## Molecular Phylogenetics and Evolution 84 (2015) 266-283

Contents lists available at ScienceDirect





## Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

## Multilocus phylogenetic inference in subfamily Chlorogaloideae and related genera of Agavaceae – Informing questions in taxonomy at multiple ranks



## Jenny K. Archibald<sup>a,\*</sup>, Susan R. Kephart<sup>b</sup>, Kathryn E. Theiss<sup>b</sup>, Anna L. Petrosky<sup>c</sup>, Theresa M. Culley<sup>d</sup>

<sup>a</sup> Biodiversity Institute and Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, KS 66045, USA

<sup>b</sup> Department of Biology, Willamette University, Salem, OR 97301, USA

<sup>c</sup> Department of Integrative Biology, University of California, Berkeley, CA 94709, USA

<sup>d</sup> Department of Biological Sciences, University of Cincinnati, Cincinnati, OH 45221, USA

### ARTICLE INFO

Article history: Received 21 July 2014 Revised 8 December 2014 Accepted 16 December 2014 Available online 10 January 2015

Keywords: Agavaceae Chlorogaloideae Chlorogalum Hesperaloe Hesperocallis ITS

## ABSTRACT

A series of taxonomic questions at the subfamilial, generic, and intrageneric levels have remained within subfamily Chlorogaloideae s.s. (comprising Camassia, Chlorogalum, Hastingsia, and Schoenolirion) and relatives in Agavaceae. We present the first phylogenetic hypotheses focused on Chlorogaloideae that are based on multiple independent loci and include a wide sampling of outgroups across Agavaceae. In addition to chloroplast regions ndhF and trnL-trnF, we used nrDNA ITS for phylogenetic inference. Incomplete concerted evolution of the latter is indicated by intra-individual site polymorphisms for nearly half of the individuals. Comparisons of four coding and analysis methods for these characters indicate that the region remains phylogenetically informative. Our results confirm that Chlorogaloideae s.s. is not monophyletic, due to the close relationship of Schoenolirion with Hesperaloe and Hesperoyucca, as well as the likely sister relationship between Hesperocallis and core Chlorogaloideae (Camassia, Chlorogalum, and Hastingsia). Chlorogalum is also not monophyletic, being divided with strong support into vespertine and diurnal clades. This study produced the first phylogenetic hypotheses across Hesperaloe, allowing initial tests of several taxonomic disagreements within this genus. Our results reveal the lack of cohesion of H. funifera, indicating that H. funifera ssp. funifera may be more closely related to H. campanulata than to H. funifera ssp. chiangii (=H. chiangii). With potential gene flow between many members of Hesperaloe and a possible hybrid origin for *H. campanulata*, the genetic relationships within this genus appear complex. Further population-level investigation of many of the taxa in Chlorogaloideae s.l. would benefit our understanding of the evolution and taxonomy of these groups; Camassia and Hastingsia are the current focus of ongoing study.

Published by Elsevier Inc.

### 1. Introduction

Although there are groups of plants whose taxonomic limits have been clearly and robustly inferred since the time of Linnaeus (e.g., Asteraceae; Funk et al., 2009), lilies and their relatives are certainly not among them. Massive family-level rearrangements of Liliaceae s.l. into potentially over 20 families have resulted from efforts to discern their evolutionary relationships (e.g., APG, 2003, 2009; Cronquist, 1981). One challenging family within this complex group of plants is Agavaceae, also known as Agavoideae in Asparagaceae (APG, 2009; Chase et al., 2009). Some members of this family have received significant attention, such as the classic study system of *Yucca* – yucca moth mutualisms (Pellmyr, 2003) and economically-relevant taxa (e.g., FNA, 1993+; Zizumbo-Villarreal et al., 2013). Within Agavaceae, the less wellstudied Chlorogaloideae s.l. (Halpin and Fishbein, 2013) provides many potential avenues for evolutionary studies. This group comprises a diverse assemblage of species that inhabit deserts to wetlands (FNA, 1993+); the distributions of some species cover broad regions of the United States while others are narrowly endemic on serpentine soils (Halpin and Fishbein, 2013). These rosette-forming plants vary greatly in morphology. For example, they range in height from 18 to 400 cm (Hochstätter, 2009; Sherman, 1969; Starr, 1997) and vary in floral shape among tubular, campanulate, and rotate types (zygomorphic and actinomorphic;

<sup>\*</sup> Corresponding author.

*E-mail addresses:* jkarchibald@yahoo.com (J.K. Archibald), skephart@willamette. edu (S.R. Kephart), kathryn.theiss@gmail.com (K.E. Theiss), annapetrosky@berkeley. edu (A.L. Petrosky), theresa.culley@uc.edu (T.M. Culley).

FNA, 1993+; Starr, 1997). As in Agavaceae as a whole, Chlorogaloideae includes many plants utilized by humans, such as those used in horticulture (e.g., McGary, 2001; Starr, 1997), for pulp and paper applications (Sanchez et al., 2011), and as historically important food sources (Mehalchick et al., 2004; Moerman, 1986).

Chlorogaloideae s.s. (Speta, 1998) has traditionally included four genera with up to 18 species and 29 total infrageneric taxa: *Camassia* (camas), *Chlorogalum* (soap plants), *Hastingsia* (rush lilies), and *Schoenolirion* (rush lilies; Fig. 1 and Table 1). The latter three genera were first recognized as subtribe Chlorogaleae (Watson, 1879). This grouping was maintained for nearly 100 years, although Engler (1887) added *Hemiphylacus* and included *Hastingsia* in *Schoenolirion* within his Chlorogalinae. A link between *Camassia* and the original three genera of subtribe Chlorogalinae was supported by the cytological and morphological study of *Schoenolirion* by Sherman (1969), who also suggested that *Hemiphylacus* be excluded from the subtribe. Later, Speta (1998) officially proposed Chlorogaloideae as one of five subfamilies in Hyacinthaceae.

Considerable taxonomic fluidity characterizes Chlorogaloideae, including at the family level. Some genera have been placed in Liliaceae (Baker, 1873; Engler, 1887; Watson, 1879), Anthericaceae (Schulze, 1982), Hyacinthaceae (Batsch, 1786; Dahlgren et al., 1985; Speta, 1998), Camassiaceae (Cupov, 1994), Chlorogalaceae (Hoogland and Reveal, 2005), and Agavaceae. The inclusion of Chlorogaloideae in Agavaceae is well supported by phylogenetic evidence and by their shared bimodal karyotype, i.e., with chromosomes of two distinct size classes (Cave, 1970; McKain et al., 2012; Sato, 1935). A potential relationship of Chlorogaloideae with Hesperaloe, Hesperocallis, and Hesperoyucca (Fig. 1 and Table 1) is also supported by recent phylogenetic analyses (Halpin, 2011; Halpin and Fishbein, 2013). For simplicity, we will refer to these three genera plus the four genera of Chlorogaloideae s.s. as Chlorogaloideae s.l. (following Halpin and Fishbein, 2013, who summarized key aspects of their taxonomic history).

Preliminary phylogenetic insights for some members of Chlorogaloideae s.l. have been provided by several studies. Based on sampling 1–3 species per genus, both a *Camassia – Chlorogalum* clade and a *Hesperaloe – Hesperoyucca* clade were supported as members of Agavaceae s.l. by Bogler et al. (2006; using ITS, *ndhF*, and *rbcL*) and Smith et al. (2008; using *trnL–trnF*), but relationships with other members of Agavaceae have not been well supported. *Hesperocallis* was also sampled by Bogler et al. (2006) and placed in Agavaceae. Other phylogenetic studies of Agavaceae, Hyacinthaceae, and Asparagales similarly resolved sampled genera of Chlorogaloideae s.l. within Agavaceae (Good-Avila et al., 2006; Pfosser and Speta, 1999; Seberg et al., 2012). Although helpful in clarifying the family-level placement of Chlorogaloideae s.l., the broad focus of each of these studies led to very low sampling within the subfamily.

Halpin and Fishbein (2013) recently provided the first phylogenetic hypotheses focused on Chlorogaloideae s.s., using four cpDNA regions (rpl16, trnD-trnT, psbJ-petA, and trnS-trnfM). They sampled at least one population from each species in Chlorogaloideae s.s. and Hesperocallis undulata, one species from Hesperaloe and Hesperoyucca, and seven outgroups; this was the first inclusion of Schoenolirion in a phylogenetic study. Their tree resolved a clade that they designated "Core Chlorogaloideae" (Camassia, Chlorogalum, and Hastingsia), with Hesperocallis weakly supported as sister, followed by their "SHH" clade comprising Schoenolirion, Hesperaloe, and Hesperoyucca. Concordantly, recent analyses of whole chloroplast genomes strongly place Camassia, Chlorogalum, and Hesperocallis in a clade sister to a Schoenolirion - Hesperaloe - Hesperoyucca clade; in that case only one species was sampled for each genus except Hesperaloe (2 spp.; Michael McKain, pers. comm.). Halpin and Fishbein (2013) highlighted the need for more data, such as independent loci, and posed the possibility of dividing Chlorogaloideae s.s. or expanding the subfamily to encompass core Chlorogaloideae, *Hesperocallis*, and the SHH clade.

The present study focuses on providing some essential missing links in our understanding of phylogenetic relationships in these genera. The monophyly of four of the seven genera in Chlorogaloideae s.l. was supported by the phylogeny of Halpin and Fishbein (2013), with three genera remaining unsupported: Hesperaloe (not tested), Hesperoyucca (not tested), and Chlorogalum (paraphyletic). At the intrageneric level, only the phylogenetic analysis of Halpin and Fishbein (2013) evaluated Chlorogalum and Schoenolirion, although separate chloroplast phylogenies are available for Camassia (Fishbein et al., 2010) and Hastingsia (Halpin, 2011); each tree provided an initial outline of relationships, but with many unanswered questions. A multilocus phylogenetic hypothesis for both genera is being developed as part of a separate integrative taxonomic study (S. Kephart, I. Archibald, T. Cullev, K. Theiss, unpub.), Hesperocallis is monotypic, but the phylogenetic relationships within Hesperaloe and Hesperoyucca are not well known. Only one or two species in these two genera had been sampled for broad phylogenetic studies, limiting assessment of relationships across all Chlorogaloideae s.l. prior to our work.

No previous phylogenetic analyses have focused on relationships within *Hesperaloe*, although the ITS phylogeny of Clary (2001) included three of the 6–9 taxa in the genus and strongly supported its monophyly. *Hesperaloe* has been consistently placed in Agavaceae (see FNA, 1993+), with some species formerly recognized as members of *Yucca* or *Aloe*. The exact number and ranks of taxa within *Hesperaloe* vary depending on the taxonomic treatment (Hochstätter, 2009; Hochstätter and Martínez-Ávalos, 2010; Starr, 1997).

*Hesperoyucca* currently encompasses three species (Clary, 2001), with *Hesperoyucca whipplei* being divided into 0–5 infraspecific taxa recognized at the variety or subspecies level (e.g., Haines, 1941; McKinney and Hickman, 1993). Formerly part of *Yucca* (e.g., Bogler et al., 2006; McKelvey, 1938, 1947), the genus is now recognized as distinct based on morphological, phenological, pollination, and phylogenetic criteria (Clary, 2001). Bogler et al. (2006) placed *Hesperoyucca* in Agavaceae s.s. as a close relative of *Hesperaloe*. The first phylogeny focused on this genus was the ITS tree of Clary (2001), with two populations of *H. whipplei* and one of each of the other two species. The genus as a whole and *H. whipplei* were each well supported as monophyletic.

Hesperocallis is a monotypic genus whose relationship to other monocots has been a long-standing puzzle. Treated alternatively in Liliaceae, Hemerocallidaceae, Funkiaceae (=Hostaceae), and Hesperocallaceae (putatively related to Alliaceae), it is placed currently in Agavaceae (Pires et al., 2004), consistent with its bimodal karyotype. Pires et al. (2004) supported this placement with the first phylogeny that included Hesperocallis. With greater sampling in the family, Bogler et al. (2006) also strongly supported Hesperocallis in an Agavaceae s.l. clade sister to an "extended Agavaceae" clade. We follow Chase et al. (2009) in considering Agavaceae to encompass both clades. Analyses by Halpin and Fishbein (2013) intriguingly placed Hesperocallis as sister to core Chlorogaloideae, but lacked strong support. To our knowledge, no prior phylogenies have strongly-supported resolution of the close relatives of Hesperocallis. Few morphological studies have explicitly linked this genus with the genera of Chlorogaloideae s.s. However, in a monograph of Chlorogalum, Hoover (1940) considered Camassia, Hastingsia, and Schoenolirion as most closely related to Chlorogalum, followed by Hesperocallis and Odontostomum Torr. The latter is now known to be well separated from Agavaceae (Stevens, 2001+). Gould (1942) similarly noted an affinity among Camassia, Chlorogalum, Hesperocallis, and Schoenolirion. Regardless, the exact placement of Hesperocallis within Agavaceae remains uncertain.



**Fig. 1.** Photographs across Chlorogaloideae s.l., representing a subset of the morphological diversity. (A) *Camassia leichtlinii* ssp. *leichtlinii*, inflorescences. (B and C) *C. angusta*, plant in bud (note keeled leaves), flower (note 3-lobed stigma). (D and E) *Hastingsia alba*, inflorescence, bulbs. (F) *H. atropurpurea*, inflorescence. (G) *Chlorogalum pomeridianum* var. *pomeridianum*, inflorescence. (H) *Ch. angustifolium*, inflorescence. (I–K) *Hesperocallis undulata*, plant, bulb, flowers. (L) *Schoenolirion wrightii*, inflorescence. (M) *S. croceum*, inflorescences. (N) *Hesperaloe funifera* spp. *chiangii*, vegetative plants. (O) *H. funifera* ssp. *funifera*, leaf cross sections. (P and Q) *H. engelmannii*, flower and dehisced capsule, flower with three tepals removed (note stout style). (R and S). *Hesperoyucca peninsularis*, plants and habitat, flowers. (T) *H. newberryi*, carpel and stamens (note apparently capitate stigma with trichomes). Photo credits: (A, B, and O) JKA; (C) Jim Kephart; (D–F) Linnea Hardlund; (G and H) Julie Kiersted Nelson; (I) Sue Carnahan via SEINet.org; (J and T) Wendy Hodgson; (K) Linda Prince; (L) Louisiana Natural Heritage Program; (M) Mason Brock; (N) Greg Starr; (P and Q) James Henrickson; (R and S) James Riley.

