Phylogenetics of a Rapid, Continental Radiation: Diversification, Biogeography, and Circumscription of the Beardtongues (*Penstemon*; Plantaginaceae)

Andrea D. Wolfe[1,*], Paul D. Blischak[1,3,4], and Laura S. Kubatko[1,2]

[1]Department of Evolution, Ecology, & Organismal Biology and [2]Department of Statistics, The Ohio State University, Columbus, OH, USA, 43210. [3]Department of Ecology & Evolutionary Biology and [4]Department of Molecular & Cellular Biology, University of Arizona, Tucson, AZ, USA, 85721.

[*]***Corresponding author:*** *Andrea D. Wolfe, Department of Evolution, Ecology, and Organismal Biology, The Ohio State University, 318 W. 12th Avenue, Columbus, OH, 43210, USA; E-mail: wolfe.205@osu.edu.*

Running Head: PHYLOGENETICS OF THE *PENSTEMON* RADIATION

Abstract.—Penstemon (Plantaginaceae), the largest genus of plants native to North America, represents a recent continental evolutionary radiation. We investigated patterns of diversification, phylogenetic relationships, and biogeography, and determined the age of the lineage using 43 nuclear gene loci. We also assessed the current taxonomic circumscription of the ca. 285 species by developing a phylogenetic taxonomic bootstrap method. *Penstemon* originated during the Pliocene/Pleistocene transition. Patterns of diversification and biogeography are associated with glaciation cycles during the Pleistocene, with the bulk of

diversification occurring from 1.0–0.5 mya. The radiation across the North American continent tracks the advance and retreat of major and minor glaciation cycles during the past 2.5 million years with founder-event speciation contributing the most to diversification of *Penstemon*. Our taxonomic bootstrap analyses suggest the current circumscription of the genus is in need of revision. We propose rearrangement of subgenera, sections, and subsections based on our phylogenetic results. Given the young age and broad distribution of *Penstemon* across North America, it offers an excellent system for studying a rapid evolutionary radiation in a continental setting.

Key words: biogeography, diversification, evolutionary radiation, *Penstemon*, taxonomy

Penstemon Schmidel (Plantaginaceae), commonly known as the beardtongues, is the largest plant genus endemic to the North American continent, containing ca. 285 species, with new taxa being described every few years (Turner 2010; Estes 2012; O'Kane and Heil 2014; Zacarías-Correa et al. 2019). The recent publication of the Flora of North America treatment for the genus (Freeman 2019) included 239 species found north of the US/Mexico border. Mexico has ca. 60 species (Zacarías-Correa 2020). This large and diverse genus is an example of a rapid evolutionary radiation, with much of the diversification hypothesized to have taken place during the Pleistocene (Wolfe et al. 2006). Most of the species are found in the Intermountain West (Holmgren 1984; Holmgren and Holmgren 2016), with the number of species decreasing dramatically east of the Rocky Mountains. For example, the Great Plains region to the Mississippi River contain ca. 39 species, with an additional 10 species found only east of the Mississippi (Lindgren and Wilde 2003). In contrast, Utah has 76 species, with ca. 29 percent endemic to the state (Holmgren 1984; Stevens et al. 2020).

Species of *Penstemon* occur in a wide variety of habitats, including edaphic specialization (e.g., deep sand, limestone derived soils, igneous soils, oil shales), with most species adapted to xeric landscapes (Lindgren and Wilde 2003; Dockter et al. 2013). The diversity of floral shapes, sizes, and colors suggests a significant role of pollinator selective pressure in the evolutionary history of the genus (Pennell 1935; Straw 1956a, 1956b, 1963). Other morphological traits with high levels of variation include size, habit, inflorescence, leaves, anthers, and staminodes. Most of the species are restricted to narrow ranges, with about 1/3 of the species restricted to a single state in the United States or Mexico. Habitat specialization, combined with narrow geographic distribution, has resulted in conservation concerns for many species (Wolfe et al. 2014, 2016; Rodriguez-Peña et al. 2018; Stone et al. 2019; Zacarías-Correa

et al. 2020). *Penstemon penlandii*, *P. haydenii,* and *P. debilis* have been Federally listed as endangered or threatened under the US Endangered Species Act (USFWS 2011), with more than 90 species under consideration for protection (USFWS 1993).

The taxonomy of the genus has been developed over the past 250 years, since its recognized designation by Schmidel (1763). Earlier work included a description by John Mitchell, which was published in 1748 (Straw 1966), but his work is not currently recognized as establishing the genus (Freeman 2019). The most comprehensive listings of *Penstemon* taxonomy have been compiled by the American Penstemon Society in a series of publications that were most recently updated nearly 20 years ago (Lodewick and Lodewick 1987; Lindgren and Wilde 2003). Because the Flora of North America treatment by Freeman (2019) did not address the taxonomy of the genus, we will refer to the taxonomic designations from Lindgren and Wilde (2003), amended with information from the updated *Penstemon* treatments in Volumes Four and Seven of *The Intermountain Flora* (Holmgren 1984; Holmgren and Holmgren 2016).

The most comprehensive phylogenetic study for *Penstemon* was published by Wolfe et al. (2006) and included 163 species representing all subgenera and sections of the genus. This study used ITS and two cpDNA loci and was able to establish some major trends in the evolutionary patterns for the genus such as 1) the rapid evolutionary radiation of a large genus in a continental setting, 2) biogeographic trends across North America, 3) a probable role in diversification due to hybridization, and 4) the independent derivation of a hummingbird pollination syndrome in at least 10 lineages. Taxonomic trends observed in the 2006 study included strong support for the circumscription of subgenus *Dasanthera* as the earliest diverging lineage, and close affinities among some sections of subgenera *Saccanthera* and *Penstemon*, and

Penstemon and *Habroanthus*. The resolution of the backbone for the tree was insufficient to fully resolve taxonomic relationships, but indications for non-monophyly for three of the four multi-taxa subgenera were present.

Recent phylogenetic studies for *Penstemon* have focused on the core group of species encompassing sections *Gentianoides*, *Coerulei*, *Spectabiles,* and members of what has traditionally been known as subgenus *Habroanthus* (Crosswhite 1967; Wessinger et al. 2016, 2019). The Wessinger et al. (2016, 2019) studies employed multiplexed shotgun genotyping (MSG) SNP data for 75 and 120 species, respectively. The phylogenies based on MSG data had more resolution than did the trees presented in the Wolfe et al. (2006) study but were mostly in agreement with clade topologies of the earlier tree, and the trees were specifically used as a framework for understanding the evolutionary trends in the shift between bee- to bird-pollination syndromes.

In the Flora of North America, including most of the species of *Penstemon*, Freeman (2019) revised the taxonomy of genus to include only two subgenera (*Dasanthera* and *Penstemon*), based on the phylogenetic results from Wolfe et al. (2006). In addition to subsuming four subgenera, the number of sections was reduced, and there were no divisions below the sectional level. This resulted in the small subgenus *Dasanthera* circumscribing only nine species including *P. personatus*, formerly placed in a monotypic subgenus (*Cryptostemon*), and the remainder of the genus circumscribing subgenus *Penstemon*.

In this study we expand the phylogenetic survey to include 239 species of *Penstemon* with the intent to examine the correspondence of historical and contemporary taxonomic boundaries. With the expanded phylogenetic survey, we also address the age of the genus,

patterns of diversification, and biogeographic trends, and we compare our results to previous phylogenetic surveys for *Penstemon*.

MATERIALS AND METHODS

Sample Collection, DNA Extraction, and Amplicon Sequencing

DNA was extracted from a combination of either field-collected, silica-dried leaf tissue or leaves sampled from herbarium specimens using a modified CTAB protocol for DNA isolation (Wolfe 2005). After extraction, all samples were quantified using a Qubit fluorometer (Invitrogen, Carlsbad, CA, USA) and normalized to a concentration of 20 ng/μL. Normalized DNA samples for the 282 accessions of *Penstemon* representing 239 species plus two hybrids (ca. 84% of the genus) used in this study, plus nine accessions from other members of the tribe Cheloneae (one *Pennelianthus*, five *Keckiella*, one *Nothochelone*, one *Chelone*, and one *Chionophila*), were sent to the IBEST Genomics Resource Core at the University of Idaho (Moscow, ID, USA) for sample preparation and amplicon sequencing (Appendix 1). All subgenera, sections, and subsections of *Penstemon* were represented in the sampling (Table 1). Amplification of targeted regions and the addition of sample barcodes and Illumina adapters was done using microfluidic PCR on the Fluidigm 48x48 Access Array (Fluidigm Corporation, South San Francisco, CA, USA), followed by 300 bp, paired-end sequencing on an Illumina MiSeq (Illumina, San Diego, CA, USA) (Uribe-Convers et al. 2016). Primers for the 48 loci used in this study were designed and tested as described in Blischak et al. (2014) and are given in Table S1. Raw, paired-end sequencing reads were demultiplexed with dbcAmplicons [https://github.com/msettles/dbcAmplicons] prior to being returned from IBEST (Uribe-Convers

Table 1. Taxonomy of *Penstemon* as recognized by the American Penstemon Society, with annotations from the Freeman (2019) Flora of North America Treatment. Sampling of species for this study is indicated in the numerator, and abbreviations are those used in Figure 1.

Subgenus	Section	Subsection	Abbreviations	Species Sampling	FNA designations
Cryptostemon			C	1/1	Subg. Dasanthera
Dasanthera			D	9/9	
Dissecti			Di	1/1	
Habroanthus	*Elmigera*		H, El	6/7	
	Glabri		H, Gl	34/44	
Penstemon	*Ambigui*		P, A	1/2	
	Baccharifolii		P, B	1/1	
	Chamaeleon		P, Ch	3/4	
	Coerulei		P, Co	20/20	
	Cristati		P, Cr	24/28	3 spp transferred from sect. *Ericopsis*
	Ericopsis	*Caespitosi*	P, E, Ca	12/12	sect. Caespitosi (2 spp transferred to sect. *Cristati)*
		Ericopsis	P, E, E	1/1	sect. *Cristati*
		Linarioides	P, E, L	3/3	sect. *Caespitosi*
	Fasciculus	*Campanulati*	P, F, Cp	5/8	
		Fasciculi	P, F, Fa	10/13	
		Perfoliati	P, F, Pf	1/3	
		Racemosi	P, F, R	4/4	
	Peltanthera	*Centranthifolii*	P, Pe, Ce	10/10	sect. *Gentianioides*
		Havardiani	P, Pe, Hv	1/3	
		Peltanthera (Spectabiles)	P, Pe, Sp	14/16	sect. *Spectabiles*
		Petiolati	P, Pe, Pt	1/1	sect. *Petiolati*
	Penstemon	*Arenarii*	P, P, Ar	1/2	sect. *Penstemon*
		Deusti	P, P, De	3/3	
		Gairdneriani	P, P, G	2/2	
		Harbouriani	P, P, Ha	1/1	
		Humiles	P, P, Hu	15/20	
		Multiflori	P, P, M	1/1	
		Penstemon	P, P, P	10/18	
		Proceri	P, P, Pr	16/17	
		Tubaeflori	P, P, T	1/1	
Saccanthera	*Bridgesiani*		S, B	1/1	
	Saccanthera	*Heterophylli*	S, S, He	20/22	sect. *Saccanthera*
		Serrulati	S, S, Se	6/6	Sect. *Saccanthera*
				Total = 239/285	

et al. 2016). We then processed the sequence data using Fluidigm2PURC v0.2.1 [https://github.com/pblischak/fluidigm2purc] (Blischak et al. 2018) to assemble and align haplotypes for phylogenetic inference (Rothfels et al. 2017). The initial, unfiltered haplotype data returned by Fluidigm2PURC was further processed in Geneious v8.1.9 (Kearse et al. 2012). Each locus was first realigned in Geneious using MUSCLE with default settings (Edgar 2004), after which we removed poorly aligned sequences, as well as a small number of duplicated accessions that were repeated across sequencing runs.

Phylogenetic Inference

To infer a phylogeny for *Penstemon*, we used a coalescent-based approach. We inferred individual gene trees for each locus using FastTree v2.1.11 (Price et al. 2010), which we then used as input for ASTRAL v5.6.3 (ASTRAL-III; Zhang et al. 2018). Gene trees were inferred using the GTR+GAMMA model and ASTRAL-III was run using default options. We also ran a concatenation-based, maximum likelihood (ML) analysis using RAxML v8.2.10 (Stamatakis 2014). For our ML analysis, we concatenated all processed and aligned sequences using Phyutility v2.7.1 (Smith and Dunn 2008), fit separate GTR+GAMMA parameters for each gene using a partitions file, and performed 1000 rounds of rapid bootstrapping to assess support (Stamatakis et al. 2008). Previous phylogenetic work has placed *Pennellianthus frutescens* as sister to the rest of the Cheloneae, and we used it here as an outgroup for both analyses (Wolfe et al. 2002, 2006).

