

PHYLOGENETIC, MORPHOLOGICAL, AND CHEMOTAXONOMIC INCONGRUENCE IN THE NORTH AMERICAN ENDEMIC GENUS *ECHINACEA*¹

LEX E. FLAGEL,² RYAN A. RAPP,² CORRINNE E. GROVER,² MARK P. WIDRLECHNER,³
JENNIFER HAWKINS,⁴ JESSIE L. GRAFENBERG,² INÉS ÁLVAREZ,⁵ GYU YOUNG CHUNG,⁶
AND JONATHAN F. WENDEL^{2,7}

²Department of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, Iowa 50011 USA; ³USDA-ARS North Central Regional Plant Introduction Station, Ames, Iowa 50011 USA; ⁴Department of Genetics, University of Georgia, Athens, Georgia 30602 USA; ⁵Departamento de Biodiversidad y Conservación, Real Jardín Botánico, CSIC, Madrid 28014 Spain; and ⁶School of Bioresource Science, Andong National University, Andong Gyeongbuk, 760-749 South Korea

The study of recently formed species is important because it can help us to better understand organismal divergence and the speciation process. However, these species often present difficult challenges in the field of molecular phylogenetics because the processes that drive molecular divergence can lag behind phenotypic divergence. In the current study we show that species of the recently diverged North American endemic genus of purple coneflower, *Echinacea*, have low levels of molecular divergence. Data from three nuclear loci and two plastid loci provide neither resolved topologies nor congruent hypotheses about species-level relationships. This lack of phylogenetic resolution is likely due to the combined effects of incomplete lineage sorting, hybridization, and backcrossing following secondary contact. The poor resolution provided by molecular markers contrasts previous studies that found well-resolved and taxonomically supported relationships from metabolic and morphological data. These results suggest that phenotypic canalization, resulting in identifiable morphological species, has occurred rapidly within *Echinacea*. Conversely, molecular signals have been distorted by gene flow and incomplete lineage sorting. Here we explore the impact of natural history on the genetic organization and phylogenetic relationships of *Echinacea*.

Key words: Asteraceae; chloroplast DNA; *Echinacea*; incomplete lineage sorting; phylogenetics; single-copy nuclear DNA.

Species of the genus *Echinacea* are geographically circumscribed within a region of North America that has undergone repeated rounds of glaciation (Clayton and Moran, 1982), with the last such round, the Wisconsinan, ending roughly 10 000 yr before the present. Presently, the genus ranges from southern Alberta, Canada to near the coast of the Gulf of Mexico in Texas and Louisiana and from the oak savannas of Ohio, the glades of Tennessee, and open habitats in the Carolinas west to the foothills of the Rocky Mountains (Urbatsch et al., 2006). Much of this range was under ice during the last glacial epoch, signifying that *Echinacea* survived in southerly refugia. Despite the expansive aggregate range of the genus, much of this range has been converted into agricultural production, resulting in an extremely fragmented modern population structure. This distributional history has many potential implications for the genetic architecture of a perennial plant species, most impor-

tantly, the disruption of natural processes of intraspecific and interspecific gene flow and the attendant increase in population fragmentation and genetic bottlenecks.

Taxonomically, *Echinacea* is delimited into nine species (Table 1), including two, *E. angustifolia* DC and *E. paradoxa* (Norton) Britton, that each are further divided into two varieties (McGregor, 1968; Flora of North America Editorial Committee, 1993+; McKeown, 1999). These species are all diploid with the exception of *E. pallida*, which is putatively a polyploid (Mechanda et al., 2004). This taxonomic treatment was devised by McGregor (1968), who spent 15 years studying the genus while making controlled, common-garden crosses, noting that many hybrids have high levels of stability, fecundity, and viability in parental backcrosses. In a recent morphological study, four species with eight subspecies were proposed (Binns et al., 2002, 2004), but McGregor's classification continues to be widely used by botanists and herbalists (see discussion in Blumenthal and Urbatsch [2006]) and serves as the basis for the recent *Flora of North America* treatment (Urbatsch et al., 2006).

McGregor's results regarding the ease of formation and fertility of interspecific hybrids suggest that *Echinacea* may either be a young genus in which rapid speciation has occurred (McKeown, 2004) or one in which, for reasons other than relative youth, genetic barriers have incompletely formed. In either case, gene flow between species has been historically common; McGregor noted hybrid swarms in natural sympatric settings, and in more recent molecular work, Mechanda et al. (2004) found evidence of natural hybrid individuals.

Assessment and maintenance of *Echinacea* genetic diversity is of interest due to the purported human health benefits from several *Echinacea* species (Speroni et al., 2002), as well as the cultivation and breeding of the plant as an ornamental

¹ Manuscript received 11 February 2008; revision accepted 7 April 2008.

This journal paper of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa, Project No. 1018, was supported by Hatch Act and State of Iowa funds and was made possible by grant number P01ES012020 from the National Institute of Environmental Health Sciences (NIEHS) and the Office of Dietary Supplements (ODS), NIH. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIEHS, NIH. Mention of commercial brand names does not constitute an endorsement of any product by the U. S. Department of Agriculture or cooperating agencies. J.-A. McCoy and other members of the Center for Research on Botanical Dietary Supplements at Iowa State University were helpful in providing feedback and comments throughout the research. L. Urbatsch kindly provided DNA samples for sequence analysis.

⁷ Author for correspondence (e-mail: jfw@iastate.edu)

TABLE 1. *Echinacea* taxa characterized, with U. S. state of origin and USDA Germplasm Resources Information Network Plant Introduction (PI) accession.^{a,b}

Taxon	State	PI accession
<i>E. angustifolia</i>	OK	PI 631267
<i>E. angustifolia</i> var. <i>angustifolia</i>	OK	PI 631272
	IA	PI 631285
	KS	PI 631318
<i>E. angustifolia</i> var. <i>strigosa</i>	OK	PI 631266, PI 631320
<i>E. atrorubens</i>	KS	PI 631255, PI 631299
	OK	PI 631260, PI 631262
<i>E. laevigata</i>	SC	PI 631310, PI 631312
	NC	PI 631314
	VA	PI 631316
<i>E. pallida</i>	OK	PI 631275
	IA	PI 631290
	AR	PI 631293
	MO	PI 631296
	NC	PI 631315
<i>E. paradoxa</i> var. <i>neglecta</i>	OK	PI 631263, PI 631264, PI 631265
	AR	PI 631292
<i>E. paradoxa</i> var. <i>paradoxa</i>	MO	PI 631301, PI 631321
	LA	PI 633669
<i>E. purpurea</i>	MO	PI 631307
	NC	PI 631313
	LA	PI 631257, PI 631258, PI 633672
<i>E. sanguinea</i>	LA	PI 631257, PI 631258, PI 633672
	LA	PI 631257, PI 631258, PI 633672
<i>E. simulata</i>	KY	PI 631249
	MO	PI 631304
	TN	PI 631308
<i>E. tennesseensis</i>	TN	PI 631250, PI 631324, PI 631325, PI 631326

Notes: AR, Arkansas; IA, Iowa; LA, Louisiana; MO, Missouri; NC, North Carolina; OK, Oklahoma; SC, South Carolina; TN, Tennessee; VA, Virginia

^a See <http://www.ars-grin.gov/npgs> to reference PI accession numbers.