In all, previous studies have contributed new insights yet also challenging questions for understanding the evolutionary relationships of the taxonomically complex 'lilies' known as Chlorogaloideae. Here we provide a nuclear and chloroplast phylogeny of Chlorogaloideae s.l., with nearly complete taxon sampling of all putative species and infraspecific taxa. By sampling three genetic regions previously used for study within Agavaceae (ITS, *ndhF*, *trnL-trnF*), we were able to incorporate a wide array of outgroups across Agavaceae. Our goals were to: (1) Provide the first phylogenetic test based on multiple independent loci of the monophyly of Chlorogaloideae s.s. and of each of the seven study genera. (2) Infer the placement of each genus of Chlorogaloideae s.l. relative to the major clades of Agavaceae, with particular focus on the enigmatic *Hesperocallis.* (3) Produce the first phylogenetic tree across *Hesperaloe*, allowing preliminary phylogenetic tests of the differing taxonomic schemes proposed for this genus. (4) Develop a broad phylogenetic framework for subsequent species-delimitation and speciation studies within *Camassia* and *Hastingsia*.

## 2. Materials and methods

## 2.1. Taxon collection and DNA extraction

Leaf material (ca. 20 mg or more per individual) was collected on silica gel from each population. DNA was extracted from leaf samples using standard CTAB methods (Doyle and Doyle, 1987) or a DNeasy<sup>®</sup> Plant Mini Kit (Qiagen, Valencia, California, USA). Initial analyses included 119 accessions, which were pruned to 92, 87, and 95 accessions for the final concatenated (ITS. *ndhF*. and *trnL–F*). ITS, and cpDNA (*ndhF* and *trnL*–*F*) analyses, respectively (Table 2 and Fig. 2). The differences in sampling among loci are due to different availability of outgroup (OG) sequences from GenBank (accessions designated "GB"). Outgroups were included in the final concatenated analysis if sequences from at least two of the three DNA regions were available. Other accessions were removed for the final analyses if there was evidence of hybridization, cultivation, unconfirmable taxonomic identification, or very long phylogenetic branches (see below). Final sampling of 87-95 accessions included 41 putative taxa of Chlorogaloideae s.l. and up to 14 OG taxa. At least one, and usually two or more, populations were sampled from all taxa within Chlorogaloideae s.l. (Tables 1 and 2), except the putative infraspecific taxa of *Hesperovucca* whipplei and the rare Mexican Hesperaloe malacophylla (Hochstätter and Martínez-Ávalos, 2010). Outgroups represent each of the major clades of Agavaceae supported in Bogler et al. (2006); we based the choice of individuals for rooting on consistent relationships in previous phylogenies focused at larger taxonomic scales (Bogler et al., 2006; Chen et al., 2013; Seberg et al., 2012).

In addition to the outgroup targeting noted above, three other taxa were included from GenBank because prior taxonomic hypotheses suggested a relationship to the ingroup (i.e., Chlorogaloideae s.l., Table 1). These included *Hemiphylacus* S. Watson, formerly a member of subtribe Chlorogalinae (Engler, 1887; Krause, 1930), and *Oziroe biflora* (Ruiz & Pav.) Speta, formerly placed in *Camassia* (Tropicos.org). Previous analyses (e.g., Bogler et al., 2006) resolved a more distant relationship for *Hemiphylacus* and *Oziroe* with Agavaceae. Currently, *Hemiphylacus* is in Asparagaceae/Asparagoideae, and *Oziroe* is in Hyacinthaceae (or Scilloideae of Asparagaceae; Chase et al., 2009). Our initial analyses confirmed this distant relationship, placing both accessions on very long branches. We excluded them from final analyses to avoid

long-branch attraction or repulsion (Felsenstein, 1978; Siddall, 1998). Finally, Yucca queretaroensis was included. Pellmyr et al. (2007) resolved this species as sister to the rest of Yucca using AFLP markers. They recommended further phylogenetic tests with greater sampling of Hesperaloe and Hesperoyucca to verify the generic placement of this species. Smith et al. (2008) sequenced trnL-trnF across Agavaceae, focusing on Yucca. They sequenced an accession of Y. queretaroensis that may be influenced by introgression (designated GB2 in this study) and included data from an accession sampled by Good-Avilla et al. (2006; GB3). The latter was sister to the remainder of Yucca, but with weak support and alternate placements depending on details of analysis (including a possible placement as sister to Hesperoyucca). The other accession was nested within Yucca. We included each sequence in separate analyses as well as an *ndhF* sequence from another population (Y. *queret*aroensis GB1). Our analyses showed no major differences in the inferred relationships regardless of the accession(s) used of Y. queretaroensis, likely due to much smaller sampling within Yucca (3 species) compared to Pellmyr et al. (2007) and Smith et al. (2008).

GenBank sequences for different gene regions from the same species were included in initial analyses as different putative individuals (numbered GB1, GB2, etc.), unless voucher citations from source papers verified origin from the same individual. Initial analyses confirmed consistent phylogenetic placement for each species, and their sequences were concatenated for subsequent combined analyses. This is indicated in accession labels; GB12 indicates concatenation of sequences GB1 and GB2 for a species.

## 2.2. DNA sequencing

We amplified one nuclear and two chloroplast regions: the nrDNA internal transcribed spacer regions (ITS-1 – 5.8S – ITS-2; Baldwin et al., 1995; Soltis and Soltis, 1998), as well as *ndhF* (Terry et al., 1997) and the *trnL* intron, *trnL* exon, and *trnL*-*trnF* spacer region (referred to here as the *trnL*-F region) from the



**Fig. 2.** Map of the United States and Mexico, showing all known locations for the sampled populations of Chlorogaloideae s.l.; the original sites for *Hesperaloe campanulata* MMsn and *H. parviflora* ssp. *bechtoldii* JH24815 are unknown and not mapped. Site code is indicated next to the symbol. Taxon or clade membership is indicated by symbol shape and color (see Fig. 3). Some symbols differ in size to facilitate visibility.

#### Table 1

Previously-hypothesized groups that include some genera of Chlorogaloideae s.l. and purported relatives. Genera that are not treated in a given grouping are indicated by an "-".

No. intrageneric taxa	Chlorogaloideae s.l.	Subfamily Chlorogaloideae s.s.	Schoenolirion and relatives	Chlorogalum and relatives	Subtribe Chlorogalinae	Subtribe Chlorogaleae
FNA (1993+) and others <sup>a</sup>	Halpin and Fishbein (2013)	Speta (1998)	Sherman (1960)	Hoover (1940)	Engler (1887)	Watson (1879)
6 spp., 10 sspp.	Camassia	Camassia	Camassia	Camassia	-	-
5 spp., 5 vars.	Chlorogalum Core Chlorogaloideae	Chlorogalum	Chlorogalum	Chlorogalum	Chlorogalum	Chlorogalum
2–4 spp., 0–2 vars.	Hastingsia	Hastingsia	Hastingsia	Hastingsia	Hastingsia <sup>b</sup>	Hastingsia
3 spp.	Schoenolirion	Schoenolirion	Schoenolirion	Schoenolirion	Schoenolirion	Schoenolirion
5–9 spp., 0–2 sspp.	Hesperaloe SHH clade	-	-	-	-	-
3 spp., 0–5 sspp., 0–4 vars.	Hesperoyucca	-	-	-	-	-
1 sp.	Hesperocallis	-	-	Hesperocallis	-	-
-	-	-	-	- Odontostomum	Hemiphylacus -	-

<sup>a</sup> See text: Becking (1986), Hochstätter (2009), Hochstätter and Martínez-Ávalos (2010), Lang and Zika (1997), Starr (1997).

<sup>b</sup> Merged within Schoenolirion.

chloroplast genome (Soltis and Soltis, 1998). Standard primers were employed for ITS (Table 3; Wen and Zimmer, 1996). The sequencing of *trnL–F* was complicated by several polyT regions, including one near the beginning of the region that was disruptive enough to require design of a new forward primer ("c2," Table 3). Amplifications also used the standard reverse primer "f" and internal primers "d" and "e" when necessary (Taberlet et al., 1991). Other polyT regions, usually near the center of *trnL–F*, disrupted sequencing in at least 64% of the individuals, but it was possible to compensate via additional sequencing. Primers for *ndhF* were either standard or slightly modified versions of those from Terry et al. (1997; Table 3), being used to amplify the region in two or three fragments. High quality DNAs were amplified in two fragments, using primer combinations 032F – 1318bR and 1101bF – 2110R.

PCR reactions (25-50 µl) consisted of 0.5-1 µl DNA, 0.4 µM of each primer, and  $1 \times$  Premium Bullseye Tag DNA Polymerase Mix (75 mM Tris-HCl pH 8.5, 20 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 1.5 mM MgCl<sub>2</sub>, 0.1% Tween 20, 0.2 mM dNTPs, 0.025 units/µl Bullseye Taq polymerase, inert red dye and stabilizer; Midsci, Saint Louis, Missouri, USA). 10% DMSO was added to ITS reactions and 5% DMSO to trnL-F reactions. A Labnet MultiGene thermocycler (Labnet International Inc., Edison, New Jersey, USA) was used with the following settings: 5 min at 94 °C; 35 cycles of 1 min at 94 °C, 1 min at 50 °C (for ITS) or 48 °C (for *trnL–F* and *ndhF*), and 2 min at 72 °C; 10 min at 72 °C. Each thermocycler run included a negative control reaction with all reagents except for the DNA. PCR products were purified using ExoSAP-IT (Affymetrix, Santa Clara, California, USA) and sequenced by Macrogen USA (Rockville, Maryland, USA). Sequences have been deposited in GenBank (KP008151-KP008366; Table 2).

#### 2.3. Phylogenetic analyses and hypothesis testing

Individual contigs were edited, assembled, and aligned in Geneious v. 6.1.7 (Biomatters, Auckland, New Zealand). Alignments used MUSCLE settings and were adjusted by eye. Gaps were coded as characters using a modified version of the complex indel coding method of Simmons and Ochoterena (2000), as implemented by SeqState v. 1.4.1 (IndelCoder: modified complex coding option; Müller, 2005). Final analyses were run on concatenated, ITS, and cpDNA datasets using maximum parsimony (MP) and Bayesian

inference (BI). Maximum likelihood (ML) analyses were run in Garli v. 2.01 (Zwickl, 2006) for comparison; the topologies of the favored trees were very similar to those from BI and will largely not be discussed. Parsimony analyses were conducted in PAUP\* v. 4.0b10-x86 (Swofford, 2003) with the heuristic search option, Mul-Trees, no MaxTrees limit, swapping on all trees, and parsimony options set to collapse branches if minimum length is zero ("amb-"). Uninformative characters were excluded before all analyses and all remaining characters were equally weighted and unordered. One thousand quick searches with random taxon addition and NNI branch swapping were used to locate multiple islands of minimum-length trees, forming a starting pool of trees for more thorough searches employing TBR (Maddison, 1991). Separate TBR searches were run in different putative islands, but always inferred the same final relationships. Relative support for clades was assessed via jackknife (JK) analyses (5000 replicates). PAUP\* was set to emulate Jac resampling with character deletion at 37%. Initial analyses were run using both nucleotide and indel characters, indel characters alone, and a subset of the indel characters that excluded potentially problematic characters. Results were consistent with or without indel characters; final MP analyses included all indels except three formed from variable single nucleotide repeats in the cpDNA dataset. Maximum likelihood and BI analyses were conducted on the nucleotide datasets only. Model parameters were estimated using the Akaike Information Criterion (AIC) of jModelTest v. 2.1.3 (Posada, 2008). Bayesian analyses were conducted in MrBayes v. 3.2.1 (Ronquist et al., 2012) and were partitioned by gene region, allowing independent evolutionary models for each. Two runs of metropolis-coupled Markov chain Monte Carlo simulations each used four linked chains (three heated and one cold) and default priors for all model parameters. The analyses were halted when the average standard deviation of split frequencies was below 0.0008, except for the cpDNA analyses, for which a cut-off of 0.008 was sufficient (Table 4). Parameter values were sampled every 100 generations. Convergence was assessed using Tracer v. 1.6 (Rambaut and Drummond, 2007) and AWTY (Nylander et al., 2008; Wilgenbusch et al., 2004; "compare" and "cumulative" utilities).

Topological incongruence was assessed by visually comparing the well-supported nodes from each analysis, considering nodes with at least 70% bootstrap/jackknife support (see Kellogg et al., 1996; Mason-Gamer and Kellogg, 1996) or 0.90 posterior probabil-

## Table 2

Information for individuals sampled for phylogenetic analyses. Individual codes include site code and individual number (with collection year for *Camassia* and *Hastingsia*), or they are designated "GB" if those sequences were downloaded from GenBank. Voucher information includes collector(s), collection number, herbarium, and herbarium accession number if available. GenBank accession numbers are given for ITS, *trnL–F*, and *ndhF*. "MonAToL" indicates sequences pulled from genomic data gathered by M. McKain et al. for the Monocot Tree of Life project (http://www.botany.wisc.edu/givnish/monocotatol.htm). Inapplicable fields for a given individual are indicated by a "–".

