot5448.
80. C. A. Hamilton, A. R. Lemmon, E. Moriarty Lemmon, J. E. Bond, Expanding anchored hybrid enrichment to resolve both deep and shallow relationships within the spider Tree of Life. *BMC Evol. Biol.* **16**, 212 (2016).
81. S. E. Banker *et al.*, Hierarchical hybrid enrichment: Multitiered genomic data collection across evolutionary scales, with application to chorus frogs (*Pseudacris*). *Syst. Biol.* **69**, 756–773 (2020).
82. C. E. Buddenhagen *et al.*, Anchored phylogenomics of angiosperms I: Assessing the robustness of phylogenetic estimates. bioRxiv [Preprint] (2016). https://doi.org/10.1101/086298 (Accessed 20 December 2023).
83. N. Mitchell, P. O. Lewis, E. Moriarty Lemmon, A. R. Lemmon, K. E. Holsinger, Anchored phylogenomics resolves the evolutionary relationships in the rapid radiation of Protea L. (Proteaceae). *Am. J. Bot.* **104**, 102–115 (2017).
84. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
85. M. Kearse *et al.*, Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinform.* **28**, 1647–1649 (2012).
86. D. A. Rokyta, A. R. Lemmon, M. J. Margres, K. Arnow, The venom-gland transcriptome of the eastern diamondback rattlesnake (*Crotalus adamanteus*). *BMC Genomics* **13**, 312 (2012).
87. J. J. Jin *et al.*, GetOrganelle: A fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Gen. Biol.* **21**, 241 (2020).
88. G. Y. De la Cerda *et al.*, Balancing read length and sequencing depth: Optimizing nanopore long-read sequencing for monocots with an emphasis on the Liliales. *Appl. Plant Sci.* **11**, e11524 (2023).
89. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinform.* **30**, 2114–2120 (2014).
90. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinform.* **25**, 1754–1760 (2009).
91. H. Li *et al.*, The sequence alignment/map format and SAMtools. *Bioinform.* **25**, 2078–2079 (2009).
92. L. T. Nguyen, H. A. Schmidt, A. von Haeseler, B. Q. Minh, IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
93. S. Kalyaanamoorthy, B. Q. Minh, T. K. Wong, A. von Haeseler, L. S. Jermiin, Model- Finder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
94. C. Zhang, M. Rabiee, E. Sayyari, S. Mirarab, ASTRAL III: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinform.* **19**, 15–30 (2018).
95. B. Q. Minh, M. W. Hahn, R. Lanfear, M. Rosenberg, New methods to calculate concord- ance factors for phylogenomic datasets. *Mol. Biol. Evol.* **37**, 2727–2733 (2020).
96. R. S. Lu *et al.*, Comparative plastome genomics and phylogenetic analyses of Liliaceae. *Bot. J. Linn. Soc.* **196**, 274–293 (2021).
97. R. Bouckaert *et al.*, BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLOS Comp. Biol.* **15**, e1006650 (2019).
98. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
99. T. J. Givnish *et al.*, Monocot plastid phylogenomics, timeline, net rates of species diversification, the power of multi-gene analyses, and a functional model for the origin of mono- cots. *Am. J. Bot.* **105**, 1888–1910 (2018).
100. E. Paradis, K. Schliep, ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinform.* **35**, 526–528 (2019).
101. D. L. Rabosky, S. C. Donnellan, M. Grundler, I. J. Lovette, Analysis and visualization of complex macroevolutionary dynamics: An example from Australian scincid lizards. *Syst. Biol.* **63**, 610–627 (2014).
102. D. L. Rabosky *et al.*, BAMMtools: An R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods Ecol. Evol.* **5**, 701–707 (2014).
103. M. Plummer, N. Best, K. Cowles, K. Vines, CODA: Convergence diagnosis and output analysis for MCMC. *R News* **6**, 7–11 (2006).
104. N. J. Matzke, Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. *Syst. Biol.* **63**, 951–970 (2014).
105. F. Ronquist, Dispersal-vicariance analysis: A new approach to the quantification of historical biogeography. *Syst. Biol.* **46**, 195–203 (1997).
106. P. J. Umhoefer, Why did the Southern Gulf of California rupture so rapidly? Oblique divergence across hot, weak lithosphere along a tectonically active margin. *GSA Today* **21**, 4–10 (2011).
107. J. M. Beaulieu, B. C. O'Meara, M. J. Donoghue, Identifying hidden rate changes in the evolution of a binary morphological character: The evolution of plant habit in campanulid angiosperms. *Syst. Biol.* **62**, 725–737 (2013).
108. R. G. Fitzjohn, W. P. Maddison, S. P. Otto, Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. *Syst Biol.* **58**, 595–611 (2009).
109. The Calflora Database, Calflora: Information on California plants for education, research and conservation (The Calflora Database, Berkeley, CA, 2023), https://www.calflora.org/.
110. I. Mayrose, M. S. Barker, S. P. Otto, Probabilistic models of chromosome number evolution and the inference of polyploidy. *Syst. Biol.* **59**, 132–144 (2010).
111. L. Glick, I. Mayrose, ChromEvol: Assessing the pattern of chromosome number evolution and the inference of polyploidy along a phylogeny. *Mol. Biol. Evol.* **31**, 1914–1922 (2014).
112. S. A. Chamberlain, C. Boettiger, R Python, and Ruby clients for GBIF species occurrence data. *PeerJ Preprints* **5**, e3304v1 (2017), 10.7287/peerj.preprints.3304v1.
113. A. Zizka *et al.*, CoordinateCleaner: Standardized cleaning of occurrence records from biological collection databases. *Methods Ecol. Evol.* **10**, 744–751 (2019).
114. E. Pebesma, Simple features for R: Standardized support for spatial vector data. *R J.* **10**, 439–446 (2018).
115. E. R. Pianka, The structure of lizard communities. *Ann. Rev. Ecol. Syst.* **4**, 53–74 (1973).
116. J. L. Zhang, K. P. Ma, *spaa:* An R package for computing species association and niche overlap. *Res. Progr. Biodiv. Conserv. China* **10**, 165–174 (2014).
117. S. W. Kembel *et al.*, *Picante*: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**, 1463–1464 (2010).
118. D. Bates, M. Mächler, B. M. Bolker, S. C. Walker, Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
119. R. McKelvery, R. Zavoina, A statistical model for the analysis of ordinal level dependent variables. *J. Math. Sociol.* **4**, 103–120 (1975).
120. M. R. Veall, K. F. Zimmermann, Evaluating Pseudo-R^2s for binary probit models. *Quality Quant.* **28**, 151–164 (1994).
121. R. S. Bivand, E. Pebesma, V. Gómez-Rubio, *Applied Spatial Data Analysis with R* (Springer, New York, 2013).
122. N. Karimi, Chromosomal evolution environmental heterogeneity and migration drive spatial patterns of species richness in Calochortus (Liliaceae) [Dataset]. NCBI SRA BioProject Accession PRJNA1069663. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1069663.Deposited 26 January 2024.
123. N. Karimi, Chromosomal evolution environmental heterogeneity and migration drive spatial patterns of species richness in Calochortus (Liliaceae) [Dataset]. Dryad. https://doi.org/10.5061/dryad.kwh70rzbw. Deposited 26 January 2024.