^b More accession information on herbarium vouchers and GenBank sequences is in Appendix 1.

(Ault, 2006). The health-promoting properties of these plants have garnered much attention from herbalists, scientists, and consumers (Speroni et al., 2002; Kim et al., 2004; Turner et al., 2005; Schoop et al., 2006), and recent usage of *Echinacea* has increased largely due to its potential application as a modulator of the human immune system (Yu and Kaarlas, 2004). Demand for *Echinacea* has generated a small industry based on wild harvesting and processing (Price and Kindscher, 2007) [particularly of *E. angustifolia*, *E. pallida* (Nutt.) Nutt., and *E. purpurea* (L.) Moench]. Such wild harvesting, coupled with habitat loss, now threatens some remaining wild populations of *Echinacea* (McKeown, 1999), two of which are federally endangered, *E. laevigata* (C. L. Boynton & Beadle) S. F. Blake and *E. tennesseensis* (Beadle) Small (see <http://www.fws.gov/endangered>).

The horticultural and medicinal promise of *Echinacea* has prompted numerous studies of genetic variation, genetic structure, and hybridity within the genus, using a suite of molecular markers including amplified fragment length polymorphism (AFLP) (Baum et al., 2001; Kim et al., 2004; Mechanda et al., 2004) and randomly amplified polymorphic DNA (RAPD) (Kapteyn et al., 2002). When compared to one another, the results from these studies are incongruent and contain conflicting assessments of gene flow within and between species. An additional hurdle in integrating these results has been the small and disparate sampling strategies employed for each study.

In contrast, a recent study (L. Wu, Iowa State University; P. Dixon, B. Nikolau, G. Kraus, M. Widrechner, and E. Wurtele, unpublished manuscript) of 40 populations of *Echinacea*, selected to encompass a broad geographical and morphological diversity, examined metabolite profiles generated by HPLC, and reported that patterns of biochemical diversity corresponded well to taxonomic circumscriptions and relationships as conveyed in McGregor's (1968) monograph. In addition, a morphological study by Binns et al. (2002), although proposing an alternative treatment, used character data to produce a clustering pattern with node support reflecting McGregor's original treatment.

In this study, we sought to elucidate a phylogenetic framework for *Echinacea* by using both nuclear and plastid loci. Our goal was to describe genetic relationships among the nine congeners and reveal the parental origin of the polyploid species. The data, however, revealed a history of secondary contact and hybridization, mirroring the glaciation-entwined history and shedding light on some of the processes giving rise to conflicting molecular assessments of phylogenetic relationships.

MATERIALS AND METHODS

Plant material—We selected 38 accessions of *Echinacea* (Table 1, Appendix 1) representing the full geographic range of the *Echinacea* germplasm collection in the U. S. National Plant Germplasm System maintained by the USDA-ARS North Central Regional Plant Introduction Station (NCRPIS), Ames, Iowa (Widrechner and McKeown, 2002). These accessions span the extremes of *Echinacea*'s geographic distribution and include several accessions from areas where species exist in sympatry (Missouri, Kansas, and Oklahoma). Accessions were keyed to species during the initial regeneration process on the basis of McGregor (1968).

Seed samples of the 38 accessions were soaked for 24 h in a 1 mmol solution of ethephon to overcome dormancy and promote rapid germination (Sari et al., 2001). After soaking, the seeds were transferred to clear plastic boxes with blotter moistened with distilled water. The germination boxes were held at 4°C for 4 weeks and then transferred to germination chambers at a constant 25°C with 14 h of light per day. Three-week-old seedlings were transferred into 20-cm pots in a growing medium consisting of 50% Canadian peat moss, 40% perlite, and 10% mineral soil, and grown under ambient light in a greenhouse at 22–25°C, with daily watering and biweekly fertilizing.

Tissue preparation and DNA extraction—Greenhouse grown plants were keyed to species (McGregor, 1968) at sexual maturity to verify identities as received from the NCRPIS. Young leaves and flower buds were collected for DNA extraction and flash frozen in liquid nitrogen. Samples were ground under liquid nitrogen, and DNA was extracted from 100 mg aliquots by using Qiagen (Valencia, California, USA) DNeasy Plant-mini DNA extraction kits.

Outgroup selection—Over 50 members of the Heliantheae and several suspected close allies from Zinnieae and Ecliptinae (Appendix 1) were used in the initial sequencing of the *trnG* plastid locus to create unrooted trees (Appendix S1, see Supplemental Data with online version of this article). From these analyses, the genus *Sanvitalia* was found to be sister to *Echinacea* and was treated as the outgroup in subsequent analyses. DNA used in outgroup selection was obtained from previous studies where vouchers have already been deposited (Urbatsch et al., 2000).

Locus amplification, molecular cloning, and sequencing—**Nuclear loci**—Because there is little sequence information available for the genus *Echinacea* (14 sequences deposited in GenBank as of 11/16/07), nuclear loci were selected that previously have demonstrated high utility in molecular phylogenetic studies. Three nuclear loci were selected: alcohol dehydrogenase (*Adh*), cellulose synthase (*CesA*), and glyceraldehyde 3-phosphate acetyl transferase (*GPAT*) (Table 2). These loci were selected based on their previous utility in species-level molecular systematic studies (*Adh* [Sang et al., 1997; Small and Wendel, 2000]; *CesA* [Cronn et al., 2002; Senchina et al., 2003]; and *GPAT* [Tank and Sang, 2001]). Degenerate primers were used to perform preliminary locus amplification, after which *Echinacea*-specific primers were designed to

TABLE 2. Amplification conditions and primers for nuclear genes used in phylogenetic analysis in *Echinacea*.

Locus	Gene portion sequenced ^a	Forward nonspecific primer sequence (Forward <i>Echinacea</i> -specific primer sequence)	Reverse nonspecific primer sequence (Reverse <i>Echinacea</i> -specific primer sequence)	Annealing temperature (°C)
<i>Adh</i>	Exon 7–9	CTGCKGKTKGCATGGGARGCAGGGAA (CTGTTGTAACCAAATGAAACC)	GCACAGCCACACCCCAACCCCTG (TTCTTGTATCAAATTCAAAAGGGT)	51
<i>CesA</i>	Exon 9–12	CTTGGCCTGGAAATAACCCGCGTG (GGGTGATGCCAGAAGGGATG)	ATGCTCACACCGCTCCATC (CATCGAATTTCAAGGACACTGGTTG)	59.5
<i>GPAT</i>	Exon 4–9	TTTGGYCAAAATTATATTCGKCC (CGGCCTTTGATCAATTTTCAGGT)	CCACCACTKGGTGCAATCCA (CCATATTATTTTGGAGCCACCC)	56

^a Full-length sequences from *Echinacea* were not obtained; exon/intron position was estimated from orthologous full-length genes in closely related species.

span exonic and intronic regions such that amplicons of ~800–1100 bp could be generated for each locus (Table 2).