Taxon	Individual	Locality	Voucher	ITS	trnL–F	ndhF
Agave dasylirioides Jacobi &	GB			U23999, U24019	-	DQ071892
Agave lechuguilla Torr.	GB1			U24015 U24000, U24020	-	DQ071893
Agave lechuguilla Torr.	GB2			-	EU092542	-
Agave striata Zucc.	GB			U24001, U24021	-	DQ071896
Anthericum liliago L.	GB			-	AF508513	AF508402
Didr.	GBI			-	-	AY191168
Behnia reticulata (Thunb.)	GB2			-	AF117007,	-
Beschorneria yuccoides K. Koch	GB			U24008, U24028	-	-
Camassia angusta (Engelm. & A. Gray) Blank.	OSA_1 (2012)	United States, Kansas, Osage Haymeadow, 38.7822, —95.67857	Archibald, King, and Sharkan 2012-4 (KANU 393280)	KP008295	KP008151	KP008223
Camassia angusta (Engelm. & A. Gray) Blank.	OTT_1M (2011)	United States, Indiana, Otterbein Prairie, 40.494493, –87.124091	Homoya and Dana 90- 06-14-66 (BUT 155014)	KP008296	KP008152	KP008224
Camassia cusickii S. Watson	HC_2 (2006)	United States, Oregon, Hell's Canyon, 45.126083. –116.836117	Kephart 600 (WILLU 50056)	KP008297	KP008153	KP008225
Camassia cusickii S. Watson	PC_3 (2006)	United States, Oregon, Pine Creek, 45.0578, –116.9035	Kephart 805 (WILLU 50117)	KP008298	KP008154	KP008226
Camassia howellii S. Watson	DC_1 (2012)	United States, Oregon, Dutcher Creek, 42.4233, –123.54135	Kephart, Bell, and Johnson 625 (WILLU 50091)	KP008299	KP008155	KP008227
Camassia howellii S. Watson	SM_1 (2006)	United States, Oregon, Sexton Mountain, 42.594583, –123.37113	Kephart 593 (WILLU 50057)	KP008300	KP008156	KP008228
Camassia leichtlinii (Baker) S. Watson ssp. leichtlinii	GL_4M (2007)	United States, Oregon, Glide, 43.301217, –123.225967	Kephart s.n. (WILLU 50067)	KP008301	KP008157	KP008229
Camassia leichtlinii (Baker) S. Watson ssp. leichtlinii	PS_23 (2007)	United States, Oregon, Popcorn Swale, 43.301317. –123.225967	Kotaich 105 (WILLU 50076)	KP008302	KP008158	KP008230
Camassia leichtlinii (Baker) S. Watson ssp. suksdorfii (Greenm.) Gould	BFV_L24 (2006)	United States, California, Butterfly Valley, 40.012283, –120.994119	Kephart and Theiss 604 (WILLU 50058)	KP008303	KP008159	KP008231
Camassia leichtlinii (Baker) S. Watson ssp. suksdorfii (Greenm) Gould	FC_1 (2006)	United States, Oregon, Fruitland Creek, 44.914334, –122.938023	LaDouceur and Dick s.n. (WILLU 50087)	KP008304	KP008160	KP008232
Camassia quamash (Pursh) Greene ssp. azurea (A. Heller) Gould	MM_1 (2007)	United States, Washington, Mima Mounds, 46.904633, —123.049567	Swift s.n. (WILLU 50059)	KP008305	KP008161	KP008233
Camassia quamash (Pursh) Greene ssp. azurea (A. Heller) Gould	WH_1 (2008)	United States, Washington, Wolf Haven, 46.90335, –122.848348	Kotaich 100 (WILLU 50085)	KP008306	KP008162	KP008234
Camassia quamash (Pursh) Greene ssp. breviflora Gould	SHM_1 (2006)	United States, California, Sagehen Meadow, 39.430217, –120.239651	Kephart 525 (WILLU 50055)	KP008307	KP008163	KP008235
Camassia quamash (Pursh) Greene ssp. breviflora	ZUM_20 (2008)	United States, Oregon, Zumwalt Prairie, 45.55966, —116.986354	Kephart 612 (WILLU 50079)	KP008308	KP008164	KP008236
Camassia quamash (Pursh) Greene ssp. intermedia	CG_11 (2008)	United States, Oregon, Cottage Grove, 43.758, -123.095	Kotaich 104 (WILLU 50080)	KP008309	KP008165	KP008237
Camassia quamash (Pursh) Greene ssp. intermedia	PS_1 (2007)	United States, Oregon, Popcorn Swale, 43.301317, –123.225967	Dennis 102 (WILLU 50071)	KP008310	KP008166	KP008238
Camassia quamash (Pursh)	LLV_1M	United States, California, Little Lake Valley,	Theiss and Hardlund s.n.	KP008311	KP008167	KP008239
Camassia quamash (Pursh)	(2012) TB_1M (2012)	39.46086, – 123.346616 United States, California, Table Bluff, 40.6953,	(WILLU 50098) Mesler 758 (WILLU 50073)	KP008312	KP008168	KP008240
Camassia quamash (Pursh)	BC_WS17	- 124.2706 United States, Oregon, Bethel Church,	Kephart 576 (WILLU	KP008313	KP008169	KP008241
Greene ssp. maxima Gould Camassia quamash (Pursh)	(2006) VIC_Q1 (2004)	45.041017, -123.183417 Canada, British Columbia, University of	Allen 1310 (WILLU 50050)	KP008314	KP008170	KP008242
Greene ssp. maxima Gould Camassia quamash (Pursh)	MTK_7 (2008)	Victoria, 48.460967, –123.3189 United States, Montana, Pattee Canyon,	Snustad-Clark s.n. (WILLU	KP008315	KP008171	KP008243
Greene ssp. quamash Camassia quamash (Pursh)	WP_1 (2008)	46.860283, –113.88695 United States, Idaho, Weippe Prairie,	50081) Kotaich 111 (WILLU 50082)	KP008316	KP008172	KP008244
Greene ssp. quamash Camassia auamash (Pursh)	CC 5 (2011)	46.349633, –115.9252 United States, Utah, Cherry Creek 41 931861	Wolf 924 (WILLI	KP008317	KP008173	KP008245
Greene ssp. <i>utahensis</i> Gould	(2011)	–111.77944	50097)	500517		

## Table 2 (continued)

Taxon	Individual	Locality	Voucher	ITS	trnI_F	ndhF
Camassia auamash (Pursh)	CM 1 (2008)	United States Oregon Carson Meadow	Kotaich 123 (WILLI 50083)	KP008318	KP008174	KP008246
Greene ssp. utahensis	civi_1 (2000)	44.922265, –117.190783	Kotaleli 125 (WILLO 50005)	11 000510	Ki 000174	NI 000240
Gould						
Camassia quamash (Pursh)	HM_1M (2013)	United States, Oregon, Hogue Meadow,	Archibald, Hardlund, and Kenhart s.n. (WILLU 50103)	KP008319	KP008175	KP008247
(Piper) Gould	(2013)	42.0308, -123.02333	Repliant S.II. (WILLO JOIOS)			
Camassia quamash (Pursh)	KR_1 (2007)	United States, Oregon, Ken Rose, 42.1286,	Dennis 113 (WILLU 50074)	KP008320	KP008176	KP008248
Greene ssp. <i>walpolei</i>		-123.6616				
(Piper) Gould Camassia scilloides (Raf.) Corv	BAR 1 (2012)	United States Kansas Barnhardt Prairie	Archibald 2011-4 (KANU	KP008321	KP008177	KP008249
	_ ( , , , , , , , , , , , , , , , , , ,	39.23107, -95.02493	393269)			
Camassia scilloides (Raf.) Cory	BIE_5 (2011)	United States, Indiana, Biesecker Prairie,	Schnabel s.n. (WILLU	KP008322	KP008178	KP008250
Camassia scilloides (Raf.) Corv	HAL 10	41.4178, -87.468283 United States Obio Hall's Creek	50075) Cartieri s.n. (CINC TMC13-	KP008323	KP008179	KP008251
cumusta semenass (nam) cory	(2011)	39.36731861, -84.15142363	1001)	10 000525	14 000170	11 000201
Chlorogalum angustifolium	AHA_1GH	United States, California, Ahart Ranch,	Ahart 1397 (UCD 138418)	KP008324	KP008180	KP008252
Kellogg Chlorogalum angustifolium	IKN 4	39.35775, -121.5136667 United States California Redding 40.661488	Nelson 2013-008 (WILLI	KP008325	KP008181	KP008253
Kellogg	JICIV_4	-122.3084	50108)	KI 000525	KI 000101	KI 000255
Chlorogalum grandiflorum	DAV182812_a	United States, California, BLM Pine Hills	CNPS SN Foothill Team	KP008326	KP008182	KP008254
Hoover Chlorogalum parviflorum S	LICP154092 a	Reserve, 38.76492237, –121.0214849	SNFN0219 (UCD 80426)	1/10002277	1/000102	KD008255
Watson	UCK154585_d	33.092, -117.288	154983)	KF006527	KF000105	KF008233
Chlorogalum parviflorum S.	SD166985_a	United States, California, Otay Mountain	Rebman, Gregory, and	KP008328	KP008184	KP008256
Watson Chlorogalum nomoridianum	4112050 1	Ecological Reserve, 32.60667, –116.88778	Rocks 10507 (SD 166985)	1/100000000	VD000105	VD000257
(DC.) Kunth var.	AH2958_1	Marine Reserve, 38,3138055, –123,0681944	(KANU 392053)	KP008329	KP008185	KP008257
divaricatum (Lindl.)						
Hoover				1/10000000	1/2000100	1/0000050
(DC) Kunth yar	JJ444_1	United States, California, Point Reyes National Seashore 37 995 –123 023	Jernstedt 369 (UCD 113455)	KP008330	KP008186	KP008258
divaricatum (Lindl.)		Scashore, 57,555, 125,525	119199)			
Hoover						
(DC) Kupth yar minus	FCsn_a	United States, California, Near Sunnyside,	Callahan s.n. (–)	KP008331	KP008187	KP008259
Hoover		55.65755, -122.57522				
Chlorogalum pomeridianum	AHA_1GH	United States, California, Ahart Ranch,	Ahart 19670 (WILLU	KP008332	KP008188	KP008260
(DC.) Kunth var.		39.346833, -121.487055	50112)			
Chlorogalum pomeridianum	MF5972 1	United States, Oregon, 8 Dollar Mountain,	Fishbein 5972 (HPSU)	KP008333	KP008189	KP008261
(DC.) Kunth var.		42.2315, -123.651				
pomeridianum		United States, California, Dadding, 40 CC1 400	Noloco 2012 000 (M/ULU	1/10000224	VD000100	1/100000000
(DC) Kunth yar	JKN_3	United States, California, Redding, 40.661488, -122 3084	Nelson 2013-009 (WILLU 50099)	KP008334	KP008190	KP008262
pomeridianum		122.500 1	50055)			
Chlorogalum purpureum	SBBG2169_a	United States, California, Near Jolon, 35.965,	Wilken 15701 (SBBG 2169)	KP008335	KP008191	KP008263
Brandegee var. purpureum	DWsn 1	–121.161 United States California Red Hill Road	Hannon 898 (SBBC	KP008336	KP008192	KP008264
Brandegee var. <i>reductum</i>	DTTSH_1	35.40293, -120.27959	114288)	10000000	10 000152	1000201
Hoover						
Chlorophytum alismifolium Baker	GB			-	-	AY191163
Furcraea cahum Trel.	GB			-	-	DQ071898
Hastingsia alba (Durand) S.	BFV_1M	United States, California, Butterfly Valley,	Kephart 661 (WILLU	KP008337	KP008193	KP008265
Watson Hastingsia alba (Durand) S	(2012) DBT 17M	40.01185, –120.99238 United States California Darlingtonia	50100) Barosh and Theiss sin	KDUU8338	KP008194	KP008266
Watson	(2012)	Botanical Trail, 41.850778, –123.907972	(WILLU 50101)	KI 000550	KI 000134	KI 000200
Hastingsia atropurpurea	WF_1 (2010)	United States, Oregon, Woodcock Fen,	Kephart 640 (WILLU	KP008339	KP008195	KP008267
Becking Hastingsia bracteosa S	DCC 19M	42.12755, -123.697 United States, Oregon, Deer Creek Center	50092) Morse I181b (_)	KD008340	KD008196	KD008268
Watson	(2012)	42.277183, -123.648136	Morse 11810 (-)	KI 000540	KI 000150	KI 008208
Hastingsia bracteosa S.	HF_1 (2010)	United States, Oregon, Howell's Fen,	Kephart 635 (WILLU	KP008341	KP008197	KP008269
Watson Hastingsia sorpontinicola	PP 2 (2010)	42.233263, -123.659023	50093) Kophart 620 (WILLU	1/1000212	1/0009109	KD008270
Becking	KK_3 (2010)	42.095183, -123.686733	50095)	111000342	NE000130	AF000270
Hastingsia serpentinicola	WF_17M	United States, Oregon, Woodcock Fen,	Theiss s.n. (WILLU 50116)	KP008343	KP008199	KP008271
Becking Herroria salsanarilha Mart	(2012)	42.12922, -123.69105				AV101170
Hesperaloe campanulata G.D.	GSAllende_1	Mexico, Coahuila de Zaragoza, Near Allende.		– KP008344	– KP008200	KP008272
Starr	_	28.30, -100.91				
Hesperaloe campanulata G.D.	GSType1_1	Mexico, Nuevo León, Mamulique Microondas, 26,5167 –100,125	Starr 93-001 (ARIZ 319675)	KP008345	KP008201	KP008273
Jian		20.3107, -100.123				

#### Table 2 (continued) -

Taxon	Individual	Locality	Voucher	ITS	trnL–F	ndhF
Hesperaloe campanulata G.D.	JH24658_1	Mexico, Nuevo León, Near Cerralvo, 26.0166,	Henrickson and	KP008346	KP008202	KP008274
Starr Hesperaloe campanulata G.D. Starr	MMsn	–99.0333 Plant grown at New York Botanical Garden	ratterson 24658 (TEX)	MonAToL	MonAToL	MonAToL
Hesperaloe engelmannii Krauskopf	JHsn_1	United States, Texas, Plant grown in Federicksburg, Texas; originally from W.		KP008347	KP008203	KP008275
Hesperaloe engelmannii Krauskonf	PSsn_3	branch of Nueces river, 29, –100 United States, Texas, Dobbs Run Ranch, 29,66672, –100,39512	Smith s.n. (KANU 392050)	KP008348	KP008204	KP008276
Hauskopi Hesperaloe funifera (K. Koch) Trel. ssp. chiangii G.D. Starr	GSMaiz_1	Mexico, San Luis Potos, Near Ciudad de Maíz, 22.451597, –99.671136		KP008349	KP008205	KP008277
Hesperaloe funifera (K. Koch) Trel. ssp. chiangii G.D. Starr	JH23741_1	Mexico, San Luis Potos, Near Pozos Santa Clara, 23.25, –100.55	Henrickson 23741 (TEX)	KP008350	KP008206	KP008278
Hesperaloe funifera (K. Koch) Trel. ssp. funifera	GB			U23978, U24038	-	DQ071899
Hesperaloe funifera (K. Koch) Trel. ssp. funifera	ARIZ319572_a	Mexico, Nuevo León, Sabinas Hidalgo, 26.319, -100.379	Starr 91-2 (ARIZ 319572)	KP008351	KP008207	KP008279
Hesperaloe nocturna Gentry	ARIZ406281_a	Mexico, Sonora, Cañada el Tejano, 30.57194, 	Reina-G., Van Devender, and Wolf 2010-217 (ARIZ 406281)	KP008354	KP008210	KP008282
Hesperaloe parviflora (Torr.) J.M. Coult. ssp. bechtoldii Hochstätter	JH24815_1	Plant grown in Sonora, Texas; original location unknown	Henrickson 24815 (TEX)	KP008355	KP008211	KP008283
Hesperaloe parviflora (Torr.) I.M. Coult. ssp. parviflora	GB			U23979, U24039	-	-
Hesperaloe tenuifolia G.D. Starr	GSType2_1	Mexico, Sonora, Cerro Agujudo, 27.11389, —108.72861	Meyer and Jenkins 90-63 (ARIZ 292741)	KP008358	KP008214	KP008286
Hesperocallis undulata A. Gray Hesperocallis undulata A. Gray	GB ARIZ376561_a	United States, Arizona, Barry M. Goldwater	McLaughlin and Bowers	– KP008359	AY561253 KP008215	AY225050 KP008287
Hesperocallis undulata A. Gray	MMsn	Range, 32.665, –113.1218333 United States, California, Desert Lily Reserve,	10283 (ARIZ 376561) Prince, Koontz, Pilapil, and	-	MonAToL	MonAToL
Hesperoyucca newberryi	DES62648_a	33.789, –115.283 United States, Arizona, 193 Mile Canyon,	Asanidze 674 (RSA 788238) Hodgson 21151 (DES	KP008352	KP008208	KP008280
(McKelvey) Clary Hesperoyucca newberryi	DES67431_a	36.00136194, -133.0038972 United States, Arizona, Mohawk Canyon,	62648) Hodgson, Makarick, Prince,	KP008353	KP008209	KP008281
(McKelvey) Clary	_	36.00379639, -112.0161333	Hahn, and Watters 21994 (DES 67431)			
Hesperoyucca peninsularis (McKelvey) Clary	SD145235_a	Mexico, Baja California, Near Catavina, 30.04916667, –115.5013889	Hodgson 9602 (SD 145235)	KP008356	KP008212	KP008284
Hesperoyucca peninsularis (McKelvey) Clary	SV1302261_1	Mexico, Baja California, Near Valle Tranquilo, 30.201833, -115.703833	Vanderplank, Riley, and Simancas 130226-1 (SD)	KP008357	KP008213	KP008285
Hesperoyucca whipplei (Torr.) Trel.	SD198048_a	United States, California, Jacumba, 32.63965, –116.20744	Hendrickson 3697 (SD 198048)	KP008360	KP008216	KP008288
Hesperoyucca whipplei (Torr.) Trel.	SD206369_a	United States, California, Cleveland National Forest, 32.95362, –116.7906	Rebman 19027 (SD 206369)	KP008361	KP008217	KP008289
Hesperoyucca whipplei (Torr.) Trel.	SV1303191_1	Mexico, Baja California, Rancho los aquajitos, 31.526192, –116.339312	Vanderplank and Arauz 1303191 (SD)	KP008362	KP008218	KP008290
Hosta ventricosa (Salisb.) Stearn	GB			U23980	AF508512	AF508401
Hosta ventricosa (Salisb.) Stearn	MM106	Plant grown in cultivation; original location unknown		-	MonAToL	MonAToL
Leucocrinum montanum Nutt. ex A. Grav	GB1			-	-	AY225052
Leucocrinum montanum Nutt. ex A. Grav	GB2			-	AF117003, AF117031	-
Leucocrinum montanum Nutt. ex A. Grav	LF_a	United States, Oregon, Lost Forest, 43.36411, –120.3293	Ruedas s.n. (OKLA)	KP008363	KP008219	KP008291
Manfreda virginica (L.) Salisb. ex Rose	GB	.20.5255		U23984, U24043	-	DQ071901
Polianthes geminiflora (Lex.)	GB1			U23989,	-	AY225048
Polianthes geminiflora (Lex.)	GB2			-	DQ500903,	-
Polianthes pringlei Rose	GB1			U23990, 1124048	-	DQ071902
Polianthes pringlei Rose	GB2			-	DQ500904, DO500938	-
Prochnyanthes mexicana	GB1			U23991, U24049	-	DQ071903
Prochnyanthes mexicana (Zucc.) Rose	GB2			-	DQ500917, DO500952	-
Schoenolirion albiflorum (Raf.) R.R. Gates	JS405_a	United States, Florida, Near Fellsmere, 27.7863611, –80.550556	Scanlon 405 (FLAS 208612)	KP008364	KP008220	KP008292