Taxonomic Bootstrap

To assess support for the current taxonomic classification of *Penstemon* at the level of subgenus, section, and subsection, we performed bootstrap resampling of taxa from within their respective taxonomic ranks. For an individual bootstrap replicate, a single representative was sampled from each named group at the chosen taxonomic rank and a phylogeny was inferred using these individuals' sequence data. At each taxonomic level (subgenus, section, and subsection), we performed 1000 rounds of bootstrap resampling with phylogenetic inference conducted using SVDQuartets (Chifman and Kubatko 2014) on each bootstrapped data set. Replicates were then combined into a majority rule extended consensus tree using RAxML v8.2.10 (Stamatakis 2014). The Python script for performing this 'taxonomic bootstrapping' is available in Supplemental Materials on Dryad (*sample_from_taxonomy.py*).

Time Calibration and Analysis of Diversification Rates

To obtain a time-calibrated phylogeny, we first used BEAST v2.6.0 with fossils for previously inferred divergence events for families within Lamiales from the literature as calibration points to date the splits between lineages in the tribe Cheloneae (Vargas et al. 2014). To do this, we downloaded publicly available sequences for the ITS region for three pairs of species with fossils supporting their minimum divergence time: *Kigelia africana* (Bignoniaceae) + *Catalpa duclouxii* (Bignoniaceae) [28.4 mya], *Veronica persica* (Plantaginaceae) + *Plantago lanceolata* (Plantaginaceae) [11.6 mya], and *Gratiola neglecta* (Plantaginaceae) + *Bacopa eisenii* (Plantaginaceae) [5.3 mya]. We then combined the sequences from these six species with the ITS region for samples from Cheloneae (Wolfe et al. 2002, 2006). In addition, we included sequences for trnC-D and trnT-L for all included Cheloneae samples (Table S1). Each locus was aligned

177 with MAFFT (Katoh 2013) using the '--auto' option and then all genes were concatenated with
178 Phyutility (Smith and Dunn 2008).

179 Program settings for the analysis with BEAST were specified within BEAUti (Bouckaert
180 et al. 2019). Each locus was given an independent substitution model (GTR + Gamma with four
181 rate categories) and clock model (relaxed log-normal; Drummond et al. 2006) but were
182 constrained to all follow a single tree topology. For the three fossil calibrations, we used log-
183 normal distributions with the following ranges: mean = 3.346 and standard deviation = 0.2 (95%
184 quantile range: 20.4–39.4 my) for *K. africana* + *C. duclouxii*; mean = 2.45 and standard
185 deviation = 0.2 (95% quantile range: 8.34–16.1 my) for *V. persica* + *P. lanceolata*; and mean =
186 1.667 and standard deviation = 0.2 (95% quantile range: 3.81–7.36 my) for *G. neglecta* + *B.*
187 *eisnii*. All other prior specifications were left at their default options. Parameters were sampled
188 for 2 x 10^7 generations and were logged every 1,000 generations. Posterior samples were
189 examined in Tracer v1.7.1 (Rambaut et al. 2018) to assess convergence and to estimate effective
190 sample sizes (ESS). Results were then summarized into a maximum clade credibility (MCC) tree
191 using TreeAnnotator after discarding 25% of the samples as burn in.

192 After estimating divergence times for the major Cheloneae lineages with BEAST, we
193 used three of the 95% highest posterior density estimates as secondary calibration bounds to date
194 the entire 293 taxon RAxML phylogeny using treePL (Smith and O'Meara 2012). The secondary
195 calibration points were defined for the most recent common ancestors for *Pennellinathus*
196 *frutenscens* + *Chelone glabra* (4.375–19.04 my), *K. breviflora* + *P. montanus* (1.89–8.539 my),
197 and *P. montanus* + *P. personatus* (1.419–6.416 my). We first ran treePL with the 'prime' option
198 to find the best optimization settings. Then, we used these settings, along with the 'thorough'
199 option, to estimate divergence times across the tree.

Using this time-calibrated phylogeny, we also estimated diversification rates with BAMM v2.5.0 (Rabasky 2014). Priors for the expectedNumberOfShifts, lambdaInitPrior, lambdaShiftPrior, and muInitPrior distributions were set using the setBAMMprior function in the BAMMtools R package v2.1.7 (Rabosky et al. 2014). All other priors were left at their default value. We then sampled parameters for 1 x 10^6 generations, logging them every 1,000 generations. After discarding 10% of the samples as burn-in, the remaining posterior samples were then processed in R v3.6.1 (R Core Team 2019) using the coda package (Plummer et al. 2006) to assess convergence and the BAMMtools package to summarize estimated rate shifts, net diversification rates, and sampled placements of shifts on the phylogeny. We also estimated diversification dynamics using MEDUSA v1.41 (Alfaro et al. 2009), implemented in the R package geiger v2.0.6.4 (Pennell et al. 2014), assuming a birth-death model.

Analysis of Biogeographic Distribution

To estimate biogeographic patterns in *Penstemon*, we used BioGeoBEARS v1.1.2 (Matzke 2013a) to infer ancestral areas as well as to compare competing biogeographic hypotheses using model selection criteria (e.g., Akaike information criterion, AIC; Akaike 1974). Within BioGeoBEARS, we fit the DIVALIKE (Ronquist 1997), DEC (Ree and Smith 2008), and BAYAREALIKE (Landis et al. 2013) models, along with their jump dispersal ("+J") counterparts, for a total of six models (Matzke 2013b, 2014). Areas were defined initially by the 12 regions from Wolfe et al. (2006) with updates based on the physiographic regions of North America (Fig. S1; Table 2; Barton et al. 2003). These were split into three separate sets to help with computational feasibility and to reflect geographic proximity of regions. The first set of analyses included areas [1-6], the second set included areas [4-7,11,12], and the third set

Table 2. Biogeographic regions for *Penstemon* (modified from Wolfe et al. 2006) used in Figure 1, and for BioGeoBears analyses.

Region Description	Designation
Intermontane Plateau – extreme north	1
Pacific Northwest, exclusive of Eastern and Western Cordilleras	2
Cascade-Sierra and Northwestern Cordillera	3
Southwestern Cordillera	4
Intermountain Region	5
Eastern Cordillera	6
Great Plains	7
Interior Lowlands	8
Appalachian Mountain System	9
Coastal Plains	10
American Southwest	11
Mexican Highlands	12

included areas [6-11]. The maximum number of ancestral areas for each analysis was set as the maximum number of areas currently occupied by the sampled taxa in a given set and all other parameters were kept at their default values. R code for running these analyses is available on dryad.

To further dissect evolutionary pattern and process for *Penstemon*, we also downloaded distribution information for every species of *Penstemon* available in the SEINet data portal (http://www.swbiodiversity.org). Only Federally- or State-listed species were excluded from our data set because these locations are restricted from public access. This information was collected as Keyhole Markup Language (KML) files which were then imported into Google Earth Pro (https://www.google.com/earth/versions/#earth-pro). Using our time-calibrated phylogeny, we plotted two patterns in Google Earth Pro: 1) the individual patterns of distribution of taxa from the earliest to most recently diverged lineages, and 2) a collective distribution map based on time of divergence estimations for the terminal lineages in each clade (Table S2). These time slices were 2.5 mya, 2.0 mya, 1.8 mya, and forward to 0.1 mya in 100,000-year increments. The time

of divergence was rounded up or down to fit into a time increment. Animations of the distributions in these two map sets were assembled from the individual or collective map images. The pattern of these species' distributions was then examined in the context of inferred glaciation cycles during the Pleistocene (Ehlers and Gibbard 2007; Ehlers et al. 2018).

RESULTS

Amplicon Sequencing

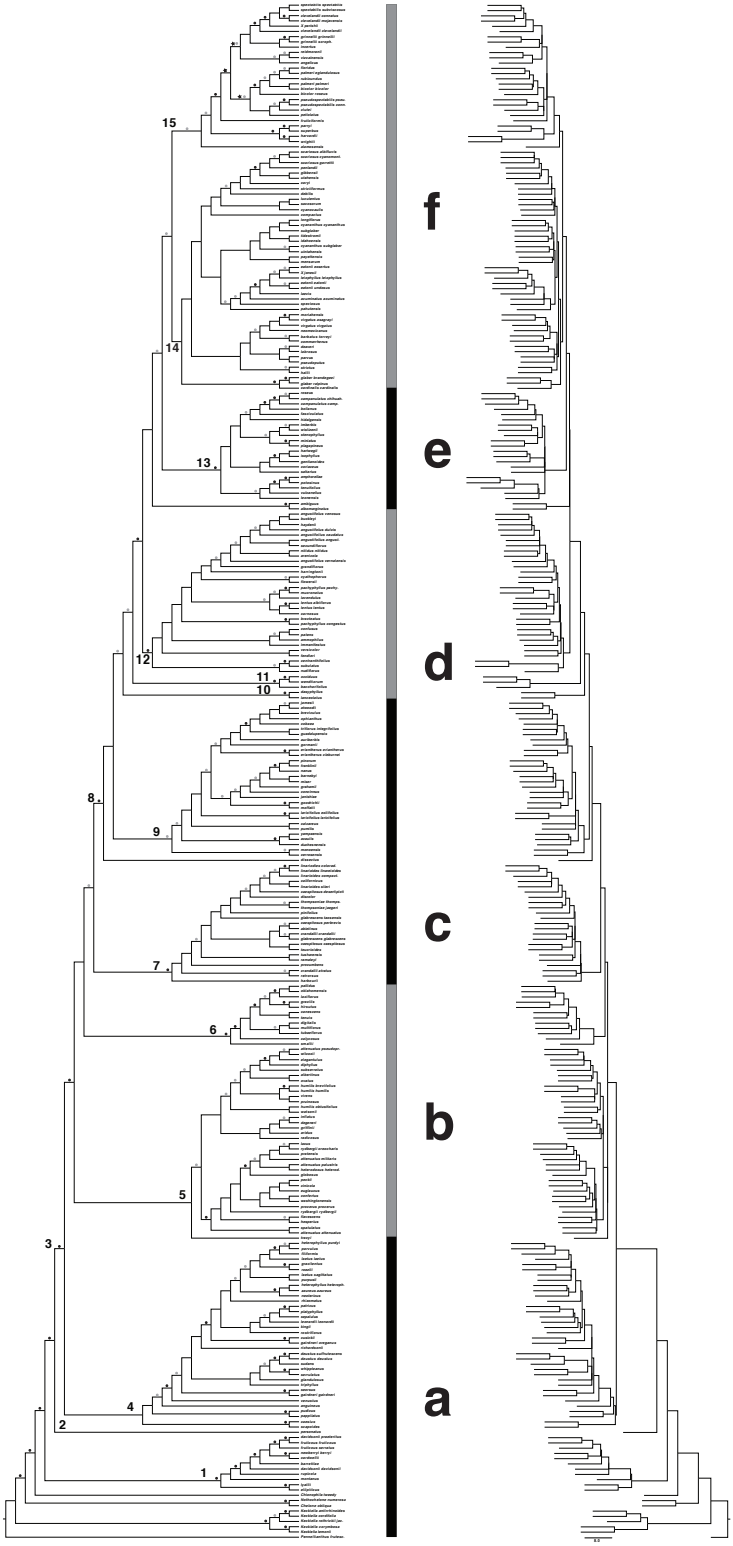
Of the 48 loci targeted for sequencing, 43 were successfully amplified in a sufficient number of our samples and were used for downstream analyses (Table S1; see filtering criteria in the Appendix 2 in Supplement Materials). Our final data matrix contained 17,518 sites (7171 variable, 4049 parsimony informative), 293 taxa (282 from *Penstemon*), and 30.4% missing data. The average number of non-missing sites per taxon was 12,191.2.

Phylogenetic Inference

Tribe Cheloneae is circumscribed by the genera *Pennellianthus, Collinsia, Keckiella, Chelone, Nothochelone, Chionophila,* and *Penstemon*. Our phylogeny based on 43 nuclear genes recovered a topology for the tribe consistent with earlier studies (Fig. 1; Wolfe et al. 2002, 2006), but with higher relative nodal support throughout the tree, and greater resolution within clades and along the backbone. Relationships among early diverging lineages in *Cheloneae* in this study also agree with previous work and support *Penstemon* as monophyletic with *Chionophila tweedyi* as sister to the entire genus (prior >0.8) (Fig. 1a).