PCR amplification of the three nuclear loci was performed using the following generalized protocol: initial denaturation phase of 95°C for 5 min, 35 cycles of amplification at 95°C for 30 s, primer-specific annealing temperature (Table 2) for 45 s, 72°C elongation for 60 s. After 35 cycles of amplification, a final elongation phase of 72°C for 7 min was used to complete polymerization. PCR reactions were conducted in a 40 µL volume of 1× *Taq* polymerase buffer, 100–500 ng total genomic DNA, 2.0 mM MgCl₂, 0.4 µM of both forward and reverse primers, 0.25 mM dNTPs, and 2 units of *Taq* polymerase (BioLine USA, Randolph, Massachusetts, USA).

Individual PCR amplification products were visualized on 1% agarose gels, and products were excised and extracted with the Qiagen QIAEX II Gel Extraction Kit. The purified product was cloned into the pGEM-T Easy vector system (Promega, Madison, Wisconsin, USA) and transformed into chemically competent Mach1 T1 *E. coli* cells (Invitrogen, Carlsbad, California, USA). Transformed cells were plated and selected via a blue-white screen on LB Agar MILLER medium (EMD Chemicals, Gibbstown, New Jersey, USA) containing 50 mg/ml X-Gal and 0.1 M isopropyl β-D-1-thiogalactopyranoside. To allow for the assessment of PCR errors and allelic sequences, 8–12 colonies were selected from each individual. These transformed colonies were grown for 20 h in 150 µL of Terrific Broth (Invitrogen, Carlsbad, California, USA). Plasmids were isolated using a standardized alkaline lysis procedure, and inserts were sequenced with vector primers T7 and M13R following the ABI-Prism Big Dye Terminator sequencing method (version 3.1; Applied Biosystems, Foster City, California, USA). Sequence reactions were run on an Applied Biosystems ABI 3730 DNA analyzer at the Iowa State University DNA Facility.

Plastid loci—We chose two plastid loci that have been shown to contain relatively high levels of sequence diversity (Shaw et al., 2005). The loci, *trnS* and *trnG*, are both noncoding spacers within the plastid genome. Loci were amplified using the protocols of Shaw et al. (2005), and PCR products were purified with Bio-Edge columns (Edge BioSystems, Gaithersburg, Maryland, USA). Amplified product was sequenced off both primers using the ABI-Prism Big Dye Terminator sequencing method (version 3.1; Applied Biosystems). The sequence reactions were run on an Applied Biosystems ABI 3730 DNA analyzer at the Iowa State University DNA Facility.

Data processing, alignment, allele calling, and sequence polishing—**Nuclear loci**—Forward and reverse reads of raw sequence data were initially trimmed of vector sequence either with the program CROSS_MATCH (Ewing et al., 1998) or manually. Ambiguous bases from the ends of the reads were removed manually or using the trimseq program from the EMBOSS software package (Rice et al., 2000) with the following parameter settings (“window = 20” and “percent = 10”). Next, a consensus read was generated from the forward and reverse sequence reads using MUSCLE 3.52 multiple alignment software (Edgar, 2004). Each output alignment (hereafter referred to as a clone sequence) was saved for further analysis.

Clone sequences were imported and manually inspected with BioEdit sequence alignment viewing software (<http://www.mbio.ncsu.edu/BioEdit/bioedit.html>). Ambiguous bases in each clone sequence were corrected manually by comparing sequence quality from trace files. Corrected clones were assembled into individual-specific files and aligned with MUSCLE 3.52. Nontarget sequences were visibly detected and removed. Once allelic variants could be visually identified, consensus sequences were segregated into allele-specific files, which were aligned with MUSCLE 3.52 and condensed into one allelic consensus sequence. All allelic variants from each nuclear locus surveyed were collated into one final file, which was used for all downstream phylogenetic analyses. For each of the three nuclear loci, the allele number and nucleotide diversity (π) (Nei, 1987) (computed with the program DNASP 4.0 [Roza et al., 2003]) were calculated and tabulated (Tables 3 and 4).

Recombinant sequences can arise naturally via homologous recombination or artificially via PCR strand swapping (Bradley and Hillis, 1997), making detection and removal of recombinant sequences important, because they increase homoplasy and confound interpretation. We used two separate recombination detection algorithms, MaxChi (Smith, 1992) and SiScan (Gibbs et al., 2000), as implemented in the RDP-V2 program (Martin et al., 2005). A *P*-value of 0.01 was used as a threshold for significance when applied to 1000 parametric bootstrap replicates for both MaxChi and SiScan. Recombinant events detected within individuals, and thus likely arising from PCR strand-swapping, were removed from the analyses. Recombinant events detected between taxa, and thus likely arising naturally through hybridization and homologous recombination, were rare. The few detected events were not significant when step-down, multiple-testing correction was applied. These few sequences, though possibly recombinant, were left in the analyses. Paralogy tests were conducted by first identifying potential paralogous sequences from a phylogenetic tree (e.g., multiple placements for clones from

TABLE 3. Sequence information for *Echinacea* nuclear loci.

Locus	Individuals sequenced (species)	Total alignment bp (coding length)	No. alleles detected	% Heterozygosity all (diploids)
<i>Adh</i>	35 (8)	880 (485)	116	94.3 (92.9)
<i>CesA</i>	71 (9)	950 (644)	141	73.2 (74.19)
<i>GPAT</i>	84 (9)	1264 (337)	148	67.9 (63.89)

Fig. 1. Phylogenetic gene-tree reconstruction for the *CesA*, *Adh*, and *GPAT* gene loci. These phylogenies represent the Bayesian consensus trees and include node support values based on Bayesian posterior probabilities and branch lengths drawn relative to sequence divergence. The outgroup is indicated by “OG” (*Sanvitalia fruticosa* for *Adh* and *GPAT*; *Zinnia violacea* for *CesA*). The retention index (RI) and homoplasy index (HI) values are documented above each tree. All *Echinacea* species are classified as in McGregor (1968) and coded as follows: *E. angustifolia* (red), *E. atrorubens* (orange), *E. laevigata* (light green), *E. pallida* (purple), *E. paradoxa* (yellow), *E. purpurea* (dark blue), *E. sanguinea* (brown), *E. simulata* (light blue), and *E. tennesseensis* (dark green).