(continued on next page)

### Table 2 (continued)

Taxon	Individual	Locality	Voucher	ITS	trnL–F	ndhF
Schoenolirion croceum (Michx.) Alph. Wood	MM103	United States, Georgia, Rock & Shoals Outcrop, 33.887633, –83.334665	McKain 103 (WILLU 50104)	MonAToL	MonAToL	MonAToL
Schoenolirion wrightii Sherman	KPN_2.1GH	United States, Arkansas, Kingsland Prairie Natural Area, 33.86133, –92.24991	Zollner s.n. (WILLU 50106)	KP008365	KP008221	KP008293
Schoenolirion wrightii Sherman	WPN_1.4GH	United States, Arkansas, Warren Prairie Natural Area, 33.58208, —91.98305	Witsell 01-0109 (ANHC 3379)	KP008366	KP008222	KP008294
Yucca brevifolia Engelm.	MMsn		Smith s.n. (–)	MonAToL	MonAToL	MonAToL
Yucca queretaroensis Pina	GB1			-	-	JX903320
Yucca queretaroensis Pina	GB2			-	EU092616	-
Yucca treculeana Carrière	GB1			U23995, U24053	-	DQ071904
Yucca treculeana Carrière	GB2			-	EU092632	-

### Table 3

Primers used for amplification and sequencing. Those modified for this study are in bold.

DNA region	Primer name	Sequence	Source
ITS	N-nc18S10	AGGAGAAGTCGTAACAAG	Wen and Zimmer (1996)
ITS	C26A	GTTTCTTTTCCTCCGCT	Wen and Zimmer (1996)
trnL–F	c2	CTCAATGGAAGCTGTTCTAA	Modified from Taberlet et al. (1991)
trnL–F	f	ATTTGAACTGGTGACACGAG	Taberlet et al. (1991)
trnL–F	d	GGGGATAGAGGGACTTGAAC	Taberlet et al. (1991)
trnL–F	e	GGTTCAAGTCCCTCTATCCC	Taberlet et al. (1991)
ndhF	032F	TACCTTTTCTTCCACTTCCAGTT	Terry et al. (1997)
ndhF	451bF	TGGGAACTTGTAGGAATGT	Modified from Terry et al. (1997)
ndhF	451bR	ACATTCCTACAAGTTCCCA	Modified from Terry et al. (1997)
ndhF	745F	TGGTTACCTGATGCTATGGAAGG	Terry et al. (1997)
ndhF	745R	CCTTCCATAGCATCAGGTAACCA	Terry et al. (1997)
ndhF	1101bF	GGAACCTATTGTTGGRTATTCGCC	Modified from Terry et al. (1997)
ndhF	1101bR	GGCGAATAYCCAACAATAGGTTCC	Modified from Terry et al. (1997)
ndhF	1318bF	GGATTAACTGCATTTTATATGTTT	Modified from Terry et al. (1997)
ndhF	1318bR	AAACATATAAAATGCAGTTAATCC	Modified from Terry et al. (1997)
ndhF	1600bF	CCTCACGAGTCGGACAATACTATG	Modified from Terry et al. (1997)
ndhF	1600bR	CATAGTATTGTCCGACTCGTGAGG	Modified from Terry et al. (1997)
ndhF	2110R	CCCCCTATATATTTGATACCTTCTCC	Terry et al. (1997)

ity (Simmons et al., 2004). Alternate phylogenetic hypotheses were tested using approximately unbiased (AU) tests in Consel (Shimodaira, 2002), using the concatenated dataset. Candidate trees for AU tests were produced in Garli via unconstrained ML and bootstrap (BSt) analyses (500 replicates) as well as those constrained to include or exclude clades of interest (100–500 replicates). Conclusions were consistent across multiple runs of Consel with differing candidate trees from unconstrained and constrained ML and bootstrap analyses.

Nearly half of the ITS sequences displayed at least one stronglysupported double peak, suggesting incomplete concerted evolution for this region in Chlorogaloideae s.l. (Alvarez and Wendel, 2003). This may leave multiple ITS types within a given individual, while others may only have one type. Potts et al. (2014) provide the abbreviation 2ISP (pronounced "twisp") for such intra-individual site polymorphisms. While complicating the interpretation of ITS trees (Alvarez and Wendel, 2003), this does not eliminate their potential utility (Potts et al., 2014). Thorough cloning of ITS types and fluorescent or genomic in situ hybridization (FISH or GISH) would provide the best estimate of all ITS types in each individual and the genomic placement of those copies. However, such work is not feasible for many phylogenetic projects (Alvarez and Wendel, 2003), including the current study; and even those time and resource intensive methods do not guarantee complete sampling of all ITS types. Regardless, the impact of this intra-individual variation on phylogenetic analyses for Chlorogaloideae s.l. may be minimal.

Several types of analyses were run on the ITS data to account for these 2ISPs in alternate ways and assess their influence on phylogenetic inference with this data set. Indel characters were treated as missing for the following initial comparisons: (1) All nucleotide characters with a 2ISP in any individual were removed. As with any analysis based on regions within DNA arrays (such as ITS), this does not completely remove the potential for complex genetic processes to confound interpretation of phylogenies (Alvarez and Wendel, 2003). However, that would be true with or without obvious 2ISPs. (2) 2ISPs were coded as ambiguities and run using the typical settings described above. (3) Haplotypes were estimated for all 2ISPs using PHASE and kept as separate units in the phylogenetic analyses. PHASE assumes that no more than two alleles are present per individual, which is not necessarily true for ITS. However, it allows for some separation of haplotypes. GenBank sequences were excluded from phasing, as were populations Chlorogalum pomeridianum var. divaricatum AH and Hesperoyucca peninsularis SV1302261. The latter two were excluded due to excessive missing data in the center of their sequences. With those exclusions, there were no missing data positions. SeqPHASE (Flot, 2010) transformed file formats (Stephens and Donnelly, 2003; Stephens et al., 2001). PHASE 2.1.1 was run five times with different random seeds in order to check for consistency, and the final run included five internal replicates. (4) 2ISPs were coded as ambiguities but run using the step-matrix method of Potts et al. (2014). This assigns costs to changes between nucleotide codes, assuming that an ambiguity code states that all rather than any of the enclosed nucleotides are present at that site (i.e., R = A and G, not

#### Table 4

Characteristics of final DNA sequence datasets and resulting maximum parsimony (MP) and Bayesian inference (BI) analyses of Chlorogaloideae s.l. (i.e., the ingroup, IG) and outgroups (OG).

	Concatenated <sup>a</sup>	ITS-unphased	ITS-phased	cpDNA <sup>b</sup>
Dataset				
No. individuals	92 (76 IG, 16 OG)	87 (75 IG, 12 OG)	71 (35 with two types; 70 IG, 1 OG)	95 (76 IG, 19 OG)
No. characters	3715 (3682 nuc., 33 indel)	781 (756 nuc., 25 indel)	765 (749 nuc., 16 indel)	2943 (2926 nuc., 17 indel)
No. potentially parsimony informative characters	425 (286 among IG)	153 (122 among IG)	133 (132 among IG)	273 (166 among IG)
MP analyses				
Length of most parsimonious trees (MPTs)	844	401	253	438
No. of MPTs	6	96	216	24
CI	0.681	0.633	0.644	0.731
RI	0.925	0.917	0.956	0.943
RC	0.63	0.581	0.616	0.689
BI analyses				
No. generations	25,198,000	103,678,000	95,659,000	695,000
Potential scale reduction factors	1.000-1.001	1.000	1.000	1.00-1.006
Total effective sample sizes for each parameter	>12,230	>100,567	>95,459	>466

<sup>a</sup> Concatenated refers to the concatenated analyses of the unphased ITS dataset, *ndhF*, and *trnL–F*.

<sup>b</sup> cpDNA refers to the concatenated analyses of *ndhF* and *trnL-F*.

A or G). Table 5 represents the step matrix used, as inferred from Fig. 1 of Potts et al. (2014) when indel characters are not included.

## 3. Results

### 3.1. Differences due to optimality criterion or DNA region

Parsimony and BI analyses result in no strongly-supported differences based on the standard criteria defined in our methods (Fig. 3 and Tables 4 and 6). The best-fit models were GTR + G for ITS, TPM2uf + I + G for *trnL*–*F*, and TVM + I + G for *ndhF*. Preliminary analyses treated *ndhF* and *trnL*–*F* as separate loci. As expected given the lack of independence of DNA regions within the chloroplast genome, resulting topologies were entirely congruent; the regions were concatenated for subsequent analyses. Results from ITS and cpDNA analyses were generally congruent; most differences were not strongly supported and concatenation usually resulted in higher resolution and support for relationships. The few strongly-supported exceptions with apparent incongruence between loci are discussed below.

### 3.2. Influence of method of analysis for ITS data

It is unlikely that our ITS data set includes pseudogenes given the presence of conserved sections of DNA without any substitutions. Atypically long branches are absent in the ingroup, although the most distantly-related OG sampled for ITS, Leucocrinum montanum, is unsurprisingly divergent. The maximum number of 2ISP positions seen in any given individual is just nine (found in one individual, Schoenolirion wrightii KPN), whereas there are over 150 potentially parsimony informative characters in ITS (Table 4). We do not have information on double peaks for GenBank sequences; of the individuals sequenced by JKA's research group, 53% have no 2ISPs. 33% have 1-2 2ISPs. and only 14% have 3-9 2ISPs. Results from all four types of coding for 2ISPs in ITS analyses are similar. Analyses 1 (removing 2ISP characters), 2 (coding 2ISPs as ambiguities), and 4 (coding 2ISPs as ambiguities with a Potts step matrix) gave nearly identical results, with increasingly higher resolution and support in a few clades. Analysis 3 (2ISP characters phased into two haplotypes) unsurprisingly gave the most divergent results, but again relationships were generally consistent. All runs in PHASE gave similar results and so haplotypes were used from the final run. There were no uncertain phase calls or genotypes and confidence probabilities for all phase calls and all haplotypes were 1.00. Further analyses and discussion are focused on comparing analysis types 4 (Potts coding) and 3 (phased) for MP, and 2 (ambiguity coding) and 3 (phased) for BI; final concatenated analyses used the Potts method for MP and ambiguity coding for BI.

Phasing of 2ISPs into separate haplotypes produced only a few differences in topology (not shown) compared to unphased analyses (Fig. 3A and C), particularly in relationships with strong clade support. In most cases, multiple haplotypes from a given individual group together in the phased analyses either as a separate clade or as members of the same polytomy. When this is not the case, it does not result in major differences in interpretation of that individual's relationship to others. Universally strongly-supported cases of non-monophyly of individuals include intermixing of haplotypes from two populations of the same species (i.e., *C. howellii* DC and SM, and *S. wrightii* KPN and WPN) and two cases where one haplotype groups with other taxa. *Hesperoyucca peninsularis* SD145\_haplotypeB forms a clade with individuals of *H. whipplei* (PP = 1, JK = 86), while haplotypeA is sister to that clade. Similarly,

Table 5
Step matrix assigning the number of steps required to transition between each
potential nucleotide character coded using IUPAC codes, under the assumptions of the
2ISP coding method of Potts et al. (2014).