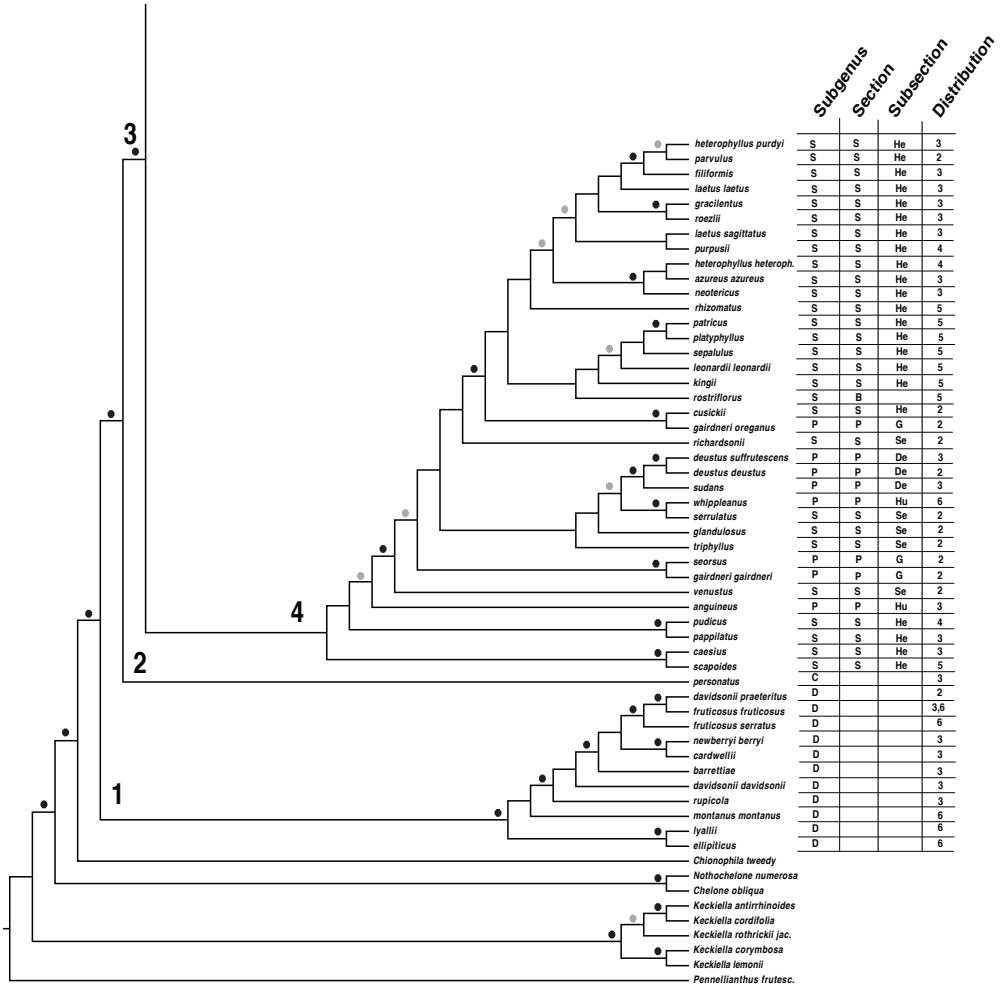
Within *Penstemon*, we infer subgenus *Dasanthera* as monophyletic, as has also been found previously (Wolfe et al. 2002, 2006). Major differences between the phylogeny presented here and earlier studies include the well-resolved placement of subgenus *Dasanthera* as sister to the rest of the genus (Fig.1a position 1), the position of *P. personatus* (subg. *Cryptostemon*; Fig. 1a position 2) as the next diverging lineage, a strongly supported clade containing the rest of the genus (referred to below as the crown clade; Fig 1a position 3), strong support for the majority of the group of penstemons classified as subgenus *Saccanthera* (Fig. 1a position 4; Freeman 2019), affinities for sect. *Penstemon* subsections *Proceri* and *Humiles* (Fig. 1b position 5), strong support for sect. *Penstemon* subsect. *Penstemon* to include subsections *Multiflori* and *Tubaeflori* (Fig. 1b position 6), and strong support for the majority of taxa traditionally in sect. *Ericopsis* (Fig. 1c position 7). *Penstemon dissectus* is the first diverging lineage for remaining group of penstemons (Fig. 1c position 8). Additional differences are the moderate to strong support for much of the backbone for the clade of penstemons sister to *P. dissectus* (Figs. 1c–1f). A clade consisting of section *Coerulei* and species from sections *Gentianoides* and *Glabri* has strong nodal support (Fig. 1d position 12). There is strong support for section *Fasciculus* (Fig. 1e position 13), and resolution of groups within the terminal clade (Fig. 1f) with moderate to strong bootstrap support. Membership of this clade is consistent with the inclusion of members of sect. *Gentianoides* with previous studies (Wessinger et al. 2016). However, our study shows members of *Habroanthus* as part of this terminal clade, also. The remaining taxa in *Habroanthus* group in a clade that includes members of subgenus *Penstemon* sections *Coerulei* and *Gentianoides* (Fig. 1f position 14), which is sister to a clade containing members of subgenus *Penstemon* sections *Peltanthera* and *Gentianoides*. Relative nodal support was moderate to strong within the *Peltanthera/Gentianoides* clade (Fig. 1f position 15) but was mostly lacking or moderate in the

287 traditional clade representing subgenus *Habroanthus*. There was no taxonomic distinction to
288 separate the red-flowered subg. *Habroanthus* sect. *Elmigera* from sect. *Glabri* (Fig. 1f position
289 14).
290
291 Figure 1. Phylogeny of *Penstemon* and related members of Cheloneae based on the Astral
292 analysis. Relative branch lengths are shown on the right. The vertical bar with lettering refers to
293 insets 1a–1f, which show details of the phylogenetic tree.



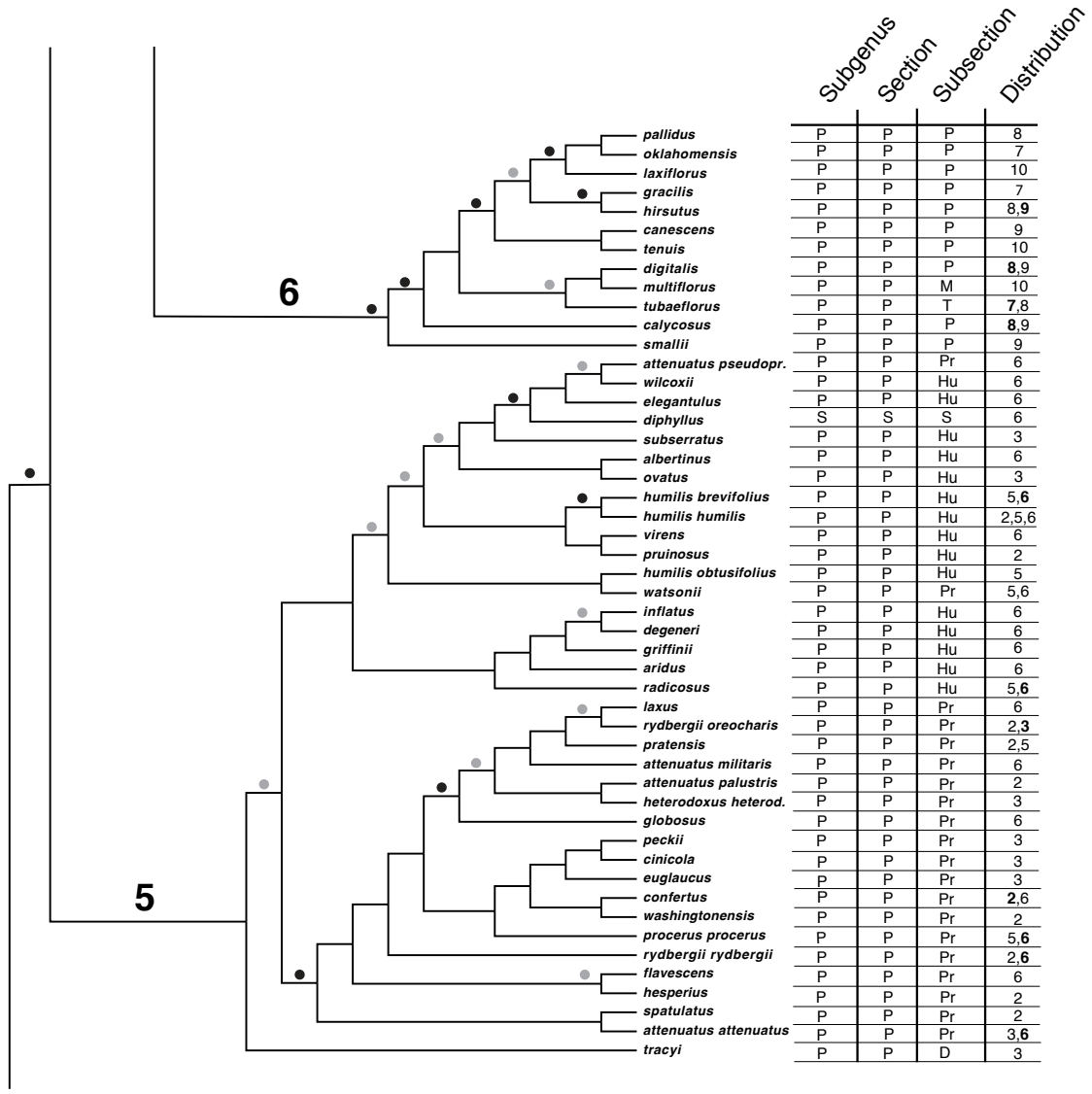
294

a)



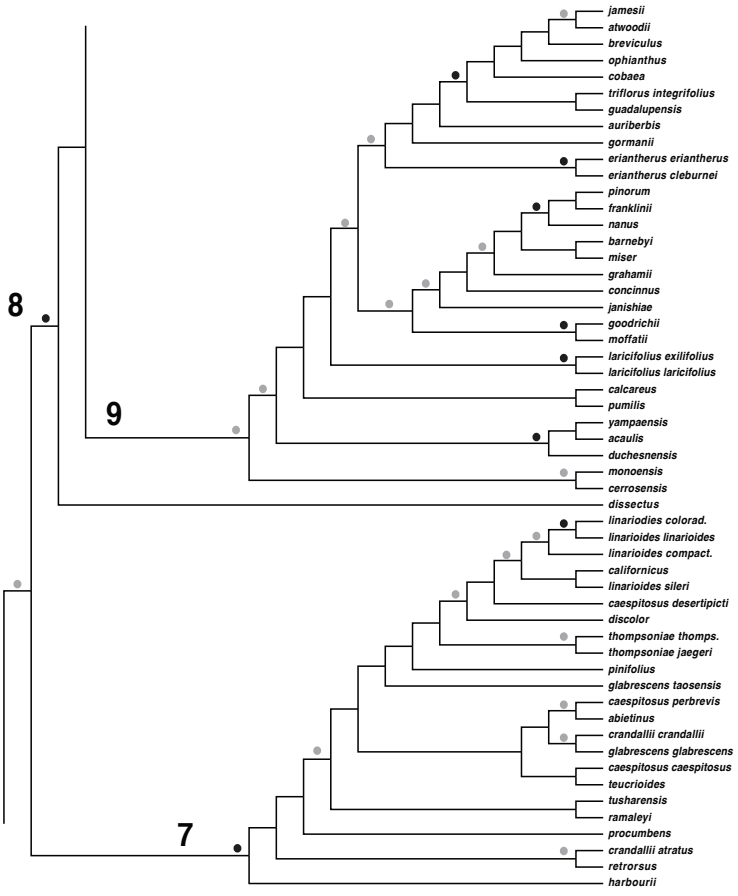
295

b)



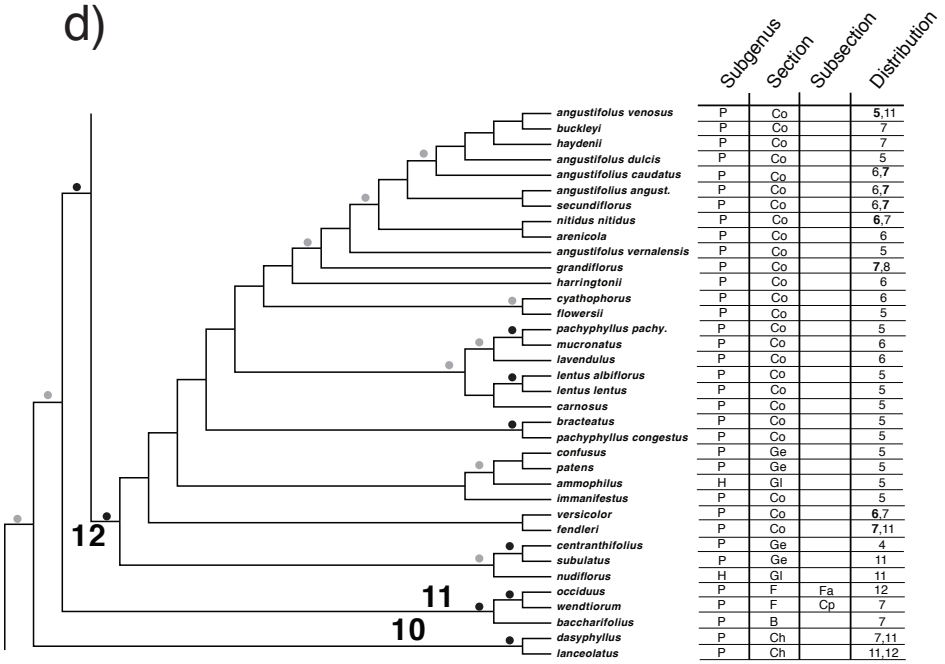
296

c)