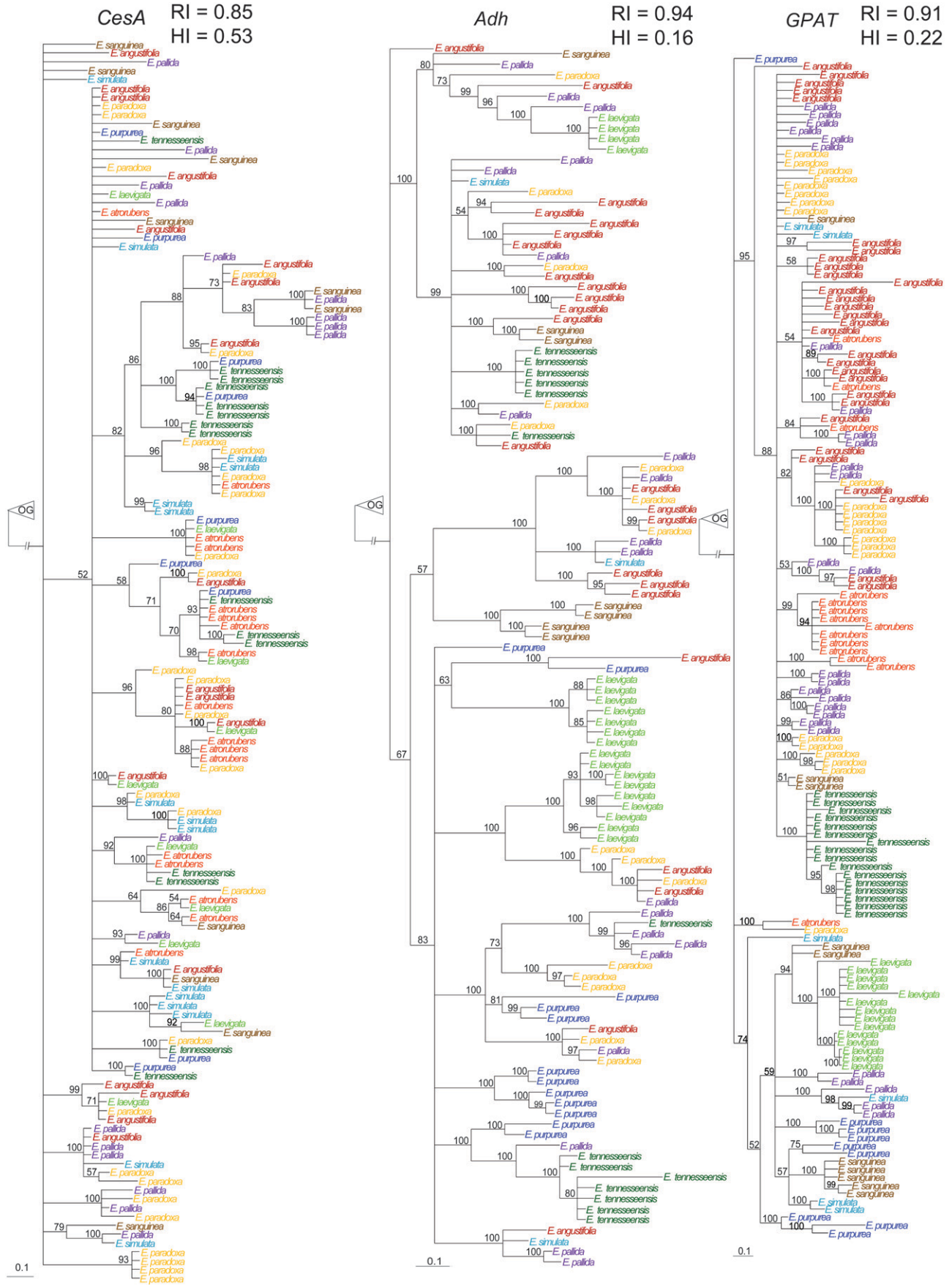


TABLE 4. Nucleotide diversity (π) partitioned between various genic domains. SD values are the standard deviations associated with each estimate of π .

Locus	π				
	Overall (SD)	Coding regions (SD)	Noncoding regions (SD)	Syn sites	Nonsyn sites
<i>Adh</i>	0.03224 (0.00082)	0.03947 (0.00103)	0.02128 (0.00100)	0.06857	0.0338
<i>CesA</i>	0.01491 (0.00035)	0.00628 (0.00031)	0.03846 (0.00111)	0.02538	0.00064
<i>GPAT</i>	0.01166 (0.00046)	0.00457 (0.00038)	0.01545 (0.00063)	0.01401	0.00194

Notes: Nonsyn, nonsynonymous sites: nucleotide sites that cause an amino acid replacement if substituted. Syn, synonymous sites: nucleotide sites that do not cause an amino acid replacement if substituted.

a given individual or phylogenetically suspicious placements) and then manually comparing sequence alignments in search of paralog-specific signatures. Following detection, primers were developed to target one paralog, and sequences from the nontargeted paralog were removed from further analyses.

Plastid loci—For the plastid sequences, forward and reverse reads were combined manually, and the resulting sequences were aligned in the program CLUSTAL_X (Thompson et al., 1997). Minimal manual adjustment was necessary because sequence diversity was low.

Phylogenetic analyses—**Nuclear loci**—Three different phylogenetic analyses were applied to the data: (1) distance-based analyses with nonparametric bootstrapping, performed with the program Phylip 3.63 (Felsenstein, 1989); (2) parsimony analyses using PAUP* 4.0 (Swofford, 2001); and (3) Bayesian-likelihood analyses with MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003). All three analyses gave highly congruent topologies (data not shown), though only Bayesian phylogenies have the advantage of retaining both branch length and node support; thus, only these phylogenies are shown (Fig. 1).

Bayesian phylogenies for each nuclear locus were estimated using a Markov chain Monte Carlo (MCMC) sampler of tree space as implemented by MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003). For all phylogenies, 2–4 MCMC runs were initiated, each with a minimum of 1 000 000 generations (more generations used when needed to reach stationarity). Prior distribution settings were left at the default values, with the exceptions of the nucleotide substitution model, which was altered to allow unique rates of substitution among or between all pairs of nucleotides (e.g., the general time reversible [GTR] model), and the rate model, which was drawn from a gamma distribution while allowing for invariant sites. Runs were started from a random tree and allowed to proceed in parallel while sampling and recording the topology every 100 generations of the MCMC chain. Performance of individual runs was assessed and phylogenies compared between runs. Majority rule (>50%) consensus trees were constructed after removing the “burn-in period” samples (the first 10% of sampled trees). Topologically, Bayesian analyses were highly congruent between runs, indicating that multiple MCMC chains consistently achieved stationarity around the same subset of possible topologies.

Further, as an exploratory tool we used the group-assignment program Structure 2.1 (Pritchard et al., 2000) to give an alternative view of the gene-sequence data. With the goal of understanding genomic levels of gene flow, the Structure algorithm performs best with randomly sampled unlinked markers, such as AFLPs and simple sequence repeats (SSRs); however, it can be used with nuclear sequence data as an exploratory tool, giving a graphical overview of population structure within these data. Analyses were run separately on polymorphic base pairs from the three nuclear gene datasets. For each data set, the number of clusters (K) was incremented from 1 to 12, and the best fit was assessed with a likelihood ratio test. Group assignments were plotted for individuals and species for each value of K (Appendix S2, see Supplemental Data with online version of article).

Plastid loci—Plastid nucleotide diversity within *Echinacea* and among outgroup species was low. We applied the same Bayesian phylogeny reconstruction method used for nuclear loci. The Bayesian tree search algorithm was allowed to run for 1 000 000 generations, achieving stationarity and the resulting consensus tree can be found in Fig. 2.