	Α	С	G	Т	М	R	W	S	Y	Κ	V	Н	D	В	Ν
Α	0	2	2	2	1	1	1	3	3	3	2	2	2	4	3
С	2	0	2	2	1	3	3	1	1	3	2	2	4	2	3
G	2	2	0	2	3	1	3	1	3	1	2	4	2	2	3
Т	2	2	2	0	3	3	1	3	1	1	4	2	2	2	3
Μ	1	1	3	3	0	2	2	2	2	4	1	1	3	3	2
R	1	3	1	3	2	0	2	2	4	2	3	3	3	3	4
W	1	3	3	1	2	2	0	3	2	2	3	1	1	3	2
S	3	1	1	3	2	2	3	0	2	2	1	3	3	1	2
Y	3	1	3	1	2	4	2	2	0	2	3	3	3	3	4
K	3	3	1	1	4	2	2	2	2	0	3	3	1	1	2
V	2	2	2	4	1	3	3	1	3	3	0	2	2	2	1
Н	2	2	4	2	1	3	1	3	3	3	2	0	2	2	1
D	2	4	2	2	3	3	1	3	3	1	2	2	0	2	1
В	4	2	2	2	3	3	3	1	3	1	2	2	2	0	1
Ν	3	3	3	3	2	4	2	2	4	2	1	1	1	1	0



**Fig. 3.** Inferred phylogenies from analyses of ITS and two cpDNA regions (*ndhF* and *trnL–F*). (A) Strict consensus topology of most parsimonious trees for concatenated analyses. Membership in Chlorogaloideae s.s. is indicated by black bars to the right of taxon names. Support values are given for concatenated/ITS/cpDNA analyses, MP jackknife values greater than 50 are shown above the branches and BI posterior probabilities are shown below. "–" indicates a clade that is not resolved in that analysis; "x" indicates a clade that is contradicted by that analysis (see B and C for alternate resolutions); "A" indicates a clade that is not relevant to that analysis due to unsampled accessions for some loci. Accessions that are missing DNA regions have the sequenced regions noted after the accession name; if two DNA regions were missing, the accession was not included in concatenated analyses. Dotted lines indicate relationships inferred by BI and/or separate locus analyses that differ from the concatenated strict consensus. More complex alternate resolutions of relationships are given in two separate topologies. (B) cpDNA MP/BI topology with focus on outgroup relationships. (C) ITS MP/BI topology for the *Camassia* clade. Genera of Chlorogaloideae s.l. are color coded, as are sub-groups in *Hesperaloe*. Site codes are listed after taxon names.

*Camassia quamash* ssp. *intermedia* PS\_haplotypeA joins the remainder of the *C. q. azurea*+ clade (PP = 1, JK = 83), excepting haplotypeB, which is sister to that clade. This may be due to deep coalescence, simply underscoring the close relationship among those subspecies given the low resolution within the *C. q. azurea*+ clade (e.g., Fig. 3). However, *C. q. intermedia* PS is sympatric with *C. leichtlinii* spp. *leichtlinii* PS, and the presence of individuals with intermediate traits such as floral color suggests gene flow. The DNA sequence of the partially "misplaced" *C. q. intermedia* PS\_haplotypeB may have been influenced by introgression, although it remains close to the *C. q. azurea*+ clade in the phylogeny. Further discussion of ITS results will focus on the unphased analyses, unless noted otherwise.

## 3.3. Phylogenetic hypotheses

Topologies are generally well resolved (Fig. 3), although some relationships require more data. Most inferences from ITS and cpDNA data are consistent. More OG sequences were available for cpDNA; we sampled five genera of the "Extended Agavaceae" clade of Bogler et al. (2006). As in that and other Family/Order-focused studies, our results strongly support this clade in the cpDNA trees (PP = 1, JK = 100; designated the "*Leucocrinum*+ clade"; Fig. 3B). Concatenated analyses also resolve this clade (PP = 1, JK = 100; Fig. 3A), with lower taxon sampling, while ITS analyses only included *Leucocrinum montanum*. Strongly inferred subclades within this group are completely consistent with Bogler et al. (2006).

The other OGs are divided among three clades: the *Hosta* clade, the *Yucca* clade, and the *Agave*+ clade (Fig. 3). The latter includes up to six of the sampled genera and is sister to the *Yucca* clade in concatenated (PP = 1, JK = 63) and cpDNA analyses (PP = 1, JK = 94), while those two clades are in a polytomy in the ITS BI tree, and *Yucca* is weakly placed as sister to a Chlorogaloideae s.l. – *Agave*+ clade (JK = 64) in the ITS MP tree. High-quality ITS sequences were not available for *Hosta*; cpDNA and concatenated analyses do not strongly resolve the placement of this genus.

Within the *Agave*+ clade, concatenated relationships are consistent with Bogler et al. (2006), although they are better supported or resolved in some cases. Relationships from ITS and cpDNA analyses are largely consistent, and the only two exceptions lack strong support in some analyses. Concatenated and ITS data resolve a *Polianthes – Prochnyanthes* clade (concat.: PP = 0.99, JK = 78; ITS: PP = 0.98, JK = 65; Fig. 3A), while *Polianthes geminiflora* is excluded from this clade in cpDNA analyses (PP = 0.97, JK = 62; Fig. 3B). Also, in concatenated and ITS analyses, the sampled species of *Agave* fall in a well-supported grade near the base of the *Agave*+ clade. In the cpDNA topologies, two of the *Agave* species are nested more deeply (PP = 0.97, JK = 62).

Chlorogaloideae s.s. is not inferred as monophyletic by any analysis nor supported by AU tests (Table 6). A separate SHH clade sensu Halpin and Fishbein (2013) is strongly supported (concat.: PP = 1, JK = 97; ITS: PP = 0.92, JK = 88; cpDNA: PP = 1, JK = 88), although almost 11% of the trees constrained to lack the SHH clade are accepted by AU tests. Core Chlorogaloideae has strong support (concat.: PP = 1, JK = 99; cpDNA: PP = 1, JK = 100), but Hesperocallis is nested within the clade by some ITS analyses (Core -Hesperocallis clade: PP = 0.99, JK = 92). The AU tests do not distinguish between trees in which the Core clade is monophyletic vs. has Hesperocallis nested, but regardless, AU tests accept fewer than 13% of trees lacking a Core - Hesperocallis clade. Although ITS sequences were only available for one population of Hesperocallis undulata, three populations were included in our final analyses of concatenated and cpDNA data and are strongly supported as monophyletic by all analyses, including AU tests. Separate analyses of the cpDNA data do not strongly resolve the relationship of *Hesperocallis* relative to Chlorogaloideae; it is placed in a weak grade/polytomy with all other major clades. Analyses of ITS strongly resolve *H. undulata* with the Core Chlorogaloideae (Core – *H. undulata*: PP = 0.99, JK = 92). *Hesperocallis* is placed sister to the *Ch. parviflorum* – *Ch. purpureum* clade by BI analyses (PP = 0.70) and sister to all of Core Chlorogaloideae by MP analyses (JK = 54). Concatenated analyses strongly resolve it as sister to the Core Chlorogaloideae clade (Core – *H. undulata*: PP = 1.0, JK = 88; Core: PP = 1.0, JK = 99). The AU tests do not significantly support or reject the monophyly of Chlorogaloideae s.l.; this group forms a polytomy or is weakly resolved either as monophyletic or not by phylogenetic analyses.

The monophyly of all tested genera in Chlorogaloideae s.l. is well supported by all phylogenies, with the exceptions of Hastingsia, Hesperaloe, and Chlorogalum. In addition, although the *Hesperovucca* clade has moderate to strong support in the phylogenies, it is not completely supported by AU tests. However, only 6% of trees constrained to lack this clade were accepted by the tests. The monophyly of Hastingsia and Hesperaloe is strongly supported by concatenated analyses, but Hastingsia forms a polytomy in ITS BI analyses (vs. ITS MP: [K = 71; cpDNA: PP = 1, [K = 100) and Hesperaloe is a polytomyin cpDNA analyses (vs. ITS: PP = 0.99, JK = 90). Our AU tests rejected all but 6% of trees lacking a Hastingsia clade and did not significantly support the monophyly of Hesperaloe. Chlorogalum forms two (to three) very strongly-supported clades. These form a grade at the base of the Camassia - Hastingsia clade, with the branch dividing Chlorogalum being strongly inferred by concatenated (PP = 1, JK = 98, AU p < 0.001) and ITS (PP = 1, JK = 78) analyses, and more weakly by cpDNA (PP = 0.8, JK = 66).

Within the SHH clade, strong support exists for a sister relationship between *Hesperaloe* and *Schoenolirion* (concat.: PP = 1, JK = 100; ITS: PP = 1, JK = 99; cpDNA: PP = 1, JK = 99). Relationships among populations of *Schoenolirion* are fully resolved and well supported, with the strongest signal from cpDNA data. In *Hesperoyucca*, two major clades are strongly supported from concatenated analyses and strongly to weakly inferred by individual DNA regions. Our preliminary analyses included three other accessions of *H. whipplei* and one of *H. newberryi* that were later excluded due to missing data, but inferences within *Hesperoyucca* were entirely consistent with the final results.

Hesperaloe funifera and its subspecies are each not monophyletic. Specifically, H. funifera ssp. funifera groups with H. campanu*lata* (concat.: PP = 1, JK = 81; JTS: PP = 0.99, JK = 58; cpDNA: PP = 0.98, JK = 62; Fig. 3) while H. f. chiangii groups with H. nocturna and *H. tenuifolia* (concat.: PP = 1, JK = 94; ITS: PP = 1, JK = 97; cpDNA: polytomy). All analyses place population H. f. funifera ARIZ319572 in a subclade of H. campanulata. The remaining accession of H. f. funifera and H. campanulata are sister to this subclade according to cpDNA; they form a polytomy with this and other Hesperaloe populations in the ITS trees. Hesperaloe parviflora s.l. is monophyletic in concatenated (PP = 1, JK = 99) and cpDNA trees (PP = 1, JK = 98), but largely unresolved in ITS. The single population of H. p. bechtoldii nests within the two populations of H. engelmannii based on concatenated and cpDNA analyses (concat.: PP = 0.99, JK = 63; ITS: polytomy; cpDNA: PP = 0.97, JK = 64); this lack of monophyly for H. engelmannii was significant according to AU tests (Table 6). Only ITS sequence data were available for H. p. parviflora; this accession grouped with H. f. funifera GB (PP = 1, |K = 100).

*Camassia* and *Hastingsia* are strongly placed as sister taxa (concat.: PP = 1, JK = 99; ITS: PP = 0.97, JK = 71; cpDNA: PP = 1, JK = 100; Fig. 3). The great majority of trees without a *Camassia* – *Hastingsia* clade were rejected by AU tests, but 3% of trees with a short branch linking *Hastingsia* and *Chlorogalum* were marginally accepted. Relationships within *Hastingsia* are largely

unresolved, whereas some clades of *Camassia* have strong support. One contrast between the two genomes involves the *C. cusickii* – *C. quamash* ssp. *utahensis* CM clade, which groups with the *C. q. breviflora*+ clade according to ITS (PP = 1, JK = 86), but with the *C. q. azurea*+ clade according to cpDNA (PP = 1, JK = 84). AU tests show no significant difference between the altering placements. Another difference, but one that is not

strongly supported by MP analyses, is the placement of the *C. angusta* – *scilloides* clade. ITS data place it sister to the rest of *Camassia*, excluding *C. howellii* (PP = 0.92, JK = 56), but cpDNA data resolve it in a polytomy with the *C. q. breviflora*+ clade (PP = 1, JK = 79). Concatenated data also group it with the *C. q. breviflora*+ clade, but AU tests do not reject the alternate resolution.

#### Table 6

-

Clade support values (MP Jackknife/BI Posterior Probability/ML Bootstrap) and maximum AU *p*-values for trees that contain ("Pos") or contradict ("Neg") a given clade. Trees are significantly less likely than best trees (i.e., rejected) if AU *p*-values are <0.05 (shown in bold). "-" indicates a clade that is not resolved in that analysis; "x" indicates a clade that is contradicted by that analysis; "^" indicates a clade that is not relevant to that analysis due to unsampled accessions for some loci.

Clade	Concat. JK/PP/BSt	ITS JK/PP	cpDNA JK/PP	AU p-values
Intergeneric				
Chlorogaloideae s.l.ª	x (SHH – OGs:<50)/0.93/–	-/-	x/x (Hesperocallis – OGs: 60/0.69)	Neg: 0.699, Pos: 0.999
Chlorogaloideae s.l. excluding <i>Hesperocallis</i> (vs. Core – <i>Hesperocallis</i> , see below)	x/x/x	x/x	<50/0.57	Pos: 0.184
Chlorogaloideae s.s. <sup>b</sup>	x/x/x	x/x	x/x	Pos: 0.018
Core Chlorogaloideae – Hesperocallis	88/1.00/96	92/0.99	x/x (Hesperocallis – OGs: 60/0.69)	Neg: 0.168
Core Chlorogaloideae <sup>c</sup>	99/1.00/99	54/x (Hesperocallis nested)	100/1.00	Neg: 0.662
SHH <sup>d</sup>	97/1.00/97	88/0.92	88/1.00	Neg: 0.422
Camassia – Hastingsia	99/1.00/99	71/0.97	100/1.00	Neg: 0.066
Generic				
Camassia	100/1.00/100	100/1.00	100/1.00	Neg: 0.011
Chlorogalum (vs. Chlorogalum grade)	x/x/x (vs. 98/1.00/ 96)	x/x (vs. 78/1.00)	x/x (vs. 66/0.80)	Pos:<0.001
Hastingsia	100/1.00/100	71/-	100/1.00	Neg: 0.193
Hesperaloe	95/0.96/53	90/0.99	-/-	Neg: 0.985
Hesperocallis	100/1.00/100	<pre>^/^ (one Hesperocallis sampled)</pre>	100/1.00	Neg: 0.047
Hesperoyucca	99/1.00/99	78/0.85	94/1.00	Neg: 0.246
Schoenolirion	100/1.00/100	92/1.00	100/1.00	Neg: 0.003
Intrageneric				
Camassia excluding C. angusta, C. scilloides, and C. howellii (vs. excluding C. leichtlinii and C. howellii)	x/x/x (vs. 80/1.00/ 84)	56/0.92	x/x (vs. 83/1.00)	Pos: 0.911
C. cusickii – C. q. breviflora+ (vs. C. cusickii – C. q. azurea+)	x/x/x (vs. 58/0.99/ 66)	86/1.00	x/x (vs. 84/1.00)	Pos: 0.995
Camassia cusickii (vs. C. cusickii HC or PC – C. q. utahensis CM)	x/x/x (vs. 87/1.00/ 91)	x/x (vs. 63/0.98)	x/x (vs. 95/1.00)	Pos:<0.001
Hesperaloe engelmannii (vs. H. engelmannii JHsn – H. p. bechtoldii JH24815)	x/x/x (vs. 63/0.99/ 61)	-/-	x/x (vs. 64/0.97)	Pos:<0.001
Hesperaloe funifera (vs. H. f. funifera ARIZ319572 – H. campanulata)	x/x/x (vs. 81/1.00/ 94)	x/x (vs. 58/0.99)	x/x (vs. 62/0.98)	Pos:<0.001
Schoenolirion croceum – S. wrightii (vs. S. albiflorum – S. wrightii)	x/x/x (vs. 99/1.00/ 99)	x/x (vs. 67/0.81)	x/x (vs. 98/1.00)	Pos:<0.001

<sup>a</sup> Chlorogaloideae s.l. = Camassia, Chlorogalum, Hastingsia, Schoenolirion, Hesperaloe, Hesperoyucca, Hesperocallis.

<sup>b</sup> Chlorogaloideae s.s. = Camassia, Chlorogalum, Hastingsia, Schoenolirion.