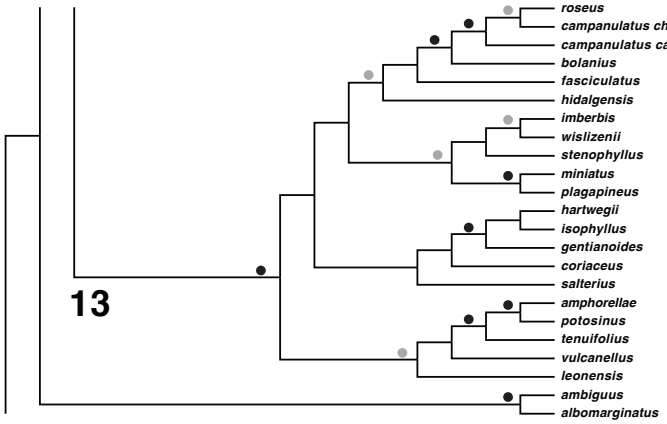
Subgenus	Section	Subsection	Distribution
P	Cr		7,**11**
P	Cr		5
P	Cr		5,**11**
P	Cr		5,11
P	Cr		7
P	Cr		7
P	Cr		7
P	Cr		7
P	Cr		1
P	Cr		6,7
P	Cr		6
P	Cr		5
P	Cr		5
P	Cr		5
P	Cr		5
P	Cr		2
P	Cr		5
P	Cr		5
P	Cr		5
P	Cr		5
P	Cr		5
P	E	E	6
P	E	E	6
P	Cr		11
P	Cr		6
P	E	Ca	6
P	E	Ca	6
P	Cr		5
P	Cr		11
P	Ge		4
Di			10
P	E	L	6,**11**
P	E	L	6,**11**
P	E	L	11
P	E	L	4
P	E	L	5,**11**
P	E	L	5,**11**
P	E	L	11
P	E	Ca	5
P	E	Ca	5
P	F	F	11
P	E	Ca	6
P	E	Ca	5
P	E	Ca	5
P	E	Ca	6
P	E	Ca	5
P	E	Ca	5,6
P	E	Ca	5
P	E	Ca	6
P	E	Ca	6
P	E	Ca	6
P	E	Ca	5
P	E	Ca	6
P	P	Ha	6

297



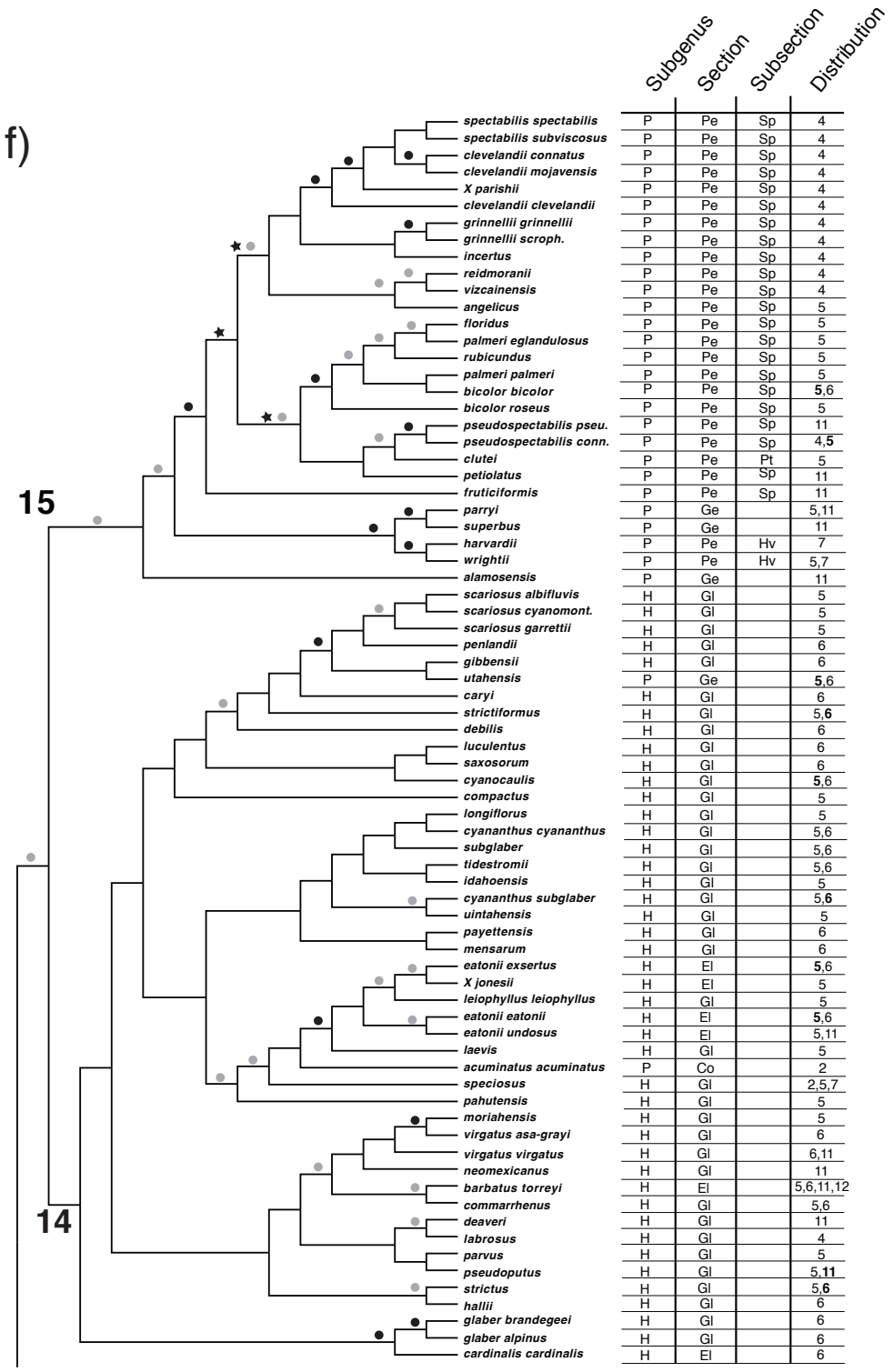
298

e)



Subgenus	Section	Subsection	Distribution
P	F	Cp	12
P	F	Cp	11,12
P	F	Cp	12
P	F	Cp	12
P	F	Fa	12
P	F	Pf	12
H	El		12
H	El		12
P	Ch		11,12
P	F	Fa	12
P	F	Fa	12
P	F	Fa	12
P	F	Fa	12
P	F	Fa	12
P	F	Cp	12
P	F	R	12
P	F	R	12
P	F	Cp	12
P	F	R	12
P	F	R	12
P	F	Fa	12
P	A		5,6,7,11
P	P	Ar	5

299



300

Taxonomic Bootstrap

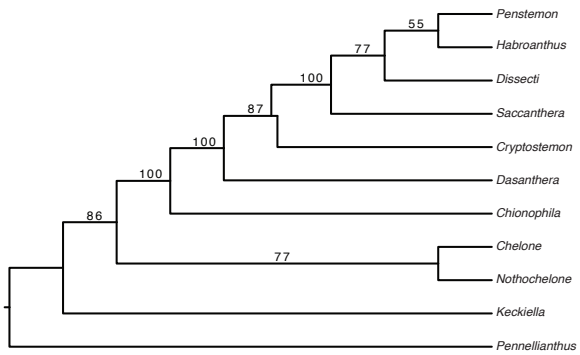
Three analyses were conducted to test the taxonomic circumscription of the genus: 1) subgenus, 2) section, and 3) subsection (Fig. 2). In the subgenus taxonomic bootstrap (Fig. 2a), support for the backbone is robust, and shows strong support for the major groupings recognized in *Penstemon,* except for the relationship among taxa in subgenera *Penstemon* and *Habroanthus*. The pattern of relationships among subgenera are as follows: subg. *Dasanthera* as the earliest diverging lineage and sister to a clade containing all the other subgenera. Subgenus *Cryptostemon* is sister to the core group of penstemons*,* and there is strong support for the grouping of taxa within subg. *Saccanthera*. It is important to note that this is a test of the current classification of major groups, and not the phylogenetic tree for the genus. However, the general order of branching for major groups matches those seen in the 43-gene phylogeny presented in Figure 1, and this is the same pattern seen for the section and subsection bootstrap analyses. For the section bootstrap analysis (Fig. 2b), the earliest diverging lineages show the same robust support as the 43-gene phylogeny, and sections within subg. *Saccanthera* are strongly supported as sister taxa. Support for taxa grouping in all recognized sections of the rest of *Penstemon* is nonexistent, but the support for subg. *Dissecti* as sister to what Wessinger et al (2016, 2019) refer to as the "crown group" of *Penstemon* has strong nodal support.

The subsection bootstrap analysis (Fig. 2c) reveals that current classification for the genus is not well supported by phylogenetic relationships. The topology of the early diverging lineages is concordant with the overall phylogeny (Fig. 2c nodes 1, 2, and 3). In this analysis, subgenus *Saccanthera* does not hold together as a monophyletic group but is grouped with members of subg. *Penstemon* sect. *Penstemon* (Fig. 2c node 5), which is sister to three of the

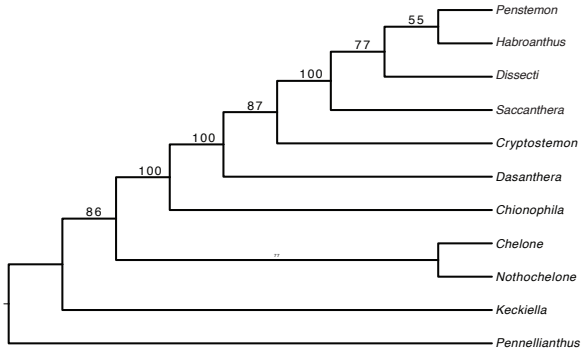
four subsections of sect. *Ericopsis* (Fig. 2c node 6). The topology of the latter group has strong bootstrap support. The monotypic sect. *Penstemon* subsect. *Harbouriani* is grouped with two subsections of sect. *Ericopsis* (Fig. 2c. node 6), which is consistent with the Wolfe et al. (2006) study and our 43-nuclear-gene tree (Fig. 1). Subgenus *Penstemon* sect. *Penstemon* is polyphyletic in the subsection bootstrap tree (Fig. 2c nodes 4 and 8). The relationship among subsections for the group representing members of subg. *Saccanthera*, subsects. *Deusti*, *Gairdneri*, *Humiles, Proceri, Caespitosi, Linarioides*, and *Harbouriani* to the rest of *Penstemon* has strong nodal support (Fig. 2c node 7). Three of the sect. *Penstemon* subsections are united by strong bootstrap support (Fig. 2c node 8). The position of subg. *Dissecti* is unresolved in the subsection bootstrap analysis (Fig. 2c node 9), but sect. *Ericopsis* subsect. *Ericopsis* is strongly supported as grouped with sect. *Cristati* (Fig. 2c. node 10). Taxonomic affinities for sect. *Coerulei* are uncertain (Fig. 2c node 11), but there is strong support for an affinity between sect. *Penstemon* subsect. *Arenarii* and sect. *Ambiguii* (Fig. 2c node 12). The subsection topology for the crown clade of *Penstemon* (Fig. 2c node 13) is not well resolved. Subsections of sect. *Fasciculus* are grouped with a bootstrap value of 50 (Fig. 2c node 14). However, two of the subsections group with strong support (Fig. 2c node 15). Subgenus *Habroanthus* sections *Glabri* and *Elmigera* have no bootstrap support (Fig. 2c, node 16), and relationships among what the American Penstemon Society refers to as sect. *Peltanthera*, which includes subsect. *Centranthifolii* (sect. *Gentianioides*), subsect. *Havardiani* (sect. *Gentianioides*), subsect. *Peltanthera* (sect. *Spectabiles*), and subsect. *Petiolati* (sect. *Spectabiles*) are grouped at node 17 of Figure 2c. The only group in this clade with strong support consists of subsect. *Peltanthera* and *Petiolati* (Fig. 2c node 18).