RESULTS

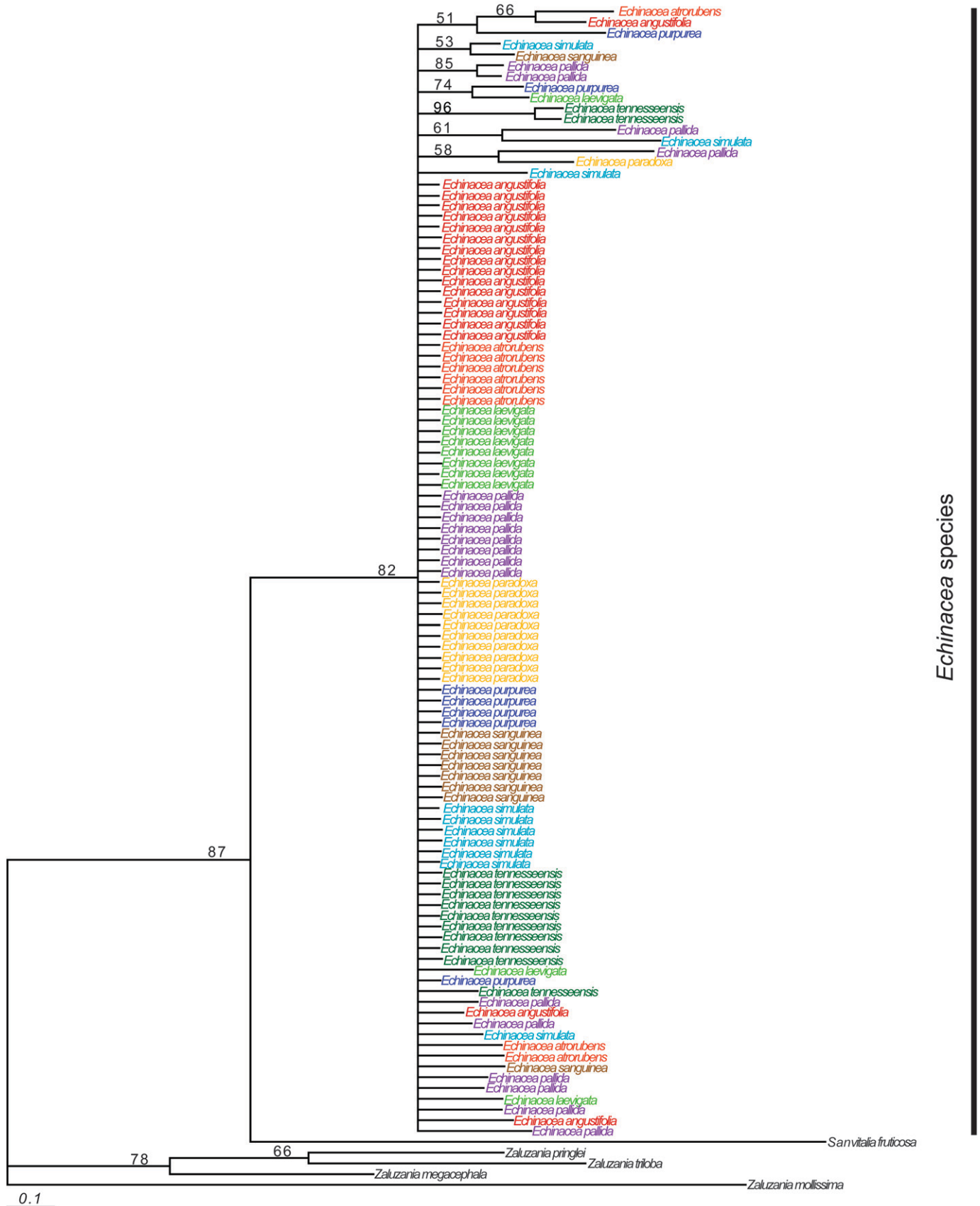
Nuclear loci sequence characteristics—In total, approximately 3.1 Mb of *Echinacea* nuclear DNA were sequenced, including 1 Mb for *Adh*, 1.2 Mb for *CesA*, and 0.93 Mb for *GPAT*. After allelic sequences were processed and identified, the raw data generated approximately 92 kb, 138 kb, and 151 kb, respectively, of total unaligned sequence data for phylogenetic analysis. For *GPAT* and *Adh*, we also amplified and sequenced nuclear loci from *Sanvitalia fruticosa* Hemsl. We were unable to amplify the *CesA* locus in *S. fruticosa* and instead used *Zinnia violacea* Cav., a close relative, (GenBank accessions AF323039, AF323040, and AF323041) as an outgroup for this data set.

Observed levels of heterozygosity were between approximately 68 and 94% (Table 3). Alleles were defined strictly by haplotype, thus alleles may differ at a minimum by a single nucleotide polymorphism. We have applied this strict assessment because we have sequenced multiple clones per individual (between 8–12), allowing us the opportunity to remove many PCR and sequencing errors. Heterozygosity values for *CesA* and *GPAT* were similar (68% and 73%). However, much higher heterozygosity was found in *Adh* (~94%), possibly due to an increased substitution rate caused by a loss of purifying selection on pseudogenized sequences at this locus. In all cases, heterozygosity was relatively high, which could result from a reported sporophytic self-incompatibility system in the genus *Echinacea* (McKeown, 2004; Stephens, in press).

Overall mean values of nucleotide diversity, π , ranged from ~0.012 to ~0.032 for the three nuclear loci (Table 4). We partitioned nucleotide diversities between coding and noncoding regions and synonymous and nonsynonymous sites within coding regions (Table 4). As expected, the levels of π for *CesA* and *GPAT* were greater in noncoding than in coding regions (6.12 and 3.38 times greater, respectively); likewise, π in synonymous sites was approximately 39.7 and 7.2 times greater, respectively, than in nonsynonymous sites.

A rather different pattern was observed at the *Adh* locus, which had the highest overall mean values of π (0.03229). In addition, levels of π were approximately 54% lower at noncoding sites (0.02128) than at coding sites (0.03947), and levels of π were only approximately two times higher when comparing synonymous (0.06857) to nonsynonymous sites (0.0338). These statistics would be unusual for a functional nuclear gene experiencing neutral evolution (Li, 1997). These factors, along with stop codons and indels in the open-reading frames of several taxa (data not shown), suggest that the *Adh* locus we sequenced represents either a pseudogene or possibly a nuclear locus with nonfunctional allelic variants. We were unable to isolate orthologous *Adh* loci

Fig. 2. Genus-level phylogenetic reconstruction, including all *Echinacea* species, using a concatenated plastid locus data set (*trnS* and *trnG*). All species within the genus *Echinacea* formed a single monophyletic group. The genus *Sanvitalia* appears sister to *Echinacea* and was used as an outgroup for the nuclear data set (Fig. 1). All *Echinacea* species are classified as in McGregor (1968); color coding follows Fig. 1.



from some taxa [*E. atrorubens* (Nutt.) Nutt.], likely due to the higher rate of loss of pseudogenized genes. For this reason, taxon sampling in the *Adh* data set remains incomplete. We report these findings regarding the limited phylogenetic utility of the *Adh* locus in the hopes that it may be avoided in future studies in *Echinacea*. Also, it serves as an example of one of the pitfalls often encountered when selecting nuclear loci for phylogenetic studies.