<sup>c</sup> Core Chlorogaloideae = Camassia, Chlorogalum, Hastingsia.

<sup>d</sup> SHH = Schoenolirion, Hesperaloe, Hesperoyucca.

#### Table 7

A subset of morphological traits discussed by Sherman (1969), as applied to Chlorogaloideae s.l. Sources are Sherman (1969) and/or FNA (1993+) unless noted otherwise.

	Rootstock	Stigma shape	Leaf shape
Camassia	Bulb	3-lobed	Keeled
Chlorogalum	Bulb	3-lobed	Keeled
Hastingsia	Bulb	3-lobed	Keeled
Schoenolirion	Rhizome with or without bulb	Unlobed to slightly 3-lobed	Flat or rounded, not prominently keeled
Hesperaloe	Rhizome	Unlobed <sup>a</sup>	Curved cross-section <sup>b</sup>
Hesperoyucca	With or without rhizomes <sup>c</sup>	Unlobed <sup>d</sup> or slightly 3-lobed <sup>e</sup>	Plano-convex, subtriquetrous, or keeled on both faces <sup>f</sup>
Hesperocallis	Bulb	Unlobed to slightly 3-lobed	Keeled

<sup>a</sup> Watson and Eaton (1871).

<sup>b</sup> Hochstätter (2009), JKA, pers. obs.

<sup>c</sup> FNA (1993+), W. Hodgson, pers. comm.

<sup>d</sup> FNA (1993+), Clary (2001), Trelease (1902).

<sup>e</sup> Watson and Eaton (1871), Engelmann (1873).

f Trelease (1902).

## 4.1. Overview and concordance among analyses

Phylogenetic analyses of independent nuclear and chloroplast loci resolved many relationships in the ecologically diverse Chlorogaloideae s.l. (Fig. 3), allowing insights into taxonomic questions at the familial to infraspecific levels. Results from MP, ML, and BI analyses were generally concordant, as were those from different gene regions. Certainly caution is important when using ITS for phylogenetic inference (Alvarez and Wendel, 2003; Edger et al., 2014), yet results in this case indicate consistent phylogenetic signal reflecting evolutionary relationships. Multiple coding methodologies produced similar phylogenetic hypotheses, and results from ITS in general are consistent with patterns evident from cpDNA regions. The coding method of Potts et al. (2014) did appear useful in retaining more information from 2ISPs (intra-individual site polymorphisms) in that the resulting topologies were slightly better resolved and supported. Even so, most clades and support values were nearly identical regardless of the coding method. Although polyploidy may also complicate patterns of evolution for molecular characters, there is no indication here that the ancient polyploidy in Agavaceae (McKain et al., 2012) nor the chromosomal changes within Chlorogaloideae s.l. (Halpin and Fishbein, 2013) confounded inference of phylogenetic relationships.

# 4.2. Relationships among outgroups and placement of Chlorogaloideae s.l. within Agavaceae

This study provides the most extensive sampling to date of taxa in Chlorogaloideae s.l., and we assess their placement within Agavaceae using a diverse array of outgroups across the major clades of the family. All strongly-supported results are consistent with previous studies focused at the family or order level (e.g., Bogler et al., 2006; Chen et al., 2013; Seberg et al., 2012). The single study with results partially inconsistent with those reported here focused on Ruscaceae/Nolinoideae (Kim et al., 2010), sampling just two species of Chlorogaloideae s.l. (in *Camassia* and *Hesperocallis*) and 16 total members of Agavaceae. *Camassia* was strongly supported as sister to the remainder of Agavaceae except *Anemarrhena*, a result not upheld here or by other phylogenetic studies (e.g., Halpin and Fishbein, 2013).

Our results confirm the phylogenetic affinity of *Yucca queretaroensis* with *Yucca* rather than with *Hesperoyucca* or *Hesperaloe*, which is consistent with its stable taxonomic placement within that genus. This provides the test suggested by Pellmyr et al. (2007), verifying its placement with greater phylogenetic sampling of these other genera.

## 4.3. Monophyly of Chlorogaloideae s.s. and s.l

Chlorogaloideae s.s. (Speta, 1998) is not inferred as monophyletic by any of our analyses (Fig. 3 and Table 6). Instead, *Schoenolirion* is separated from Core Chlorogaloideae (*Camassia*, *Chlorogalum*, *Hastingsia*) and placed in a strongly-supported "SHH" clade with *Hesperaloe* and *Hesperoyucca*. Halpin and Fishbein (2013) shared these results, supporting the monophyly of Core Chlorogaloideae based on SH tests (Shimodaira and Hasegawa, 1999) of four cpDNA regions. Core Chlorogaloideae is strongly affirmed by our concatenated and cpDNA analyses, and also by ITS with the inclusion of *Hesperocallis* weakly nested or sister to it (see below). The SHH clade is strongly supported by all phylogenetic analyses but without complete support from AU tests, similar to results of Halpin and Fishbein (2013). Previous phylogenetic taxon sampling in the SHH clade had been sparse, with *Schoenolirion* included only by Halpin and Fishbein (2013), and minimal sampling from *Hesperaloe* and *Hesperoyucca* even in that study. *Schoenolirion* is strongly inferred as sister to *Hesperaloe* and the recognition of *Hesperoyucca* as a separate genus from *Yucca* is upheld by new phylogenetic evidence (see also FNA, 1993+).

Schoenolirion has been classified with some genera of Chlorogaloideae since Watson (1879; Table 1), but its separation from Core Chlorogaloideae is not entirely unexpected based on morphology. Sherman (1969) noted only a "weak affinity" between Schoenolirion and its purported cohorts of the time (Chlorogalum, Hastingsia, and *Hemiphylacus*). He hypothesized a closer relationship with Camassia and compared the morphology of the genera in Chlorogaloideae s.s., suggesting that Schoenolirion was less closely allied than the other three but without mentioning Hesperaloe or Hesperovucca as possible relatives. Schoenolirion is unique in Chlorogaloideae s.s. for its unusual, vertical rhizome (vs. bulbs), entire stigma (vs. clearly three-lobed), and leaf shape that is not prominently keeled (Table 7; FNA, 1993+; Halpin and Fishbein, 2013; Sherman, 1969). In contrast, the three genera of Core Chlorogaloideae and Hesperocallis have bulbs (Fig. 1), while Hesperaloe, Hesperoyucca (excepting H. newberryi, Wendy Hodgson, pers. comm.), and much of the rest of Agavaceae are rhizomatous (Halpin and Fishbein, 2013; Stevens, 2001+). Also, all three genera in the SHH clade have largely unlobed stigmas (Clary, 2001; Engelmann, 1873; FNA, 1993+; Sherman, 1969; Trelease, 1902; Watson and Eaton, 1871). Stigmas of Hesperoyucca seen by JKA (herbarium specimens) and W. Hodgson (pers. comm.) appear capitate with trichomes (Fig. 1), although minute lobing is possible and may be the source of some mixed reports (Table 7). Hesperocallis stigmas are capitate to slightly three-lobed (FNA, 1993+). Finally, keeled leaves characterize Core Chlorogaloideae and Hesperocallis, while leaves from the SHH clade tend to lack keels (FNA, 1993+; Hochstätter, 2009; Sherman, 1969; Trelease, 1902; JKA, pers. obs.). In all, these morphological characters are consistent with a Core Chlorogaloideae clade (possibly joined by Hesperocallis) and a SHH clade (Table 7).

Although strongly placed in Agavaceae (Bogler et al., 2006; Pires et al., 2004), the specific affinities of monotypic *Hesperocallis* remained poorly supported. Our results from concatenated analyses now strongly place *H. undulata* sister to Core Chlorogaloideae, but AU tests are not conclusive regarding this resolution (Table 6). The exact position of *Hesperocallis* varies somewhat with DNA region, but these differences are not strongly supported and may be due to differing OG sampling. Halpin and Fishbein (2013) resolved a similar pattern of relationships with their analyses of four cpDNA regions, but with lower support. While the monophyly of Chlorogaloideae s.l. is not clearly supported or rejected by our data, it is clear that Chlorogaloideae s.s. is not supported by molecular phylogenetic data. Delimitation of alternate taxonomic subfamilies or other such groups would benefit from a global analysis focused at the family level.

# 4.4. Monophyly of genera of Chlorogaloideae s.l., with focus on and within Chlorogalum

The monophyly of each genus in Chlorogaloideae s.l. is strongly affirmed by all phylogenetic trees and AU tests, with the following exceptions. Despite strong support from concatenated MP and BI analyses, AU tests do not significantly support the monophyly of *Hastingsia*, *Hesperaloe*, or *Hesperoyucca*. The former two genera also form polytomies in at least one of the separate analyses, although the monophyly of *Hastingsia* is supported by the cpDNA phylogenies of Halpin (2011) and Halpin and Fishbein (2013). The only genus consistently supported as non-monophyletic is *Chlorogalum*  (Fig. 3), whose two (or three) strongly-supported clades form a grade at the base of the *Camassia – Hastingsia* clade. The main clades are *Ch. pomeridianum – angustifolium – grandiflorum* and *Ch. parviflorum – purpureum*; the species of the latter clade form a polytomy in the cpDNA tree. Halpin and Fishbein (2013) similarly resolved three groups of *Chlorogalum* using cpDNA data, although SH tests suggested that a monophyletic *Chlorogalum* was not statistically less likely (Halpin and Fishbein, 2013), in contrast to our AU test results.

Molecular phylogenetic data thus back the division of Chlorogalum into at least two genera, an assertion that also has morphological and cytological support. Both Ch. parviflorum and Ch. purpureum differ from other species of Chlorogalum based on smaller diurnal (vs. vespertine) flowers and styles longer than the perianth (Hoover, 1940). However, Ch. pomeridianum and Ch. grandiflorum can develop exserted styles (FNA, 1993+). Regardless, Hoover (1940) argued that "These differences are so correlated with some difference in general aspect that one is led to suspect that the genus as accepted is composed of two separate lines of descent." He nevertheless retained all five species in one genus because they were "so much alike morphologically and so close geographically" (p. 140). Chromosome number also distinguishes the diurnal species (n = 30) from the vespertine Ch. po. pomeridianum (n = 18 or 15), Ch. po. divaricatum and Ch. po. minus (n = 18), and Ch. angustifolium (n = 17; Cave, 1970). Chromosomes of Ch. grandiflorum have not been counted to our knowledge. Cave (1970) states that the *n* = 15 karyotype within *Chlorogalum* is similar to that of *Camassia* (n = 15), consistent with a close phylogenetic relationship and similar floral characters (FNA, 1993+). Overall, while molecular, morphological, and chromosomal data support division of Chlorogalum, a closer look with detailed field or morphological study is recommended to determine the classification that best reflects evolutionary relationships.

At least four taxa in *Chlorogalum* s.l. are of conservation concern due to restricted distributions (Ch. grandiflorum, Ch. po. minus, Ch. pu. purpureum, and Ch. pu. reductum). Each tested species in the genus is strongly supported as monophyletic by concatenated analyses: only one accession was available for *Ch. grandiflorum*. Hoover (1940) also noted that each species is easily distinguished, possibly in part due to allopatry of species pairs except Ch. pomeridianum and Ch. angustifolium. He detected no hybridization between these two species, although Cave (1970) observed cytological disturbances (e.g., irregular meiosis) in individuals of both species at one of several sites of sympatry. Our analyses revealed a closer relationship between Ch. pomeridianum var. pomeridianum and Ch. po. divaricatum compared to Ch. po. minus. This is consistent with morphology; Ch. po. minus is distinctive in lacking numerous coarse bulb fibers and is relatively small, although its short stature may reflect growth on serpentine soils (Hoover, 1964). Sometimes confused with Ch. grandiflorum (FNA, 1993+), Ch. po. minus is supported as a member of Ch. pomeridianum by concatenated analyses.

# 4.5. Intrageneric relationships within the SHH clade, with focus on taxonomic questions in Hesperaloe

Of three recognized species in *Hesperoyucca*, the populations of *H. newberryi* are resolved as sister to a *H. whipplei* – *H. peninsularis* clade. This is consistent with geography: *H. newberryi* is disjunct in Arizona, whereas the distributions of *H. whipplei* (California and Baja California) and *H. peninsularis* (Baja California) overlap (Fig. 2). The ITS phylogeny of Clary (2001) separates *H. whipplei* (2 populations, BSt = 73) and *H. peninsularis* (1 pop.), whereas our five populations of these two species form a polytomy. Population genetic studies may further resolve species boundaries for this genus.

In Schoenolirion, Sherman (1969) recognized three species, suggesting that S. croceum is the progenitor of derivative S. wrightii. Instead, our analyses strongly support a S. albiflorum - wrightii clade. Halpin and Fishbein (2013) inferred a S. croceum clade sister to a clade containing one population of each of the three species. Still, all these results might be consistent with Sherman's (1969) hypothesis given that S. albiflorum is polyploid (n = 24?, 2n = 49), compared to S. wrightii (n = 12) and S. croceum (n = 15 or rarely)16; Sherman, 1969). As such, a bifurcating tree may not completely describe its relationships. Sherman (1969) proposed that S. albiflorum was an allopolyploid between a hypothetical now-extinct species and S. croceum (or its ancestor), but that remains to be tested. The current distribution of *S. albiflorum* partially overlaps with *S.* croceum but not with S. wrightii. However, the only two morphological traits given by Sherman (1969) to distinguish S. croceum and S. wrightii link the latter with S. albiflorum (white vs. vellow flowers and leaves shorter than scape vs. longer). If instead S. wrightii is a diploid parent of S. albiflorum, it would explain their close relationship in our phylogenies. Despite the small size of this genus, chromosomal evolution appears complex.

Taxonomic revisions of *Hesperaloe* provide testable hypotheses (Hochstätter, 2009; Starr, 1997; Turner and Turner, 2002). Currently, the main areas of disagreement involve whether *H. funifera* ssp. *chiangii* represents a distinct species (*H. chiangii*), whether to segregate *H. engelmannii* or treat it as a synonym of *H. parviflora*, and potential recognition of new taxa, *H. parviflora* ssp. *bechtoldii* (Hochstätter, 2009) and *H. malacophylla* (Hochstätter and Martínez-Ávalos, 2010).

The monophyly of *H. funifera* and each of its subspecies is not supported by our data (Fig. 3 and Table 6). Populations of H. f. funifera are intermixed in a clade with H. campanulata (including the type population: H. campanulata GSType1). The distributions of the two taxa overlap in northern Mexico (Fig. 2), where Starr (1997, pers. comm.) noted an unusual, potentially hybridizing population of H. f. funifera characterized by intermediate floral coloration and plant sizes. Our H. f. funifera ARIZ319572 is less than 20 km from that site and has flowers of similar color (Greg Starr. pers. comm.). Starr also produced hybrids between H. f. funifera and *H. campanulata*, as well as  $F_1$  and  $F_2$  hybrids of *H. f. funifera* and *H. parviflora*. In morphology, the latter hybrids resembled *H*. campanulata, raising the hypothesis of a hybrid origin for that species (G. Starr, pers. comm.). Placing hybrids in phylogenetic reconstructions is not straightforward (McDade, 1992), and population genetic or morphological work is needed. Regardless, either hybridization or deep coalescence may explain the phylogenetic intermixing of H. campanulata and H. f. funifera, suggesting incomplete genetic segregation of these taxa.