347 Figure 2. Taxonomic bootstrap results f or a) subgenera, b) sections, and c) subsections.

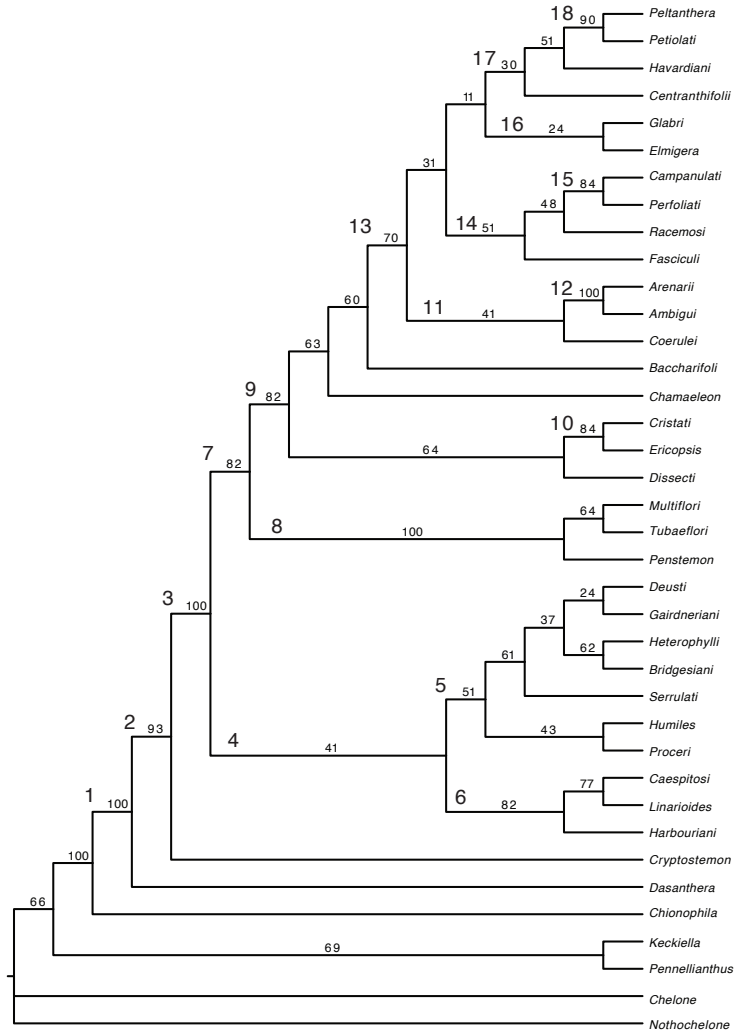
a)



b)



c)

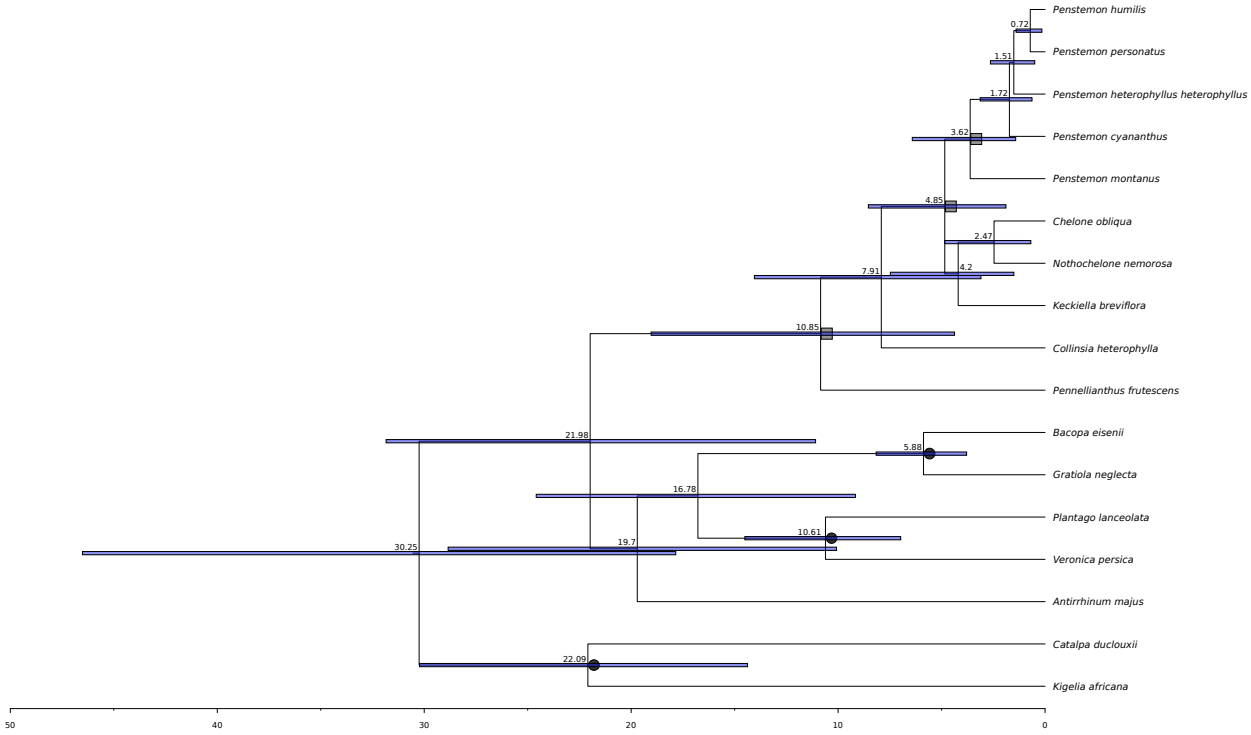


348

Time Calibration and Analysis of Diversification Rates

Divergence dating with BEAST using three fossils with minimum age estimates for members of the families Bignoniaceae and Plantaginaceae produced a well-supported phylogeny for the Cheloneae, with all sampled parameters having ESS values greater than 200 and no signs of a lack of convergence in the trace plots (Fig. 3). Based on this time calibration, we infer the age of the tribe to be 10.85 my (95% HPD interval: 4.375–19.04 my), with an estimated age for the origin of *Penstemon* 3.62 mya (95% HPD interval: 1.419–6.416). Using the 95% HPD intervals for these two age estimates as secondary calibration points, plus a third calibration point for members of the Cheloneae without *Pennellianthus*, we were able to time calibrate our entire 293-taxon phylogeny with treePL. Dates inferred by treePL were slightly older for the origin of the tribe (14.94 mya) but remained roughly the same for the origin of Penstemon (3.66 mya). The divergence for the crown clade of Penstemon (excluding subg. *Dasanthera* and *P. personatus*) was placed at 2.51 mya, demonstrating that the majority of the diversification within *Penstemon* has occurred since the early Pleistocene.

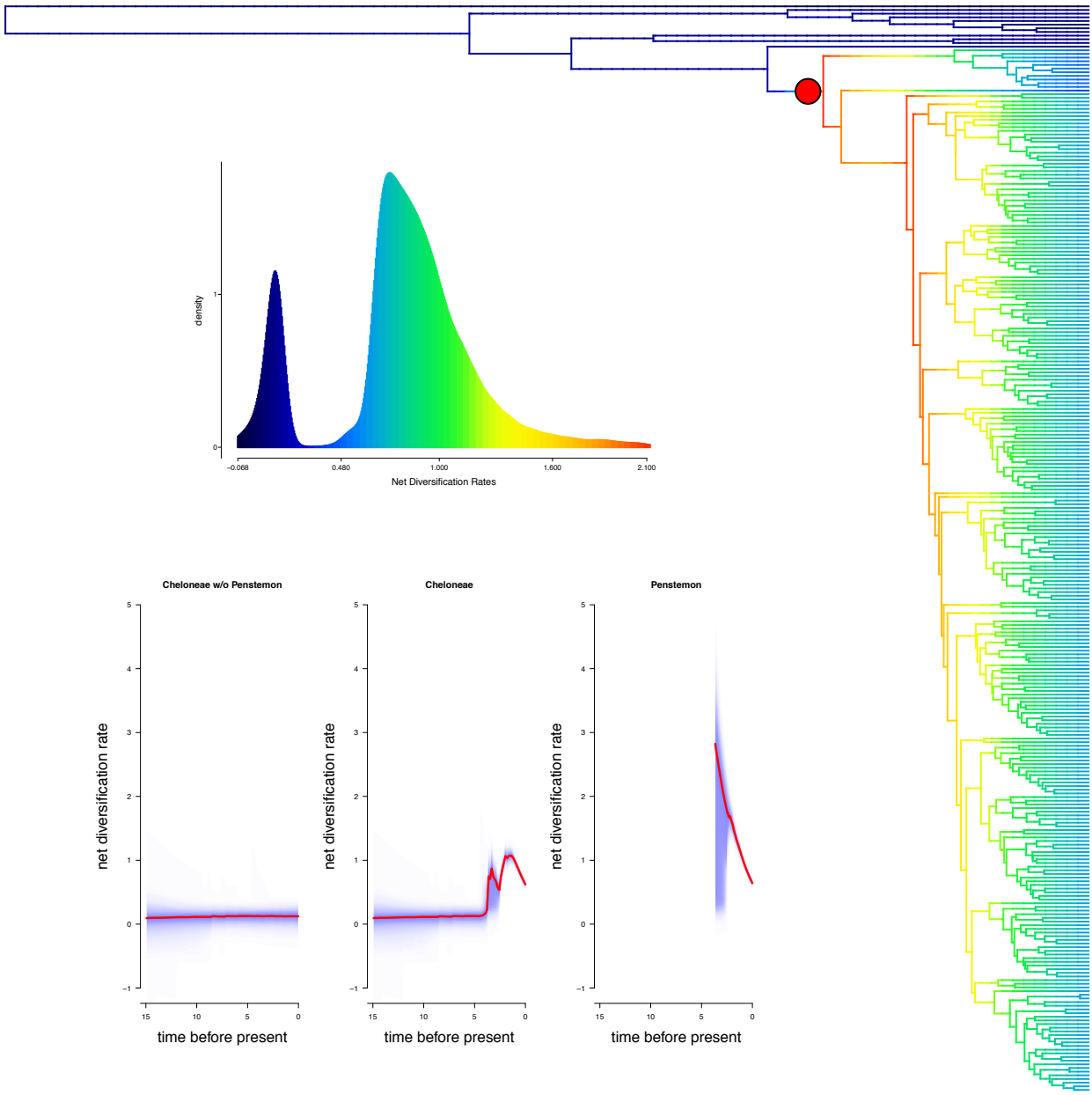
Figure 3. Time-calibrated tree for Chelonae and *Penstemon* based on fossils for members of Lamiales. Dark circles mark nodes that were calibrated using fossils from Vargas et al. (2014). Grey squares mark nodes that were used as secondary calibration points for dating the entire 293-taxon phylogeny with treePL.



Diversification analyses with BAMM and MEDUSA inferred similar patterns of macroevolutionary dynamics, showing a large increase in speciation rate at the base of *Penstemon* (Fig. 4). There was also some support in the posterior sample in BAMM for the placement of this shift in diversification at the base of the crown clade of *Penstemon* (Fig. S2). However, regardless of placement, both analyses support a single shift in diversification rate over more complex models with multiple shifts. While there is still a fair amount of controversy regarding the analysis of speciation and extinction rates from time-calibrated, molecular phylogenies (Moore et al. 2016; Louca and Pennell 2020), this result is consistent with the hypothesis from Wolfe et al. (2006) that *Penstemon* is a recent and rapid evolutionary radiation.

382 Figure 4. Inferred patterns of diversification using BAMM. The phylogeny shows the
383 reconstructed rates of net diversification inferred across the tree with a single shift in
384 diversification (best shift configuration) at the base of *Penstemon*. The top inset plot shows the
385 distribution of net diversification rates for the entire phylogeny and the bottom inset plot shows
386 the inferred net diversification rate through time for Chelonae without *Penstemon* (left), the
387 Chelonae with *Penstemon* (center), and *Penstemon* alone (right).
388

389



390

391

392 *Biogeographic Distribution*

393 Analyses with BioGeoBEARS consistently chose the BAYAREALIKE+J model as the best fit
394 for all three sets of areas (Table 3; Fig. S3). The DEC+J and DIVALIKE+J models were also
395 either the second or third best models, demonstrating that founder-event speciation events were
396
397 Table 3. Model comparison results for BioGeoBEARS analyses.