Phylogenetic results—Nuclear loci—Topologies of the three nuclear gene trees are shown in Fig. 1. Overall, few species form monophyletic groups with respect to these gene trees. The exceptions are *E. laevigata* and *E. tennesseensis*, which both form monophyletic groups in the *GPAT* tree. In addition, we observed no phylogenetic differentiation between varietal groups within either *E. angustifolia* or *E. paradoxa* (data not shown), and thus, we removed varietal designations from Fig. 1. The Bayesian *GPAT* gene tree divides the genus into two clades, with alleles from *E. angustifolia*, *E. atrorubens*, *E. paradoxa*, and *E. tennesseensis* in one clade and alleles from *E. laevigata*, *E. purpurea*, *E. sanguinea* Nutt., and *E. simulata* McGregor in the other. This split was also observed in our parsimony and distance-based trees, with 100% bootstrap support in the latter (data not shown). There are, however, a few exceptions to this division, i.e., one *E. sanguinea* and two *E. simulata* alleles can be found within the *E. angustifolia*, *E. atrorubens*, *E. tennesseensis*, and *E. paradoxa* clade. Such a division has been documented by others (Kim et al., 2004), though these authors found *E. laevigata* sister to *E. tennesseensis*. Barring this exception, the first chronological divergence that takes place in the *GPAT* phylogeny is well-supported by multiple phylogenetic methods and by the AFLP data from Kim et al. (2004). Resolution beyond this initial division in the *GPAT* phylogeny becomes less clear because many of the taxa share alleles with other taxa; the exceptions are *E. tennesseensis* and *E. laevigata* as noted.

The topological patterns of the *Adh* and *CesA* gene phylogenies are more complex than that of *GPAT*. Neither phylogeny has a single monophyletic species; furthermore, there is often reliable node support for polyphyletic associations in both phylogenies. Notably though, in both phylogenies there frequently are small clades of alleles from the same species, although these clades are paraphyletic with regard to species in all cases. An additional confounding factor is a high level of homoplasy found in the *CesA* phylogeny. The most parsimonious class of tree scores (of which there were many equally parsimonious trees) overall had a homoplasy index (HI) of 0.53. Comparatively, the *Adh* and *GPAT* loci had HI values of 0.16 and 0.22 respectively. All three nuclear phylogenies are populated by both short internal and terminal branches; thus, it is not surprising that there are several unresolved polytomies.

Plastid loci—Plastid analysis found the relationship of genera allied with *Echinacea*. The Mexican and Southwest US endemic genus *Sanvitalia* appears as sister to *Echinacea* with good node support (Fig. 2; online Appendix S1). Within the genus nucleotide diversity is extremely low and results in a phylogenetic hypothesis rich in polytomies (Fig. 2). Using this plastid phylogeny of the genus, we compared the genetic distances to geographic distances via a Mantel test as implemented in the program PASSaGE (Rosenberg, 2001). This test demonstrated that the genetic structure of these *Echinacea* plastid loci is statistically correlated to their relative geographic distances from one another ($P < 0.05$) and not to taxon label.

DISCUSSION

The primary goals of this study were to generate a large collection of sequence data for diverse populations of *Echinacea* and to use these data to reconstruct a species-level phylogeny. Previous attempts to reconstruct the genetic and evolutionary relationships of *Echinacea* (Kapteyn et al., 2002; Binns et al., 2004; Kim et al., 2004; Mechanda et al., 2004) have provided phylogenetic resolution among particular taxa but have been limited by their depth of population sampling and/or number of phylogenetically useful characters. Our approach to determine the evolutionary history of this genus was to use plastid and nuclear loci, the latter typically offering excellent resolution at the species level due to relatively high rates of sequence divergence when the assumptions of the phylogenetic model are met. Nuclear sequence data, however, may also be phylogenetically problematic or misleading. For example, hybridization, paralogy, incomplete lineage sorting, and secondary contact are all features or processes capable of obfuscating organismal-level relationships in a phylogenetic framework. In addition to these biological features, technical issues, such as PCR recombinants (Small et al., 2004) and automated sequencer base-calling errors, play an increasingly troubling role as nucleotide diversity decreases among the taxa being sampled.

Although no definitive species-level relationships may be formed from the topologies generated, several features of the data set are particularly striking, given the connections of this genus to North American glacial history and geography. In the plastid data, close ties between genetic structure and geographical distribution suggest a prominent role of past rounds of glaciation. First, the cytotypes suggest southerly refugia on either side of the Mississippi River, with both containing a unique cytotype along with other cytotypes present in both refugia. The nuclear data further corroborate the idea of secondary, postglacial, contact between species with incomplete reproductive barriers. Telling aspects include low sequence diversity but a high number of alleles, broad taxonomic distribution of nearly identical alleles, and incongruent topologies between loci.

Notwithstanding the general absence of species-level monophyly in the trees generated from nuclear loci, these trees do offer some insight into species origins and history. One example concerns the origins of the polyploid species *E. pallida*, which appears interspersed throughout the diploid phylogeny. The placement of *E. pallida* with *E. angustifolia*, *E. atrorubens*, and *E. laevigata* alleles is frequent, though *E. pallida* can be found sister to other taxa as well (Fig. 1). Additionally, based on our population-structure analysis, designations for *E. pallida* tended to be assorted evenly among several groups (online Appendix S2B, D, and F). These results indicate that the polyploid *E. pallida* was either formed more than once from different parental origins or that the formation of *E. pallida* took place at a basal level in the genus and the observed patterns at the tips of the trees are artifacts of subsequent hybridization and incomplete lineage sorting (Wendel and Doyle, 1998; Small et al., 2004). Given the shallow nature of these trees, which yield poor basal resolution, it is difficult to determine which hypothesis is correct.

In stark contrast to the sequence data reported here, a recent chemotaxonomic study by Wu et al. (L. Wu, Iowa State University; P. Dixon, B. Nikolau, G. Kraus, M. Widrlechner, and E. Wurtele, unpublished manuscript) finds strong support for McGregor's (1968) taxonomic treatment when individuals are clustered based on their metabolite profiles. Additionally, a

recent morphological analysis by Binns et al. (2002) also contains node support for McGregor's classification. It is interesting to note that although our analysis of neutral gene variation suggests supraspecific, landscape-scale processes at work, the physiology and morphology are consistent with well-differentiated and adapted species, perhaps reflecting specific ecological niches. It is possible that the high degree of physiological and morphological integrity has been maintained by selection on relatively few loci, which were not sampled during this study. Alternatively, in the absence of selection it is possible that neutral processes fixed these traits during glacial maxima, when population sizes were presumably small.

In either case, future studies seeking to elucidate genetic relationships within this genus should probably employ marker technology that has broad genomic coverage, such as AFLPs, as the phylogenetic signal within the nuclear and plastid genome appears to be extremely weak. This study has also highlighted the importance of including geographically representative individuals from all species, as using a subset of these data could easily generate an incorrect yet well-supported topology.