Our results separate *H. chiangii* from *H. funifera*. The ITS and concatenated analyses strongly support a clade with two accessions of *H. f. chiangii* (including JH23741 from the type locality) and a *H. nocturna – H. tenuifolia* subclade. Since the *H. campanulata – H. f. funifera* clade is strongly inferred by cpDNA and concatenated analyses, all analyses divide *H. funifera*. However, given the complex relationships in this group, further study of genetic and morphological patterns in multiple populations is needed prior to making taxonomic decisions. The distribution of *H. f. chiangii* is disjunct from *H. f. funifera* but also from *H. nocturna* and *H. tenuifolia* (Hochstätter, 2009; Fig. 2).

Hesperaloe parviflora s.l. includes H. p. ssp. parviflora, H. p. ssp. bechtoldii, and putative segregate H. engelmannii. The population of H. p. parviflora (represented by a single available ITS sequence) groups strongly with population H. f. funifera GB. Our population of H. p. bechtoldii is nested within two populations of H. engelmannii based on concatenated and cpDNA analyses with moderate to high support (Fig. 3). While this raises questions about the potential recognition of H. engelmannii as a species, not all sister species

are reciprocally monophyletic (e.g., Luckow, 1995). In fact, all progenitor-derivative species pairs are expected to display a paraphyletic phylogenetic pattern, at least initially (Crawford, 2010; Rieseberg and Brouillet, 1994).

Turner and Turner (2002) noted that most taxonomists had not recognized H. engelmannii, although Trelease (1902) treated it as a variety of H. parviflora. Hochstätter (2009) followed Turner and Turner (2002) in accepting H. engelmannii, in contrast to G. Starr (pers. comm.) and James Henrickson (pers. comm.), who have also worked extensively with this genus (e.g., Starr, 1997). Hesperaloe engelmannii was originally distinguished from H. parviflora based on a shorter, thicker style and longer anthers (Fig. 1; Baker, 1880; Krauskopf, 1878; Starr, 1997). Turner and Turner (2002) confirmed floral differences, stating that H. engelmannii styles are "mostly 1-2(3) times as long as the ovary" (p. 41) in contrast to 3-5 times for *H. parviflora*. They further stated that individuals of H. engelmannii are larger, with longer and darker leaves, but also noted significant variation in both putative species. This is corroborated by a detailed study revealing carpel lengths that varied as much as 44% in one population of H. parviflora s.s. (Pellmyr and Augenstein, 1997; identified as H. p. bechtoldii, J. Henrickson, pers. comm.). James Henrickson (pers. comm.) also noted extensive variation in H. parviflora s.l. and discovered that individuals transplanted from a highly exposed population of *H. parviflora* s.s. to a shaded habitat eventually produced much longer leaves, in the range expected for H. engelmannii. Our work confirms the close relationship between H. engelmannii and H. p. bechtoldii, but further studies are needed. In particular, the separate ITS placement of H. p. parviflora s.s. should be tested with other populations and loci.

Natural or artificial hybrids have been observed between multiple species of *Hesperaloe*, including *H. nocturna* and *H. parviflora*, *H. parviflora* and *H. funifera*, *H. campanulata* and *H. funifera*, *H. campanulata* and *H. parviflora* and *H. parviflora* x *H. funifera*, and *H. parviflora* and *H. eng-elmannii* (Hochstätter, 2009; Starr, 1997). One factor that may provide partial reproductive isolation among some species is flowering time. This genus includes day flowering species (*H. parviflora* s.l.), night flowering species (*H. funifera*, *H. nocturna*, and *H. tenuifolia*), as well as *H. campanulata*, whose flowers are visited by bats and hawkmoths at night but partially close into floral tubes during the day that are visited by hummingbirds (Starr, 1997; J. Henrickson, pers. comm; G. Starr, pers. comm.).

## 4.6. Phylogenetic framework for Camassia and Hastingsia

The sister relationship of *Camassia* and *Hastingsia* is strongly supported in all phylogenies. Species delimitation within *Hastingsia* has been contentious in prior and current taxonomies (Becking, 1986; FNA, 1993+; Lang and Zika, 1997; Theiss, in press). None of our analyses strongly support intrageneric resolution. This suggests a rapid radiation, also supported by lack of resolution and intermixing of species in a chloroplast (*rpl16*, *trnD–T*, *psbJ–petA*) phylogeny focused on *Hastingsia* (Halpin, 2011).

In contrast, major clades in *Camassia* are resolved. For a cpDNA phylogeny of *Camassia*, Fishbein et al. (2010) analyzed *rpl16* and *trnD–trnT*. Thirty-seven of our sampled populations overlap with theirs, and we additionally conducted a 5-region concatenated analysis with those accessions. Results (not shown) are entirely concordant with those in Fig. 3; differences are simply changes in support values, which almost always increased. In Fishbein et al. (2010), *C. leichtlinii* and *C. howellii* formed a very weakly supported clade (the "L clade"; PP = 0.51, BSt < 50) sister to the remainder of *Camassia* (the "Q clade"; PP = 1, BSt = 72). Both clades are contradicted in our ITS trees because *C. angusta – scilloides* is sister to the rest of the Q clade and *C. leichtlinii* (with weak to moderate support, Fig. 3C). Our concatenated and cpDNA results confirm a strongly-supported Q clade, but infer that *C. leichtlinii* and

*C. howellii* form a strongly-supported grade rather than the L clade (Fig. 3A). Consistent in all of our analyses, *C. howellii* is sister to the rest of *Camassia*. Both *C. leichtlinii* and *C. howellii* share the rare presence of branches in their racemes, a trait otherwise absent in *Camassia*, but found in *Chlorogalum* and *Hastingsia* (Gould, 1942).

Subspecies of C. quamash occurring west and east of the Cascade Mountains fall in separate clades (i.e., the C. q. azurea+ and C. q. breviflora+ clades, respectively; Fig. 3), consistent with prior genetic and phylogenetic studies (Fishbein et al., 2010; Tomimatsu et al., 2009). Camassia cusickii has a very narrow range in northeastern Oregon and adjacent Idaho, overlapping with the eastern group of C. quamash (C. q. breviflora, C. q. quamash, and C. q. utahensis). The same two populations of C. cusickii as in Fishbein et al. (2010) were again not monophyletic, a result now supported by AU tests and both nuclear and chloroplast data. although the placement of C. cusickii differs in the two trees. ITS data strongly place C. cusickii with other taxa distributed east of the Cascades (Fig. 3C), whereas our and Fishbein et al. (2010)'s cpDNA analyses clearly group it with the western subspecies of C. quamash, along with geographically-proximate population(s) of C. q. utahensis that remain sister to a population of C. cusickii (Fig. 3A). Recent fieldwork and specimen vouchers imply that C. q. utahensis and C. cusickii may grow within dispersal range, with potential for intermixing.

Midwestern *C. angusta* and *C. scilloides* form a strongly supported clade in our concatenated and ITS trees, as in the cpDNA tree of Fishbein et al. (2010). Our cpDNA data also inferred a nested *C. angusta* OTT – *scilloides* BAR subclade, disrupting the monophyly of both species. Although hybridization could cause this pattern, it is unlikely for these populations, with one occurring in Indiana and the other in Kansas. Regardless, close genetic ties characterize these species.

Overall, our results are consistent with many aspects of prior taxonomic and phylogenetic studies of *Camassia* (Gould, 1942; Fishbein et al., 2010; FNA, 1993+; Kephart, in press), despite the complexity of some taxonomic boundaries. Many questions remain regarding diversification within this clade. Detailed phylogenetic study of *Camassia* and *Hastingsia* is underway and further discussion of intrageneric relationships are reserved for separate papers. Combining those results with morphological, ecological, and reproductive isolation studies will allow a more comprehensive understanding of the patterns and processes of speciation in the *Camassia – Hastingsia* clade.

## 5. Conclusion

Within Chlorogaloideae s.l., our results demonstrate lack of monophyly at the subfamily, genus, species, and infraspecific levels – while suggesting stability of the Core Chlorogaloideae and SHH clades as well as many of their genera. Complex patterns of evolution for many of these groups were revealed by these data. A detailed investigation of forces behind diversification in Chlorogaloideae s.l. is outside the scope of this study, but it appears that a variety of factors may have been involved, including allopatry, hybrid speciation, divergence in floral traits and pollinator interactions, and chromosomal changes. While it is difficult to say whether some of the differences emerged before or after speciation of the relevant taxa (Templeton, 1982), regardless they may be important in maintaining isolation and allowing further divergence.

## Acknowledgments

This research was supported primarily by NSF DEB 1145847 Grant awarded to SRK, JKA, and TMC, as well as by NSF DBI

1262795 REU site Grant awarded to Jennifer Gleason and Mark Mort, and by the Biodiversity Institute and the Department of Ecology and Evolutionary Biology at the University of Kansas (KU). The Earthwatch Institute, M.J. Murdock Trust, Native Plant Society of Oregon, and Oregon Community Foundation also provided Grant support to SRK.

We thank Ryan Beck, Sidrah Sheikh, and Apollonia Shreders for assistance with lab work at KU; Theresa Barosh, Francis Cartieri, Jon King, Theresa Melhem, Justin Sharkan, Sidrah Sheikh, and Sunita Yadav for assistance with fieldwork in the Midwest US; Geraldine Allen, Mike Mesler, Julie Morgan, Andy Schnabel, Paul Wolf, many Willamette University students, and Earthwatch volunteers for field and lab assistance nationwide to SRK; Gordon Barnhardt, Roger Boyd, Craig Freeman, Jacalyn Goetz, Dean Kettle, Caleb Morse, Ken O'Dell, and J. Sleichter, for assistance in locating and accessing field sites in Kansas; the Bureau of Land Management, the Nature Conservancy, the U.S. Forest Service, and other agencies and private landowners for permission and logistical support in the northwest US; Craig Freeman and Caleb Morse (KANU) and Richard Halse (OSC) for herbarium support; Katie Sadler for assistance in greenhouse cultivation of plants at KU; Mark Holder for consultation on analytical methods; Jim Kephart for production of the map figure; and Daniel Crawford, Mark Mort, and two anonymous reviewers for helpful comments on an earlier version of the manuscript.

Inclusion of taxa outside of the Camassia – Hastingsia clade was greatly facilitated by the generous assistance of many workers and herbaria. Specimen loans were graciously made available from the following herbaria, with permission to sample for DNA extraction: ARIZ (via Shelley McMahon and Sarah Hunkins), DAV (via Ellen Dean), DES (via Wendy Hodgson), UC (via Andrew Doran), and SD (via Jon Rebman and Layla Aerne). The following aided in locating sources of plants: Pablo Carrillo-Reyes, Amanda Neill, Jackie Poole, Tiana Rehman, Tom Wendt, and Theo Witsell. Each of the following provided at least one individual included in our study: Lowell Ahart, Mark Fishbein (including DNAs obtained via herbaria: UCR and SBBG), James Henrickson, Ann Howald, Judy Jernstedt. Julie Nelson, Paula Smith, Greg Starr, Sula Vanderplank, and Dieter Wilken. Mark Fishbein additionally shared DNA aliquots from a number of members of Chlorogaloideae. Genomic data for a small subset of taxa was provided by Karolina Heyduk and Michael McKain. Personal expertise on Hesperaloe and Hesperoyucca was made available by James Henrickson, Wendy Hodgson, and Greg Starr and is greatly appreciated.

#### References

- Alvarez, I., Wendel, J.F., 2003. Ribosomal ITS sequences and plant phylogenetic inference. Mol. Phylogenet. Evol. 29, 417-434.
- APG (Angiosperm Phylogeny Group), 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. Bot. I. Linn. Soc. 141. 399-436.
- APG (Angiosperm Phylogeny Group), 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. Bot. J. Linn. Soc. 161, 105-121.
- Baker, J.G., 1873. Revision of the genera and species of Scilleae and Chlorogaleae. Bot. J. Linn. Soc. 13, 209-292.
- Baker, J.G., 1880. A synopsis of Aloineae and Yuccoideae. Bot. J. Linn. Soc. 18, 148-263.
- Baldwin, B.G., Sanderson, M.J., Porter, J.M., Wojciechowski, M.F., Campbell, C.S., Donoghue, M.J., 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. Ann. Mo. Bot. Gard. 82. 247-277.
- Batsch, A.I.G.C., 1786. Dispositio Generum Plantarum Jenensium Secundum Linnaeum et Familias Naturales. Jenae.
- Becking, R.W., 1986. Hastingsia atropurpurea (Liliaceae: Asphodeleae), a new species from southwestern Oregon. Madroño 33, 175-181.
- Bogler, D.J., Pires, J.C., Francisco-Ortega, J., 2006. Phylogeny of Agavaceae based on ndhF, rbcL, and ITS sequences: implications of molecular data for classification. Aliso 22, 313-328.
- Cave, M.S., 1970. Chromosomes of the Liliaceae. Univ. Calif. Publ. Bot. 57, 1-58.