	Model	Log-Likelihood	AICc Weight	d	e	j
Set 1	BAYAREASLIKE+J	-372.4	0.59	0. 019	1e-7	0.039
	DEC+J	-372.9	0.38	0.023	1e-12	0.036
	DIVALIKE+J	-375.6	0.025	0.026	1e-12	0.036
	DEC	-442.5	6.2e-31	0.053	0.038	0
	BAYAREALIKE	-446.9	7.2e-33	0.023	0.52	0
	DIVALIKE	-449.7	7.0e-34	0.068	0.024	0
Set 2	BAYAREALIKE+J	-424.0	0.92	0.032	1.0e-07	0.033
	DEC+J	-426.5	0.079	0.039	1.0e-12	0.029
	DIVALIKE+J	-431.8	0.0004	0.042	1.0e-12	0.029
	DEC	-479.6	1.9e-24	0.061	0.012	0
	BAYAREALIKE	-489.1	1.4e-28	0.022	0.54	0
	DIVALIKE	-492.1	6.9e-30	0.078	1.0e-12	0
Set 3	BAYAREALIKE+J	-160.3	1.0	0.0072	0.0028	0.030
	BAYAREALIKE	-250.9	1.4e-39	0.013	0.26	0
	DEC+J	-378.3	2.1e-95	0.10	1.0e-12	0.025
	DIVALIKE+J	-385.8	1.3e-98	0.11	1.0e-12	0.025
	DEC	-388.5	2.4e-99	0.11	1.0e-12	0
	DIVALIKE	-402.7	1.7e-105	0.13	0.016	0

398
399 likely to be an important process during the diversification of *Penstemon*. For '+J' models,
400 estimates of the rate of dispersal, extinction, and founder-events were all fairly consistent across
401 sets of analyses, with extinction rates estimated to be close to 0. For analyses without the
402 founder-event parameter, estimates of these rates varied more greatly. In particular, the
403 BAYAREALIKE model tended to infer a higher extinction rate by almost an order of magnitude
404 when founder-event speciation was not modeled. Despite this variation, ancestral area

405 reconstruction for these analyses placed the most likely origin of *Penstemon* in the Eastern
406 Cordillera (region 6), which agrees with previous hypotheses (Wolfe et al. 2002, 2006).
407 One hundred sixty-five distribution map images were assembled across the phylogeny of
408 *Penstemon* from this study. Individual images are in Figure S4, and the collective maps for the
409 100,000-year time slices inferred from a dated tree are in Figure S5. The time slices were
410 inferred from the dated tree for the genus (Fig. S6). Animations of these maps in the order of
411 appearance for either the clade position or the time slice can be seen at
412 https://www.youtube.com/watch?v=lMLn7Gq_ZPw. Biogeographic regions (Fig. S1, Table 2)
413 for taxa were determined by the distribution of populations shown in the individual KML files
414 collected from the SEINet data portal (www.swbiodiversity.org). These regions were mostly
415 based on the physiographic regions of North America (Barton et al. 2003) and were modified
416 from Wolfe et al. (2006) to better reflect digitized collection data.
417
418 DISCUSSION
419
420 *Phylogenetic Inference and Time Calibration*
421 This study is the most comprehensive phylogenetic analysis of *Penstemon* to date, with 239 of
422 ca. 285 species included. Previous studies have established that this genus represents a large
423 continental radiation of recent origin. Our time calibration and divergence time analyses confirm
424 this pattern, with *Penstemon* originating around the Pliocene/Pleistocene boundary. Given the
425 large number of species diverging in a relatively short amount of time, we estimate that
426 *Penstemon* has one of the highest rates of diversification of any plant genus in a continental
427 setting (Breitkopf et al. 2015; Schwery et al. 2015; Tank et al. 2015; Verboom et al. 2015;

Kriebel et al. 2019). We hypothesize that this may be due to a pattern of adaptive radiation, which will be examined thoroughly in another study.

The difference in resolution from phylogenetic analyses based on ITS and cpDNA (Wolfe et al. 2006), and the current 43-nuclear-gene loci is notable, but not surprising. One would expect better tree topology resolution with more data. However, with a relatively young genus such as *Penstemon*, it is also not surprising that relative nodal support is lacking in some areas of the tree, despite the use of a large dataset. The Wessinger et al. (2019) study with the largest number of taxa (120 total species, 104 species in the crown clade) included 2306 and 2051 loci, respectively, with 72% and 68% missing data. A consensus locus was aligned "if there were at least 20 taxa per locus, no more than 50 SNPs and 8 indels per locus, and no more than 8 shared heterozygous sites across samples" (Wessinger et al. 2019). The trees in Wessinger et al. (2016, 2019) had relatively strong nodal support along the backbone of the tree, strong bootstrap support in some terminal clades, but not complete resolution of relationships throughout the tree. Comparing nuclear-gene amplicon sequences to MSG SNP data is difficult, but there are areas of agreement between the phylogenies produced by these studies as well as areas where the topologies differ. The general patterns from the phylogenies are similar enough that where backbone support is not as strong in our 43-gene tree, but in agreement with the MSG SNP results, we infer that the clade topologies represent relationships as shown in our results (Figs. 1c–1f).

Patterns of genealogical discordance and rapid rates of evolution are also apparent in our data, as evidenced by the short branches in the ASTRAL-III tree (Figs. 1, S7). Given the recent radiation of the group, such patterns are expected and are likely a result of incomplete lineage sorting caused by rapid speciation. Other processes such as hybridization and allopolyploidy

451 have also been documented in *Penstemon* (Keck 1945; Wolfe et al. 1998a,b; Broderick et al.
452 2011) and further complicate the accurate inference of phylogeny where they occur. However,
453 the coalescent branch lengths in the ASTRAL-III tree (Figs. 1, S7) indicate that high levels of
454 discordance are more prevalent towards the tips of the tree, with little evidence of discordance
455 among the major lineages of the Cheloneae. In general, it is important to consider the effect of
456 these sources of discordance on downstream inferences. For the analyses we conducted here
457 (divergence dating, diversification, and biogeography), we are primarily focused on nodes deeper
458 in the phylogeny that are less affected by the incongruence seen in other parts of our tree.
459 Because of this, we argue that our results regarding the timing and diversification of the genus
460 are robust, especially considering the agreement between our work and previous studies, as well
461 as our dense sampling and the comparatively low amount of missing data in our analyses.
462 Our inferred timing for the biogeographic and diversification history (Table S2) of
463 *Penstemon* in the Pliocene/Pleistocene supports previous hypotheses regarding its spread across
464 North America during these epochs' dynamic periods of glaciation (Wolfe et al. 2006, Ehlers and
465 Gibbard 2007; Ehlers et al. 2018). And while the age we infer for *Penstemon* places its origin
466 slightly farther back in time than previously thought (3.66 Mya versus ~2.5 Mya; Wolfe et al.
467 2006), the timing of divergence for the majority of species in the crown clade still coincides with
468 the onset of the Pleistocene. Absent more closely related fossils than those from Vargas et al.
469 (2014) at the family level, our divergence dating with secondary calibration points represents the
470 only fossil-based divergence estimation conducted for the tribe. This time-calibrated phylogeny
471 was crucial for understanding the pattern and timing of diversification, showing that *Penstemon*
472 not only diversified at a much higher rate than other members of Cheloneae (Fig. 4) but that the
473 shift in diversification rate occurred at the base of the clade, coinciding with its inferred origin in

474 the Eastern Cordillera and subsequent dispersal through founder events across the continent.
475 Compared with other groups of angiosperms, the net diversification rate for *Penstemon* is
476 consistent with estimates for the Lamiales (Magallón and Sanderson 2002; Magallón and Castillo
477 2009); however, *Penstemon* is noteworthy as a particularly exceptional radiation when
478 considering that its diversity has arisen in only the last 3.66 My, with the bulk of diversification
479 occurring within the past 1.8 My. And while criticisms of diversification analyses are continuing
480 to raise important considerations for the interpretability of absolute estimates of diversification
481 dynamics (Moore et al. 2016; Louca and Pennell 2020), especially with regard to extinction
482 rates, the relative pattern of rapid diversification in *Penstemon* compared to its sister lineages
483 remains clear.
484
485 *Taxonomic Implications*
486 Taxonomic bootstrapping of the subgenera, sections, and subsections of *Penstemon* reveals some
487 interesting patterns, which speak to the need for taxonomic revision of the genus. In the
488 subgenus taxonomic bootstrap (Fig. 2a), the traditional categories for subg. *Dasanthera*,
489 *Cryptostemon*, *Saccanthera*, and *Dissecti* are not in conflict. The low bootstrap value for the
490 node representing subgenera *Penstemon* and *Habroanthus* indicates some taxonomic conflict
491 with placement of taxa in these categories. The taxonomic bootstrap of sections (Fig. 2b) shows
492 the same early branching pattern support for *Dasanthera*, *Cryptostemon*, and the grouping of
493 sections within subg. *Saccanthera*. Indications of taxonomic conflict occur throughout the
494 remainder of the tree as evidenced by low to moderate bootstrap values for all nodes beyond
495 subg. *Saccanthera*.

The subsection bootstrap analysis (Fig. 2c) yields insights into current taxonomic structure of the genus. Nodes 1–3 show the same strong support for the early diverging lineages (subg. *Dasanthera* and subg. *Cryptostemon* relationship and their placement as sister to the rest of the genus). Node 4 reveals conflict in taxonomic placement of some taxa within subgenera *Saccanthera* and *Penstemon*, particularly with the inclusions of subg. *Penstemon* sect. *Penstemon* subsections *Deusti* and *Gairdneri* with sections and subsections of subg. *Saccanthera*. Low bootstrap values for the node representing sect. *Penstemon* subsections *Humiles* and *Proceri* also indicate taxonomic problems within those groups. Node 6 indicates that two of the subsections of section *Ericopsis* are well defined, but these also group with the monotypic sect. *Penstemon* subsect. *Harbouriani*. Three of the subsections of section *Penstemon* have high nodal support (Fig. 2c, Node 8). Other indicators of taxonomic conflict are represented by nodes 10 (placement of sect. *Ericopsis* subsect. *Ericopsis* as sister to sect. *Cristati*), 12 (sect. *Penstemon* subsect. *Arenarii* as sister to sect. *Ambiguii*), 14 (low nodal support for taxa within section *Fasciculus*), 16 (no support for subg. *Habroanthus* sections *Glabri* and *Elmigera*), and 17 (placement of taxa within section *Gentianioides* with no nodal support).

In the context of the phylogeny presented here (Fig. 1), clearly there is a need to revise the circumscription of the genus *Penstemon.* Similar patterns of taxa grouping outside their assigned sections and subsections have been seen in other recent studies with subsets of species (Wessinger et al. 2016, 2019). Based on the current study, together with the Wessinger et al studies (2016, 2019), the number of subgenera should be decreased from six to four: *Dasanthera*, *Cryptostemon*, *Saccanthera*, and *Penstemon*. The traditional subgenus *Habroanthus* should be designated as a section, sans designated subsections *Glabri* and *Elmigera*. The traditional monotypic subgenus *Dissecti*, should be section *Dissecti*, given its phylogenetic placement (Fig.

519 1c). Many taxa need to be shuffled from their traditional assigned subgenera, subsections, or
520 sections to other groupings. Recommended changes based on phylogenetic studies can be found
521 in Table S3.
522
523 *Biogeography*
524 Biogeographic analysis of *Penstemon* with BioGeoBEARS (Fig. S3) confirmed previous
525 hypotheses for the origin of the genus in the eastern Cordillera of North America and its
526 subsequent spread across the continent (Straw 1966; Wolfe et al. 2002). The importance of
527 founder-event speciation for this process is demonstrated here for the first time, indicating that
528 dispersal to new, ancestrally unoccupied areas was a key mechanism for the evolution and
529 expansion of *Penstemon*. Given the dynamic nature of glaciation in North America during the
530 diversification of the genus (see more details below), this pattern makes sense and is likely
531 connected with the recurrent formation of "sky islands" in western and southwestern North
532 America (Rehfeldt 1999; Knowles 2001; Hewitt 2004). As with our analyses of diversification
533 with BAMM and MEDUSA, we find that the inferred rate of extinction is effectively zero for the
534 most well-supported models. However, we do note that the inference of extinction rates from
535 molecular phylogenies can be problematic (Rabosky 2010; Louca and Pennell 2020), so we
536 interpret this absence of extinction with caution, especially considering the paucity of fossil
537 evidence to corroborate diversification patterns for the Cheloneae. Nevertheless, the consistent
538 pattern of elevated net diversification in *Penstemon* and its probable connection with the
539 importance of founder-event speciation as inferred by BioGeoBEARS helps point to a possible
540 mechanism for its rapid radiation.