LITERATURE CITED

- AULT, J. 2006. Coneflower, *Echinacea* species. In N. Anderson [ed.], Flower breeding and genetics: Issues, challenges, and opportunities for the 21st century, 799–822. Springer, Dordrecht, Netherlands.
- BAUM, B. R., S. MECHANDA, J. F. LIVESEY, S. E. BINNS, AND J. T. ARNASON. 2001. Predicting quantitative phytochemical markers in single *Echinacea* plants or clones from their DNA fingerprints. *Phytochemistry* 56: 543–551.
- BINNS, S. E., J. T. ARNASON, AND B. R. BAUM. 2004. Taxonomic history and revision of the genus *Echinacea*. In S. C. Miller and H.-c. Yu [eds.], *Echinacea: The genus Echinacea*, 3–12. CRC Press, Boca Raton, Florida, USA.
- BINNS, S. E., B. R. BAUM, AND J. T. ARNASON. 2002. A taxonomic revision of *Echinacea* (Asteraceae: Heliantheae). *Systematic Botany* 27: 610–632.
- BLUMENTHAL, M., AND L. E. URBATSCH. 2006. Echinacea taxonomy—Is the re-classification of the genus warranted? *HerbalGram* 72: 30–31.
- BRADLEY, R. D., AND D. M. HILLIS. 1997. Recombinant DNA sequences generated by PCR amplification. *Molecular Biology and Evolution* 14: 592–593.
- CLAYTON, L., AND S. R. MORAN. 1982. Chronology of late Wisconsinan glaciation in middle North America. *Quaternary Science Reviews* 1: 55–82.
- CRONN, R. C., R. L. SMALL, T. HASELKORN, AND J. F. WENDEL. 2002. Rapid diversification of the cotton genus (*Gossypium*: Malvaceae) revealed by analysis of sixteen nuclear and chloroplast genes. *American Journal of Botany* 89: 707–725.
- EDGAR, R. C. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32: 1792–1797.
- EWING, B., L. HILLIER, M. C. WENDL, AND P. GREEN. 1998. Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. *Genome Research* 8: 175–185.
- FELSENSTEIN, J. 1989. PHYLIP—Phylogeny inference package (version 3.2). *Cladistics* 5: 164–166.
- FLORA OF NORTH AMERICA EDITORIAL COMMITTEE [EDS.]. 1993+. Flora of North America North of Mexico, Oxford University Press, New York, New York, USA.
- GIBBS, M. J., J. S. ARMSTRONG, AND A. J. GIBBS. 2000. Sister-scanning: A Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics (Oxford, England)* 16: 573–582.
- KAPTEYN, J., B. GOLDSBROUGH, AND E. SIMON. 2002. Genetic relationships and diversity of commercially relevant *Echinacea* species. *Theoretical and Applied Genetics* 105: 369–376.
- KIM, D. H., D. HEBER, AND D. W. STILL. 2004. Genetic diversity of *Echinacea* species based upon amplified fragment length polymorphism markers. *Genome* 47: 102–111.
- LI, W. 1997. Molecular evolution. Sinauer, Sunderland, Massachusetts, USA.
- MARTIN, D. P., C. WILLIAMSON, AND D. POSADA. 2005. RDP2: Recombination detection and analysis from sequence alignments. *Bioinformatics (Oxford, England)* 21: 260–262.
- MCGREGOR, R. 1968. The taxonomy of the genus *Echinacea* (Compositae). *University of Kansas Science Bulletin* 48: 113–142.
- MCKEOWN, K. A. 1999. A review of the taxonomy of the genus *Echinacea*. In J. Janick [ed.], Perspectives on new crops and new uses, 482–489. American Society for Horticultural Science Press, Alexandria, Virginia, USA.
- MCKEOWN, K. A. 2004. A review of preliminary *Echinacea* genetics and the future potential of genomics. In S. C. Miller and H.-c. Yu [eds.], *Echinacea: The genus Echinacea*, 13–20. CRC Press, Boca Raton, Florida, USA.
- MECHANDA, S., B. R. BAUM, D. A. JOHNSON, AND J. T. ARNASON. 2004. Analysis of diversity of natural populations and commercial lines of *Echinacea* using AFLP. *Canadian Journal of Botany* 82: 461–484.
- NEI, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York, New York, USA.
- PRICE, D. H., AND K. KINDSCHER. 2007. One hundred years of *Echinacea angustifolia* harvest in the Smoky Hills of Kansas, USA. *Economic Botany* 61: 86–95.
- PRITCHARD, J. K., M. STEPHENS, AND P. DONNELLY. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.
- RICE, P., I. LONGDEN, AND A. BLEASBY. 2000. EMBOSS: The European molecular biology open software suite. *Trends in Genetics* 16: 276–277.
- RONQUIST, F., AND J. P. HUELSENBECK. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics (Oxford, England)* 19: 1572–1574.
- ROSENBERG, M. S. 2001. PASSAGE: Pattern analysis, spatial statistics, and geographic exegesis. Department of Biology, Arizona State University, Tempe, Arizona, USA.
- ROZAS, J., J. C. SÁNCHEZ-DELBARRIO, X. MESSEGUER, AND R. ROZAS. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics (Oxford, England)* 19: 2496–2497.
- SANG, T., M. J. DONOGHUE, AND D. ZHANG. 1997. Evolution of alcohol dehydrogenase genes in peonies (*Paeonia*): Phylogenetic relationships of putative nonhybrid species. *Molecular Biology and Evolution* 14: 994–1007.
- SARI, A. O., M. R. MORALES, AND J. E. SIMON. 2001. Ethephon can overcome seed dormancy and improve seed germination in purple coneflower species *Echinacea angustifolia* and *E. pallida*. *HortTechnology* 11: 202–205.
- SCHOOP, R., P. KLEIN, A. SUTER, AND S. L. JOHNSTON. 2006. *Echinacea* in the prevention of induced rhinovirus colds: A meta-analysis. *Clinical Therapeutics* 28: 174–183.
- SENCINA, D. S., I. ALVAREZ, R. C. CRONN, B. LIU, J. RONG, R. D. NOYES, A. H. PATERSON, R. A. WING, T. A. WILKINS, AND J. F. WENDEL. 2003. Rate variation among nuclear genes and the age of polyploidy in *Gossypium*. *Molecular Biology and Evolution* 20: 633–643.
- SHAW, J., E. B. LICKEY, J. T. BECK, S. B. FARMER, W. LIU, J. MILLER, K. C. SIRIPUN, C. T. WINDER, E. E. SCHILLING, AND R. L. SMALL. 2005. The tortoise and the hare II: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany* 92: 142–166.
- SMALL, R. L., R. C. CRONN, AND J. F. WENDEL. 2004. Use of nuclear genes for phylogeny reconstruction in plants. *Australian Systematic Botany* 17: 145–170.
- SMALL, R. L., AND J. F. WENDEL. 2000. Copy number lability and evolutionary dynamics of the *Adh* gene family in diploid and tetraploid cotton (*Gossypium*). *Genetics* 155: 1913–1926.
- SMITH, J. M. 1992. Analyzing the mosaic structure of genes. *Journal of Molecular Evolution* 34: 126–129.