- Chase, M.W., Reveal, J.L., Fay, M.F., 2009. A subfamilial classification for the expanded asparagalean families Amaryllidaceae, Asparagaceae and Xanthorrhoeaceae. Bot. J. Linn. Soc. 161, 132–136.
- Chen, S.C., Kim, D.K., Chase, M.W., Kim, J.H., 2013. Networks in a large-scale phylogenetic analysis: reconstructing evolutionary history of Asparagales (Lilianae) based on four plastid genes. PLoS ONE 8, 1–18.
- Clary, K.H., 2001. The genus Hesperoyucca (Agavaceae) in the western United States and Mexico: new nomenclatural combinations. Sida, Contribut. Bot. 19, 839-847
- Crawford, D.J., 2010. Progenitor-derivative species pairs and plant speciation. Taxon 59, 1413-1423.
- Cronquist, A., 1981. An Integrated System of Classification of Flowering Plants. Columbia University Press, New York.
- Cupov, V., 1994. Filogenija i sistema porjadkov Liliales i Asparagales. Bot. Zhurn. (Moscow & Leningrad) 79, 1-12.
- Dahlgren, R., Clifford, H.T., Yeo, P., 1985. The Families of the Monocotyledons. Springer Verlag, New York.
- Doyle, J.J., Doyle, J.L., 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bullet. 19, 11-15.
- Edger, P.P., Tang, M., Bird, K.A., Mayfield, D.R., Conant, G., et al., 2014. Secondary structure analyses of the nuclear rRNA internal transcribed spacers and assessment of its phylogenetic utility across the Brassicaceae (Mustards). PLoS ONE 9, e101341. http://dx.doi.org/10.1371/journal.pone.0101341.
- Engelmann, G., 1873. Notes on the genus Yucca. Trans. Acad. Sci. St. Louis 3, 17–54. Engler, A., 1887. Liliaceae. In: Engler, A., Prantl, K. (Eds.), Die Natürlichen Pflanzenfamilien. Wilhelm Engelmann, Leipzig, Germany, pp. 10–91.
- Felsenstein, J., 1978. Cases in which parsimony or compatibility methods will be positively misleading. Syst. Zool. 27, 401-410.
- Fishbein, M., Kephart, S.R., Wilder, M., Halpin, K.M., Datwyler, S.L., 2010. Phylogeny of Camassia (Agavaceae) inferred from plastid rpl16 intron and trnD-trnY-trnEtrnT intergenic spacer DNA sequences: implications for species delimitation. Syst. Bot. 35, 77-85.
- Flot, J.F., 2010. SeqPHASE: a web tool for interconverting PHASE input/output files and FASTA sequence alignments. Mol. Ecol. Resour. 10, 162-166.
- FNA (Flora of North America Editorial Committee) (Eds.), 1993+. Flora of North America North of Mexico. 16+ vols. Oxford University Press, New York, New York and Oxford, United Kingdom.
- Funk, V.A., Susanna, A., Stuessy, T.F., Robinson, H., 2009. Classification of compositae. In: Funk, V.A., Susanna, A., Stuessy, T.F., Bayer, R.J. (Eds.), Systematics, Evolution, and Biogeography of Compositae. International Association for Plant Taxonomy, Institute of Botany, University of Vienna, Vienna, Austria, pp. 171-189.
- Good-Avila, S.V., Souza, V., Gaut, B.S., Eguiarte, L.E., 2006. Timing and rate of speciation in Agave (Agavaceae). Proc. Natl. Acad. Sci. USA 103, 9124-9129.
- Gould, F.W., 1942. A systematic treatment of the genus Camassia Lindl. Am. Midl. Nat. 28, 712-742.
- Haines, L., 1941. Variation in Yucca whipplei. Madroño 6, 33-45.
- Halpin, K.M., 2011. A chloroplast phylogeny of Agavaceae subfamily Chlorogaloideae with a focus on species relationships in Hastingsia. Botany, Oklahoma State University, Stillwater, Oklahoma.
- Halpin, K.M., Fishbein, M., 2013. A chloroplast phylogeny of Agavaceae subfamily Chlorogaloideae: implications for the tempo of evolution on serpentine soils. Syst. Bot. 38, 996–1011.
- Hochstätter, F., 2009. The genus Hesperaloe (Agavaceae). Cactus World 27, 97-106.
- Hochstätter, F., Martínez-Ávalos, J.G., 2010. Una nuova specie di Hesperaloë (Agavaceae): Hesperaloë malacophylla Hochstätter & J. G. Martínez-Ávalos spec. nov. Piante Grasse 30, 20-22.
- Hoogland, R.D., Reveal, J.L., 2005. Index nominum familiarum plantarum vascularium. Bot. Rev. 71, 1–291.
- Hoover, R.F., 1940. A monograph of the genus *Chlorogalum*. Madroño 5, 137–147. Hoover, R.F., 1964. Some noteworthy Californian Liliaceae. Leaflets West. Bot. 10, 121-128
- Kellogg, E.A., Appels, R., Mason-Gamer, R.J., 1996. When genes tell different stories: the diploid genera of Triticeae (Gramineae). Syst. Bot. 21, 321-347
- Kephart, S.R., in press. Camassia. In: Meyers, S.C., Cook, T.J., Mitchel, K.E., Hardison, L.K. (Eds.), Flora of Oregon, Volume 1. Botanical Research Institute of Texas Press, Fort Worth, Texas.
- Kim, J.H., Kim, D.K., Forest, F., Fay, M.F., Chase, M.W., 2010. Molecular phylogenetics of Ruscaceae sensu lato and related families (Asparagales) based on plastid and nuclear DNA sequences. Ann. Bot. 106, 775-790.
- Krause, K., 1930. Liliaceae. In: Engler, A., Prantl, K. (Eds.), Die Natürlichen Pflanzenfamilien. Wilhelm Engelmann, Leipzig, Germany, pp. 227–386.
- Krauskopf, E., 1878. Notice to Botanists, Botanical Gardens and Nurserymen. Circular.
- Lang, F.A., Zika, P.F., 1997. A nomenclatural note on Hastingsia bracteosa and Hastingsia atropurpurea (Liliaceae). Madroño 44, 189-192.
- Luckow, M., 1995. Species concepts: assumptions, methods, and applications. Syst. Bot. 20, 589-605.
- Maddison, D.R., 1991. The discovery and importance of multiple islands of mostparsimonious trees. Syst. Zool. 40, 315-328.
- Mason-Gamer, R.J., Kellogg, E.A., 1996. Testing for phylogenetic conflict among molecular data sets in the tribe Triticeae (Gramineae). Syst. Biol. 45, 524-545.
- McDade, L.A., 1992. Hybrids and phylogenetic systematics II. The impact of hybrids on cladistic analysis. Evolution 46, 1329-1346.
- McGary, J. (Ed.), 2001. Bulbs of North America. Timber Press and North American Rock Garden Society, Portland, Oregon.

- McKain, M.R., Wickett, N., Zhang, Y., Ayyampalayam, S., McCombie, W.R., Chase, M.W., Pires, J.C., dePamphilis, C.W., Leebens-Mack, J., 2012. Phylogenomic analysis of transcriptome data elucidates co-occurrence of a paleopolyploid event and the origin of bimodal karyotypes in Agavoideae (Asparagaceae). Am. I. Bot. 99. 397-406.
- McKelvey, S.D., 1938. Yuccas of the Southwestern United States, Part One. Arnold Arboretum of Harvard University, Jamaica Plain, Massachusetts.
- McKelvey, S.D., 1947. Yuccas of the Southwestern United States, Part Two. Arnold Arboretum of Harvard University, Jamaica Plain, Massachusetts.
- McKinney, K.K., Hickman, J.C., 1993. Yucca. In: Hickman, J.C. (Ed.), The Jepson Manual: Higher Plants of California. University of California Press, Berkeley, California.
- Mehalchick, G., Boyd, D.K., Kibler, K.W., Ringstaff, C.W., 2004. Shifting Sands and Geophytes: Geoarcheological Investigations at Paluxy sites on Fort Hood, Texas. United States Army Fort Hood Archeological Resource Management Series Research Report No. 48. Directorate of Public Works, Environmental Management Office, Fort Hood, Texas.
- Moerman, D.E., 1986. Medicinal Plants of Native America. University of Michigan Press, Ann Arbor, Michigan.
- Müller, K., 2005. SeqState primer design and sequence statistics for phylogenetic DNA data sets. Appl. Bioinform. 4, 65-69.
- Nylander, J.A.A., Wilgenbusch, J.C., Warren, D.L., Swofford, D.L., 2008. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. Bioinformatics 24, 581-583.
- Pellmyr, O., 2003. Yuccas, yucca moths, and coevolution: a review. Ann. Mo. Bot. Gard. 90. 35-55.
- Pellmyr, O., Augenstein, E.J., 1997. Pollination biology of Hesperaloe parviflora. Southw. Natural. 42, 182–187.
- Pellmyr, O., Segraves, K.A., Althoff, D.M., Balcazar-Lara, M., Leebens-Mack, J., 2007. The phylogeny of yuccas. Mol. Phylogenet. Evol. 43, 493-501.
- Pfosser, M., Speta, F., 1999. Phylogenetics of Hyacinthaceae based on plastid DNA sequences. Ann. Mo. Bot. Gard. 86, 852-875.
- Pires, J.C., Maureira, I.J., Rebman, J.P., Salazar, G.A., Cabrera, L.I., Fay, M.F., Chase, M.W., 2004. Molecular data confirm the phylogenetic placement of the enigmatic Hesperocallis (Hesperocallidaceae) with Agave. Madroño 51, 307-311.
- Posada, D., 2008. jModelTest: phylogenetic model averaging. Mol. Biol. Evol. 25, 1253-1256.
- Potts, A.J., Hedderson, T.A., Grimm, G.W., 2014. Constructing phylogenies in the presence of intra-individual site polymorphisms (2ISPs) with a focus on the nuclear ribosomal cistron. Syst. Biol. 63, 1-16.
- Rambaut, A., Drummond, A.J., 2007. Tracer. <a href="http://beast.bio.ed.ac.uk/Tracer">http://beast.bio.ed.ac.uk/Tracer</a>>.
- Rieseberg, L.H., Brouillet, L., 1994. Are many plant species paraphyletic? Taxon 43, 21-32.
- Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542.
- Sanchez, R., Rodriguez, A., Navarro, E., Requejo, A., Jimenez, L., 2011. Integrated utilization of the main components of Hesperaloe funifera. Biochem. Eng. J. 56, 130-136.
- Sato, D., 1935, Analysis of the karvotypes in Yucca, Agave and the related genera with special reference to the phylogenetic significance. Japanese J. Genet, 11, 272-278.
- Schulze, W., 1982. Beitrage zur taxonomie der Liliifloren. IX. Anthericaceae. Wiss. Z. Friedrich-Schiller-Univ. Jena, Math.-Naturwiss. Reihe 31, 291-307.
- Seberg, O., Petersen, G., Davis, J.I., Pires, J.C., Stevenson, D.W., Chase, M.W., Fay, M.F., Devey, D.S., Jorgensen, T., Sytsma, K.J., Pillon, Y., 2012. Phylogeny of the Asparagales based on three plastid and two mitochondrial genes, Am. J. Bot, 99. 875-889.
- Sherman, H.L., 1969. A systematic study of the genus Schoenolirion (Liliaceae). General Biology. Vanderbilt University, Nashville, Tennessee.
- Shimodaira, H., 2002. An approximately unbiased test of phylogenetic tree selection. Syst. Biol. 51, 492-508.
- Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16, 1114-1116.

- Siddall, M.E., 1998. Success of parsimony in the four-taxon case: long-branch repulsion by likelihood in the Farris zone. Cladistics 14, 209-220.
- Simmons, M.P., Ochoterena, H., 2000. Gaps as characters in sequence-based phylogenetic analyses. Syst. Biol. 49, 369-381.
- Simmons, M.P., Pickett, K.M., Miya, M., 2004. How meaningful are Bayesian support values? Mol. Biol. Evol. 21, 188-199.
- Smith, C.I., Pellmyr, O., Althoff, D.M., Balcazar-Lara, M., Leebens-Mack, J., Segraves, K.A., 2008. Pattern and timing of diversification in Yucca (Agavaceae): specialized pollination does not escalate rates of diversification. Proc. R. Soc. B - Biol. Sci. 275, 249-258.
- Soltis, D.E., Soltis, P.S., 1998. Choosing an approach and an appropriate gene for phylogenetic analysis. In: Soltis, D.E., Soltis, P.S., Doyle, J.J. (Eds.), Molecular Systematics of Plants II: DNA Sequencing, Kluwer Academic Publishers, Boston, Massachusetts, pp. 1-42.
- Speta, F., 1998. Systematische analyse der gattung Scilla L. s.l. (Hyacinthaceae). Phyton-Ann. REI Bot. 38, 1-141.
- Starr, G., 1997. A revision of the genus Hesperaloe (Agavaceae). Madroño 44, 282-296
- Stephens, M., Donnelly, P., 2003. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. Am. J. Hum. Genet. 73, 1162-1169
- Stephens, M., Smith, N.J., Donnelly, P., 2001. A new statistical method for haplotype reconstruction from population data. Am. J. Hum. Genet. 68, 978-989
- Stevens, P.F., 2001+. Angiosperm Phylogeny Website. Version 12, July 2012 [and More or Less Continuously Updated Since]. <http://www.mobot.org/MOBOT/ research/APweb/
- Swofford, D.L., 2003. PAUP\*. Phylogenetic Analysis Using Parsimony (\*and Other Methods). Version 4. Sinauer, Sunderland, Massachusetts.
- Taberlet, P., Gielly, L., Pautou, G., Bouvet, J., 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. 17.1105-1109.
- Templeton, A.R., 1982. Genetic architectures of speciation. In: Barigozzi, C. (Ed.), Mechanisms of Speciation. Alan R. Liss, Inc., New York, New York, pp. 105-121.
- Terry, R.G., Brown, G.K., Olmstead, R.G., 1997. Examination of subfamilial phylogeny in Bromeliaceae using comparative sequencing of the plastid locus ndhF. Am. J. Bot. 84, 664-670.
- Theiss, K.E., in press. Hastingsia. In: Meyers, S.C., Cook, T.J., Mitchel, K.E., Hardison, L.K. (Eds.), Flora of Oregon, Volume 1. Botanical Research Institute of Texas Press, Fort Worth, Texas.
- Tomimatsu, H., Kephart, S.R., Vellend, M., 2009. Phylogeography of Camassia quamash in western North America: postglacial colonization and transport by indigenous peoples. Mol. Ecol. 18, 3918-3928.
- Trelease, W., 1902. The Yucceae. Annual Rep. Missouri Bot. Gard. 13, 27-133.
- Tropicos.org, Missouri Botanical Garden. (Accessed 2.04.14). <http://www.tropicos. org>.
- Turner, B.L., Turner, M.W., 2002. Natural populations of Hesperaloe (Agavaceae) in Texas. Lundellia 5, 34-43.
- Watson, S., 1879. Contributions to American botany. I. Revision of the North American Liliaceae. II. Descriptions of some new species of North American plants. Proc. Am. Acad. Arts Sci. 14, 213-303.
- Watson, S., Eaton, D.C., 1871. Botany. G.P.O, Washington, DC. Wen, J., Zimmer, E.A., 1996. Phylogeny and biogeography of *Panax* L. (the ginseng genus, Araliaceae): inferences from ITS sequences of nuclear ribosomal DNA. Mol. Phylogenet. Evol. 6, 167–177.
- Wilgenbusch, J.C., Warren, D.L., Swofford, D.L., 2004. AWTY: A System for Graphical Exploration of MCMC Convergence in Bayesian Phylogenetic Inference. <a href="http://www.style.com">http://www.style.com</a> ceb.csit.fsu.edu/awtv>.
- Zizumbo-Villarreal, D., Vargas-Ponce, O., Rosales-Adame, J.J., Colunga-GarciaMarin, P., 2013. Sustainability of the traditional management of Agave genetic resources in the elaboration of mezcal and tequila spirits in western Mexico. Genet, Resour, Crop Evol. 60, 33-47.
- Zwickl, D.J., 2006. Genetic Algorithm Approaches for the Phylogenetic Analysis of Large Biological Sequence Datasets under the Maximum Likelihood Criterion. The University of Texas, Austin, Texas, <a href="http://www.bio.utexas.edu/faculty/">http://www.bio.utexas.edu/faculty/</a> antisense/garli/Garli.html>.