The mapped biogeographic patterns for *Penstemon* are associated with the glaciation cycles of the Pleistocene (Table S2, Figs. S4, S5; https://www.youtube.com/watch?v=lMLn7Gq_ZPw). There were 13 major glaciations and nine minor glaciations during the 2.7 years of the Pleistocene (Ehlers et al. 2018). Although most biogeographic histories emphasize the effect of the last glacial maximum, we argue that each cycle of glaciation had an impact on the diversification of *Penstemon*. The first four major glaciations occurred between 2.7 mya and 2.0 mya. During that time frame the earliest diverging lineage, subgenus *Dasanthera*, spread through the northeastern cordillera into the Cascade-Sierra cordillera and Pacific Northwest. The next major glaciation occurred between 2.0–1.8 mya. During this timeframe, *P. dissectus*, endemic to Georgia, appears in the eastern coastal plain. Given the phylogenetic position of *P. dissectus* in the middle of the phylogenetic tree (Figs. 1c, S2), one can confidently infer that distribution of species was more widespread throughout the southern reaches of the North American continent, but that extinction events occurred during the first million years of the Pleistocene. This is supported in our BAMM analyses (Fig. 4) where diversification shows an early burst along the backbone of the phylogenetic tree, followed by a decrease and/or extinction as the genus evolved. Figure 4 also illustrates two periods of diversification for tribe Cheloneae including *Penstemon*. There was an initial burst followed by a dramatic decline in net diversification rate, and a second burst and decline corresponding with the phylogenetic time frame represented in Table S2.

Between 1.8 mya and 1.2 mya there was one major and four minor glaciation episodes. During this timeframe, most of the diversification was taking place in the eastern cordillera and intermountain region (Table S2), except for the appearance of *P. smallii* in the Appalachian region. Extension of *Penstemon* into the southwest region, southern Great Plains, and highlands

564 of Mexico had just begun. From 1.2 to 1.1 mya two major glaciation cycles took place. The
565 diversification into regions shows a north-south oscillation from the Cascade-Sierra and eastern
566 cordillera ranges as far north as Alaska, and into the southwest and Mexico highlands. The
567 southern Great Plains in the west also showed activity as did the southern Intermountain Region.
568 From 1.1 mya to 1.0 mya there was one minor glaciation. There appeared to be a pattern of
569 diversification from the eastern Pacific Northwest, southeast into the eastern cordillera and
570 western Great Plains, and east into the Intermountain region.

571 Diversification of *Penstemon* was most active between 1.0 mya to 0.5 mya (Table S2,
572 Figs. S4, S5). During this time frame there were two major glaciations (0.9–0.8 mya and 0.7–0.6
573 mya) and one minor glaciation (1.0–0.9 mya). During the latter time period, there was an
574 expansion from the southern Cascade-Sierra and intermountain regions north and east into the
575 eastern cordillera and northern Cascade-Sierra. In the east there was an expansion from the
576 Appalachian region into the southern interior lowlands and eastern Great Plains, and expansion
577 through the coastal plain. From the southeastern cordillera there was an expansion into the
578 Intermountain Region, then southwest and back into the Eastern Cordillera. During the first
579 major glaciation cycle of this time period, there was an expansion south from the Cascade-Sierra
580 to the Southwest Cordillera and Baja California. This was concurrent with a trajectory from the
581 Intermountain region into the Southwest and Southwest Cordillera. During the subsequent
582 interglacial period, the diversification followed a track northward into the Eastern Cordillera and
583 Intermountain Region, and south into the Mexico Highlands. There was also an expansion from
584 the Eastern Cordillera into the Great Plains, and in the east, there was an expansion through the
585 eastern interior lowlands and northward in the Appalachian region. During the second major
586 glaciation of this period, there was another cycle of diversification to the south (Intermountain

587 region to Southwest, and Mexico Highlands), and diversfication into the southern Great Plains.
588 The next glacier-free 100,000 years had activity in the Pacific Northwest and Cascade-Sierra,
589 and an eastward movement of diversification in the Great Plains. There was also diversification
590 in the Interior Lowlands.

591 The diversification of *Penstemon* slowed down during the most recent 500,000 years,
592 with four major glaciations and two minor glaciation cycles, including the Last Glacial
593 Maximum period. During this time period, most of the action was in the southern areas of the
594 North America continent, notably in the Intermountain Region, Southwest, Southwest Cordillera
595 and Baja California, and the Mexico Highlands.

596

597 ## CONCLUSIONS

598

599 *Penstemon* is a remarkable genus in its rapid diversification into most ecological regions found
600 in North America (Fig. S1). The diversification, phylogeny, and biogeographic history suggest a
601 rapid evolutionary radiation throughout North America during the Pleistocene. The pulses of
602 diversification associated with major and minor glaciation periods, in the context of the models
603 supported by our BioGeoBEARS analyses, are consistent with the colonization of newly
604 available ecological niches during interglacial cycles. Considering the immense diversity in
605 corolla, anther, staminode, leaf, inflorescence, and habit morphology, taken together with the
606 large amount of variation in habitats, it is likely that *Penstemon* has undergone an adaptive
607 radiation in a continental setting.

608

609 ## FUNDING

This work was supported by a grant from the National Science Foundation under award DEB-145539 to A.D.W. and L.S.K.

ACKNOWLEDGEMENTS

The authors thank members of the Wolfe and Kubatko labs for comments that helped to improve this manuscript. We also thank N. Holmgren, M. Stevens, G. Moffit, past and current members of the Wolfe Lab, and members of the American Penstemon Society for help with collecting specimens. In addition, we thank the Wolfe Lab graduate and undergraduate students, past and current, for their assistance with DNA extractions, database management, and sequencing preparation. Without their dedicated contributions this study would never have been possible.

DATA AVAILABILITY

Raw sequencing reads, aligned DNA sequences, data matrices, tree files, and code for all analyses is available on Dryad (doi:XXXX).

628 REFERENCES

629

630 Akaike H. 1974. A new look at the statistical model identification. IEEE Transaction on
631 Automatic Control 19:716–723.
632 Alfaro M.E., Santini F., Brock C., Alamillo H., Dornburg A., Rabosky D.L., Carnevale G.,
633 Harmon L.J. 2009. Nine exceptional radiations plus high turnover explain species diversity in
634 jawed vertebrates. Proc. Natl. Acad. Sci. U.S.A.106:13410–13414..
635 Barton K.E., Howell D.G, Vigil J.F. 2003. The North American Tapestry of Time and Terrain.
636 USGS Geologic Investigations Series map I-2781.
637 Blischak P.D., Latvis M., Morales-Briones D.F., Johnson J.C., Di Stilio V.S., Wolfe A.D., Tank
638 D.C. 2018. Fluidigm2PURC: automated processing and haplotype inference for double-
639 barcoded PCR amplicons. Appl. Plant Sci. 6:e1156.
640 Blischak P.D., Wenzel A.J., Wolfe A.D. 2014. Gene prediction and annotation in *Penstemon*
641 (Plantaginaceae): a workflow for marker development from low-coverage genome
642 sequencing. Appl. Plant Sci 2:1400044.
643 Bouckaert R., Vaughan T.G., Barido-Sottani J., Duchêne S., Fourment M., Gavryushkina A., et
644 al. 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis.
645 PLoS Comput. Biol. 15:e1006650.
646 Breitkopf H.R., Onstein E., Cafasso D., Schlüter P.M., Cozzolino S. 2015. Multiple shifts to
647 different pollinators fueled rapid diversification in sexually deceptive *Ophrys* orchids. New
648 Phytol. 207:377-389.
649 Broderick S.R., Stevens M.R., Geary B., Love S.L., Jellen E.N., Dockter R.B., Daley S.L.,
650 Lindgren D.T.. 2011. A survey of Penstemon's genome size. Genome 54:160–173.

651 Chifman J., Kubatko L.S. 2014. Quartet Inference from SNP Data Under the Coalescent Model.
652 Bioinformatics 30:3317–3324.
653 Crosswhite F.S. 1967. Revision of *Penstemon* section *Habroanthus* (Scrophulariaceae) I:
654 Conspectus. Am. Midl. Nat. 77:1-11.
655 Dockter, R. B. D. B. Elzinga, B. Geary, P J. Maughan, L. A. Johnson, d. Tumbleson, J.L. Franke,
656 K. Dockter, and M. R. Stevens. 2013. Developing molecular tools and insights into the
657 *Penstemon* genome using genomic reduction and next-generation sequencing. BMC Genet.
658 14: 66.
659 Drummond, A.J, S. Y. W. Ho, M. J. Phillips, and A. Rambaut. 2006. Relaxed Phylogenetics and
660 Dating with Confidence. PLoS Biol. 4: e88.
661 Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high
662 throughput. Nucleic Acids Res. 32: 1792–1797.
663 Ehlers, J. and P. L. Gibbard. 2007. The extent and chronology of Cenozoic Global Glaciation.
664 Quat. Int. 164-165: 6–20.
665 Ehlers, J., P. L. Gibbard, and P. D. Hughes. 2018. Quaternary glaciations and chronology. *In:*
666 Past Glacial Environments (Second Edition), eds. J. Menzies and J. J. M. van der Meer.
667 Elsevier, Ltd., Amsterdam, Netherlands. pp. 77-101.
668 Estes, D. 2012. *Penstemon kralii* (Plantaginaceae), a new species from Alabama and Tennessee,
669 with an updated key to the southeastern U.S. Taxa. J. Bot. Res. Inst. Texas 6: 1-18.
670 Freeman, C. E. 2019. *Penstemon*. In: Flora of North America Editorial Committee, eds. 1993+.
671 Flora of North America North of Mexico. 21+ vols. New York and Oxford. Vol. 17, pp 82–
672 255.
673 Hewitt, G. M. 2004. Genetic consequences of climatic oscillations in the Quaternary. Philos.

674 Trans. R. Soc. London, Ser. B 359: 183–195.

675 Holmgren, N. H. 1984. *Penstemon*. *In* A. Cronquist, A. H. Holmgren, N. H. Holmgren, J. L.
676 Reveal, and P. K. Holmgren [eds.], Intermountain flora: vascular plants of the intermountain
677 west, vol. 4, 370-457. New York Botanical Garden, Bronx, New York, USA.

678 Holmgren, N. H., and P. K. Holmgren. 2016. Intermountain flora. Vascular plants of the
679 intermountain West, U.S.A. Volume Seven. Potpourri: keys, history, authors, artists,
680 collectors, beardtongues, glossary, indices. New York Botanical Garden, New York, New
681 York. 312 pages.

682 Katoh, S. 2013. MAFFT multiple sequence alignment software version 7: Improvements in
683 performance and usability. Mol. Biol. Evol. 30: 772–780.

684 Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, et al.
685 2012. Geneious Basic: an integrated and extendable desktop software platform for the
686 organization and analysis of sequence data. Bioinformatics 28: 1647–1649.

687 Keck, D. D. 1945. Studies in *Penstemon*–VIII. A cytotaxonomic account of the section
688 Spermunculus. Am. Midl. Nat. 33: 128–206.

689 Knowles, L. L. 2001. Did the Pleistocene glaciation promote divergence? Tests of explicit
690 refugial models in montane grasshoppers. Mol. Ecol. 10: 691-701.

691 Kriebel, R., B. T. Drew, C. P. Drommond, J. G. González-Gallegos, F. Celep, M. M. Mahdjoub,
692 J. P. Rose, C-L. Xiang, G-X. Hu, J. B. Walker, E. M. Lemmon, A. R. Lemmons, and K. J.
693 Sytsma. 2019. Tracking temporal shifts in area, biomes, and pollinators in the radiation of
694 *Salvia* (sages) across continents: leveraging anchored hybrid enrichment and targeted
695 sequence data. Am. J. Bot. 106: 573-597.

696 Landis, M. J., N. J. Matzke, B. R. Moore, and J. P. Huelsenbeck. 2013. Bayesian Analysis of

Biogeography when the Number of Areas is Large. Syst. Biol. 62: 789–804.

Lindgren, D., and E. Wilde. 2003. Growing penstemons: species, cultivars and hybrids. American Penstemon Society, Infinity Publishing, Haverford, PA.

Lodewick K., and R. Lodewick. 1987. *Penstemon* nomenclature. American Penstemon Society, Eugene, OR.

Louca, S., and M. W. Pennell. 2020. Extant timetrees are consistent with a myriad of diversification histories. Nature 580: 502–505.

Magallón, S., and M. J. Sanderson. 2002. Absolute diversification rates in angiosperm clades. Evolution 55: 1762–1780.

Magallón, S., and A. Castillo. 2009. Angiosperm diversification through time. Am. J. Bot. 96: 349–365.

Matzke, N. J. 2013a. BioGeoBEARS: BioGeography with Bayesian (and Likelihood) Evolutionary Analysis in R Scripts. University of California, Berkeley. Berkeley, CA.

Matzke, N. J. 2013b. Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. Frontiers in Biogeography 5: 242–248.

Matzke, N. J. 2014. Model Selection in Historical Biogeography Reveals that Founder-event Speciation is a Crucial Process in Island Clades. Syst. Biol. 63: 951–970.

Moore, B. R., S. Höhna, M. R. May, B. Rannala, and J. P. Huelsenbeck. 2016. Critically evaluating the theory and performance of Bayesian analysis of macroevolutionary mixtures. Proc. Natl. Acad. Sci. U.S.A. 113: 9569–9574.

O'Kane, S. L. Jr., and K. D. Heil. 2014. *Penstemon bleaklyi* (Plantaginaceae), a new high-elevation species from north-central New Mexico. Phytoneuron 2014-61: 1-4.

720 Pennell, F. W. 1935. The Scrophulariaceae of eastern temperate North America. Academy of
721 Natural Sciences of Philadelphia, Philadelphia, Pennsylvania, USA.
722 Pennell, M. W., J. M Eastman, G. J. Slater, J. W. Brown, J. C. Uyeda, R. G. FitzJohn, M. E.
723 Alfaro, and L. J. Harmon. 2014. geiger v2.0: an expanded suite of methods for fitting
724 macroevolutionary models to phylogenetic trees. Bioinformatics 30: 2216–2218.
725 Plummer, M., N. Best, K. Cowles, and K. Vines. 2006. CODA: Convergence Diagnosis and
726 Output Analysis for MCMC. R News 6: 7–11.
727 Price, M.N., P. S. Dehal, and A. P. Arkin. 2010. FastTree 2 – Approximately Maximum-
728 Likelihood Trees for Large Alignments. PLoS ONE 5: e9490.
729 R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for
730 Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
731 Rabosky, D. L. 2010. Extinction rates should not be estimated from molecular phylogenies.
732 Evolution 64: 1816–1824.
733 Rabosky, D. L. 2014 Automatic Detection of Key Innovations, Rate Shifts, and Diversity-
734 Dependence on Phylogenetic Trees. PLoS ONE 9: e89543.
735 Rabosky, D.L., M. Grundler, C. Anderson, P. Title, J. J. Shi, J. W. Brown, H. Huang, and J. G.
736 Larson. 2014. BAMMtools: An R package for the analysis of evolutionary dynamics on
737 phylogenetic trees. Methods Ecol. Evol. 5: 701–707.
738 Rambaut, A., A. J. Drummond, D. Xie, G. Baele, and M. A. Suchard. 2018. Posterior
739 Summarization in Bayesian Phylogenetics Using Tracer 1.7. Syst. Biol. 67: 901–904.
740 Ree, R. H., and S. A. Smith. 2008. Maximum Likelihood Inference of Geographic Range
741 Evolution by Dispersal, Local Extinction, and Cladogenesis. Syst. Biol. 57: 4–14.
742 Rehfeldt, G. E. 1999. Systematics and genetic structure of *Ponderosae* taxa (Pinaceae) inhabiting

743 the mountain islands of the southwest. Am. J. Bot. 86: 741-752.

744 Rodríguez–Peña, R. A., R. L. Johnson, L. A. Johnson, C. D. Anderson, N. J. Ricks, K. M. Farley,
745 M. D. Robbin, A. D. Wolfe, and M. R. Stevens. 2018. Investigating the genetic diversity and
746 differentiation patterns in the *Penstemon scariosus* species complex under different sample
747 sizes using AFLPs and SSR. Conserv. Genet. 19: 1335–1348.

748 Ronquist, F. 1997. Dispersal-Vicariance Analysis: A New Approach to the Quantification of
749 Historical Biogeography. Syst. Biol. 46: 195–203.

750 Rothfels, C. J., K. M. Pryer, and F.-W. Li. 2017. Next-generation polyploid phylogenetics: rapid
751 resolution of hybrid polyploid complexes using PacBio single-molecule sequencing. New
752 Phytol. 213: 413–429.

753 Schmidel, C.C. 1763-1771: Icones Plantarum. ed. Keller. Fleischmann, Nürnberg.

754 Schwery, O. R. E. Onstein, Y. Bouchenak-Khelladi, Y. Xing, R. J. Carter, and H. P. Linder.
755 2015. As old as the mountain: the radiations of the Ericaceae. New Phytol. 207: 355-367.

756 Smith, S. A., and C. Dunn. 2008. Phyutility: a phyloinformatics utility for trees, alignments, and
757 molecular data. Bioinformatics 24: 715–716.

758 Smith, S. A., and B. C. O'Meara. 2012. treePL: divergence time estimation using penalized
759 likelihood for large phylogenies. Bioinformatics 28:2689–2690.

760 Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of
761 large phylogenies. Bioinformatics 30: 1312–1313.

762 Stamatakis, A., P. Hoover, and J. Rougemont. 2008. A rapid bootstrap algorithm for the RAxML
763 web servers. Syst. Biol. 57: 758–771.

764 Stevens, M. R., S. L. Love, and T. McCammon. 2020. The heart of *Penstemon* country: a natural
765 history of penstemons in the Utah region. Sweetgrass Books, Helena, MT. 394 pp.

766 Stone, B. W., A. Ward, M. Farenwald, A. Lutz, and A. D. Wolfe. 2019. Genetic diversity and
767 population structure in Cary's Beardtongue, *Penstemon caryi* (Plantaginaceae), a rare plant
768 endemic to the eastern Rocky Mountains of Wyoming and Montana. Conserv. Genet. 20:
769 1149–1161.
770 Straw, R. M. 1956a. Adaptive morphology of the *Penstemon* flower. Phytomorphology 6: 112-
771 119.
772 Straw, R. M. 1956b. Floral isolation in *Penstemon*. Am. Nat. 90: 47-53.
773 Straw, R. M. 1963. Bee-fly pollination of *Penstemon ambiguous*. Ecology 44: 818-819.
774 Straw, R. M. 1966. A redefinition of *Penstemon* (Scrophulariaceae). Brittonia 18: 80-95.
775 Tank, D. C., J. M. Eastman, M. W. Pennell, P. S. Soltis, D. E. Soltis, C. E. Hinchliff, J. W.
776 Brown, E. B. Sessa, and L. J. Harmon. 2015. Nested radiations and the pulse of angiosperm
777 diversification: increased diversification rates often follow whole genome duplications. New
778 Phytol. 207: 454-467.
779 Turner, B. L. 2010. Taxonomy of the *Penstemon campanulatus* complex (Scrophulariaceae) and
780 description of a new species from its midst. Phytoneuron 2010-31: 1-5.
781 Uribe-Convers, S., M. L. Settles, and D. C. Tank. 2016. A Phylogenomic Approach Based on
782 PCR Target Enrichment and High Throughput Sequencing: Resolving the Diversity within
783 the South American Species of Bartsia L. (Orobanchaceae). PloS One 11: e0148203.
784 USFWS (United States Fish and Wildlife Service). 1993. Endangered and threatened wildlife
785 and plants; review of plant taxa for listing as endangered or threatened species. Federal
786 Register 58: 51144–51180.
787 USFWS (United States Fish and Wildlife Service). 2011. Endangered and threatened wildlife
788 and plants; designation of critical habitat for *Ipomopsis polyantha* (Pagosa Skyrocket) and

789 threatened status for *Penstemon debilis* (Parachute Beardtongue) and Phacelia submutica
790 (DeBeque Phacelia). Federal Register 76: 45078–45128.
791 Vargas, P., L. M. Valente, J. L. Blanco-Pastor, I. Liberal, B. Guzman, E. Cano, A. Forrest, and
792 M. Fernandez-Mazuecos. 2014. Testing the biogeographical congruence of palaeofloras
793 using molecular phylogenetics: snapdragons and the Madrean–Tethyan flora. J. Biogeogr.
794 41:932–943.
795 Verboom, G. A., N. G. Bergh, S. A. Haiden, V. Hoffman, and M. N. Britton. 2015. Topography
796 as a driver of diversification in the Cape Floristic Region of South Africa. New Phytol. 207:
797 368-376.
798 Wessinger, C. A., C. C. Freeman, M. E. Mort, M. D. Rausher, and L. C. Hileman. 2016.
799 Multiplexed shotgun genotyping resolves species relationships within the North American
800 genus *Penstemon*. Am. J. Bot. 103: 912–922.
801 Wessinger, C. A., M. D. Rauscher, and L. C. Hileman. 2019. Adaptation to hummingbird
802 pollination is associated with reduced diversification in *Penstemon*. Evol. Lett. 3: 521–533.
803 Wolfe, A. D. 2005. ISSR techniques for evolutionary biology. Methods Enzymol. 395: 134–144.
804 Wolfe, A. D., Q.-Y. Xiang, and S. R. Kephart. 1998a. Assessing hybridization in natural
805 populations of *Penstemon* (Scrophulariaceae) using hypervariable intersimple sequence
806 repeat (ISSR) bands. Mol. Ecol. 7: 1107–1125.
807 Wolfe, A. D., Q.-Y. Xiang, and S. R. Kephart. 1998b. Diploid hybrid speciation in *Penstemon*
808 (Scrophulariaceae). Proc. Natl. Acad. Sci. U.S.A. 95: 5112–5115.
809 Wolfe, A. D., S. L. Datwyler, and C. P. Randle. 2002. A phylogenetic and biogeographic
810 analysis of the Cheloneae (Scrophulariaceae) based on ITS and matK sequence data. Syst.
811 Bot. 27: 138–148.

812 Wolfe, A. D., A. McMullen-Sibul, V. J. Tepedino, L. Kubatko, T. Necamp, and S. Fassnacht.
813 2014. Conservation genetics and breeding system of *Penstemon debilis* (Plantaginaceae), a
814 rare beardtongue endemic to oil shale talus in western Colorado, USA. J. Syst. Evol. 52:
815 598–611.
816 Wolfe, A. D., T. Necamp, S. Fassnacht, P. Blischak, and L. Kubatko. 2016. Population genetics
817 of *Penstemon albomarginatus* (Plantaginaceae), a rare Mojave Desert species of deep sand
818 habitats. Conserv. Genet. 17: 1245–1255.
819 Wolfe, A. D., C. P. Randle, S. L. Datwyler, J. J. Morawetz, N. Arguedas, and J. Diaz. 2006.
820 Phylogeny, taxonomic affinities, and biogeography of *Penstemon* (Plantaginaceae) based on
821 ITS and cpDNA sequence data. Am. J. Bot. 93: 1699–1713.
822 Zacarías-Correa, A. G. 2020. Diversidad, estructura genética y modelos de distribucíon del
823 genero *Penstemon* (Plantaginaceae) en México. Ph.D. thesis, Instituto de ecologia, Xalapa,
824 Veracruz, México.
825 Zacarías-Correa, A. G., A. Lira-Noriega, E. Pérez-Calix, M.-S. Samain, and A. D. Wolfe. 2020.
826 Back to the future of a rare plant species of the Chihuahuan desert: tracing distribution
827 patterns across time and genetic diversity as a basis for conservation actions. Biodivers.
828 Conserv. 291821–1840.
829 Zacarías-Correa, A. G., A. D. Wolfe, E. M. Salas, and M-S. Samain. 2019. *Penstemon*
830 *reidmoranii* (Plantaginaceae), a new species from Baja California. Phytotaxa 387: 63-70.
831 Zhang, C., M. Rabiee, E. Sayyari, and S. Mirarab. 2018. ASTRAL-III: polynomial time species
832 tree reconstruction from partially resolved gene trees. BMC Bioinf. 19: 153.
833
834

835 Figure 1. Phylogeny of *Penstemon* and related members of Cheloneae based on the Astral
836 analysis. Relative branch lengths are shown on the right. The vertical bar with lettering refers to
837 insets 1a–1f, which show details of the phylogenetic tree.
838
839 Figure 2. Taxonomic bootstrap results for a) subgenera, b) sections, and c) subsections.
840
841 Figure 3. Time-calibrated tree for Chelonae and *Penstemon* based on fossils for members of
842 Lamiales. Dark circles mark nodes that were calibrated using fossils from Vargas et al. (2014). Grey
843 squares mark nodes that were used as secondary calibration points for dating the entire 293-taxon
844 phylogeny with treePL.
845
846 Figure 4. Inferred patterns of diversification using BAMM. The phylogeny shows the
847 reconstructed rates of net diversification inferred across the tree with a single shift in
848 diversification (best shift configuration) at the base of *Penstemon*. The top inset plot shows the
849 distribution of net diversification rates for the entire phylogeny and the bottom inset plot shows
850 the inferred net diversification rate through time for Chelonae without *Penstemon* (left), the
851 Chelonae with *Penstemon* (center), and *Penstemon* alone (right).
852