- SPERONI, E., P. GOVONI, S. GUIZZARDI, C. RENZULLI, AND M. C. GUERRA. 2002. Anti-inflammatory and cicatrizing activity of *Echinacea pallida* Nutt. root extract. *Journal of Ethnopharmacology* 79: 265–272.
- STEPHENS, L. C. In press. Self incompatibility in *Echinacea purpurea*. *HortScience*.
- SWOFFORD, D. L. 2001. PAUP*: Phylogenetic analysis using parsimony (*and other methods), version 4. Sinauer, Sunderland, Massachusetts USA.
- TANK, D. C., AND T. SANG. 2001. Phylogenetic utility of the glycerol-3-phosphate acyltransferase gene: Evolution and implications in *Paeonia* (Paeoniaceae). *Molecular Phylogenetics and Evolution* 19: 421–429.
- THOMPSON, J. D., T. J. GIBSON, F. PLEWNIAC, F. JEANMOUGIN, AND D. G. HIGGINS. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* 25: 4876–4882.
- TURNER, R. B., R. BAUER, K. WOELKART, T. C. HULSEY, AND J. D. GANGEMI. 2005. An evaluation of *Echinacea angustifolia* in experimental rhinovirus infections. *New England Journal of Medicine* 353: 341–348.
- URBATSCH, L. E., B. G. BALDWIN, AND M. J. DONOGHUE. 2000. Phylogeny of the coneflowers and relatives (Heliantheae: Asteraceae) based on nuclear rDNA internal transcribed spacer (ITS) sequences and chloroplast DNA restriction site data. *Systematic Botany* 25: 539–565.
- URBATSCH, L. E., K. M. NEUBIG, AND P. B. COX. 2006. *Echinacea* Moench, Methodus. In Flora of North America Editorial Committee [eds.], Flora of North America North of Mexico, vol. 21, 88–92, Oxford University Press, New York, New York and Oxford UK. Available at website http://www.efloras.org/florataxon.aspx?flora_id=1&taxon_id=111203 [accessed 15 January 2008].
- WENDEL, J. F., AND J. J. DOYLE. 1998. Phylogenetic incongruence: Window into genome history and molecular evolution. In D. Soltis, P. Soltis, and J. Doyle [eds.], Molecular systematics of plants II: DNA sequencing, 265–296. Kluwer, Boston, Massachusetts, USA.
- WIDRLECHNER, M., AND K. A. MCKEOWN. 2002. Assembling and characterizing a comprehensive *Echinacea* germplasm collection. In J. Janick and A. Whipkey [eds.], Trends in new crops and new uses, 506–508. American Society for Horticultural Science Press, Alexandria, Virginia, USA.
- YU, H.-C., AND M. KAARLAS. 2004. Popularity, diversity, and quality of *Echinacea*. In S. C. Miller and H.-c. Yu [eds.], *Echinacea: The genus Echinacea*, 127–150. CRC Press, Boca Raton, Florida, USA.

APPENDIX 1. The taxa analyzed in this study, collection locality, U.S. Department of Agriculture GRIN database Plant Introduction (PI) accession, voucher reference followed by the Herbarium of voucher deposition, and the corresponding GenBank accessions for nuclear and plastid loci. Herbaria abbreviations: Iowa State University Ada Hayden Herbarium = ISC, Louisiana State University Herbarium = LSU, University of Texas Herbarium = TEX, University of California, Berkeley Jepson Herbarium = UC/JEPS. Sequential GenBank accession numbers have been shortened with a hyphen; e.g., EU423454-6 indicates accessions EU423454, EU423455, and EU423456. Nonsequential GenBank accessions are separated with a comma. All nonapplicable values are indicated with a dash (—).

Taxon—Collection locality; PI accession; *Voucher*; Herbaria; GenBank accessions: *Adh*; *Cesa*; *GPAT*; *trnG*; *trnS*.

Acmella lundellii R.K.Jansen; Belize; —; Turner 0-81; TEX; —; —; —; EU586863; —. *Acmella radicans* (Jacquin) R.K.Jansen; Mexico; —; Vazquez 274; TEX; —; —; —; EU586867; EU440302, EU440322. *Aspilia purpurea* Greenm.; Mexico; —; Urbatsch 3340, LSU; —; —; —; —; EU440310.

Balsamorhiza careyana A.Gray; Washington; —; Baldwin s.n.; UC/JEPS; —; —; —; EU440128; —. *Balsamorhiza hirsuta* Nutt.; Washington; —; Baldwin s.n.; UC/JEPS; —; —; —; EU440127; —. *Balsamorhiza hookeri* Nutt.; California; —; Baldwin s.n.; UC/JEPS; —; —; —; EU440134; —. *Balsamorhiza serrata* A.Nelson & J.F.Macbr.; Nevada; —; Baldwin s.n.; UC/JEPS; —; —; —; EU440131; —. *Balsamorhiza serrata*; Oregon; —; Baldwin s.n.; UC/JEPS; —; —; —; EU440121; —. *Baldwinia angustifolia* B.L.Rob.; Florida; —; Urbatsch s.n.; LSU; —; —; —; EU440116; —. *Berlandiera lyrata* Benth.; Arizona; —; Urbatsch 7098; LSU; —; —; —; EU440126; —. *Berlandiera pumila* Nutt.; Florida; —; Urbatsch s.n.; LSU; —; —; —; EU440129; —. *Berlandiera subcaulis* (Nutt.) Nutt.; Florida; —; Urbatsch s.n.; LSU; —; —; —; EU440120; —.

Echinacea angustifolia; Oklahoma; PI 631267; —; EU423454-6; EU423562-7; EU423705-12; EU440100, EU440164; EU440198, EU440251, EU440285. *Echinacea angustifolia* var. *angustifolia*; Oklahoma; PI 631272; Rapp et al. 66; ISC; EU423457-9; EU423568; EU423713-6; EU440152-3; EU440155; EU440228, EU440242, EU440271. *Echinacea angustifolia* var. *angustifolia*; Iowa; PI631285; Rapp et al. 78; ISC; EU423460-3; EU423569; EU423717-22; EU440150, EU440173, EU440186; EU440213, EU440229, EU440265. *Echinacea angustifolia* var. *angustifolia*; Kansas; PI 631318; —; —; EU423570-3; EU423723-6; EU440089, EU440143, EU440149; EU440218, EU440232, EU440278. *Echinacea angustifolia* var. *strigosa* McGregor; Oklahoma; PI 631266; —; EU423446-53; EU423574-77; EU423727-8; EU440094, EU440098, EU440162; EU440191, EU440223, EU440258, EU440263, EU440273. *Echinacea angustifolia* var. *strigosa*; Oklahoma; PI 631320; —; EU423464-69; EU423578-80; EU423729-34; EU440090, EU440110, EU440165; EU440193, EU440203, EU440209, EU440277. *Echinacea atrorubens*; Kansas; PI 631255; —; —; EU423584-8; EU423735-8; EU440083, EU440103; EU440236,

EU440256, EU440262, EU440276. *Echinacea atrorubens*; Kansas; PI 631299; —; —; EU423594-6; EU423747-8; EU440092, EU440179, EU440189; EU440296, EU440305. *Echinacea atrorubens*; Oklahoma; PI 631260; —; —; EU423581-3; EU423739-43; EU440109, EU440170, EU440174; EU440194, EU440239, EU440247. *Echinacea atrorubens*; Oklahoma; PI 631262; —; —; EU423589-93; EU423744-6; EU440096; EU440190, EU440292. *Echinacea laevigata*; South Carolina; PI 631310; Rapp et al. 67; ISC; EU423470-6; EU423597-8; EU423749-50; EU440078, EU440138; EU440196-7, EU440205, EU440211. *Echinacea laevigata*; South Carolina; PI 631312; Rapp et al. 65; ISC; EU423477-82; EU423599-600; EU423751-4; EU440107; EU440217, EU440267, EU440283. *Echinacea laevigata*; North Carolina; PI 631314; —; EU423483-6; EU423601-2, EU423606; EU423755-57; EU440085, EU440101; EU440201, EU440207, EU440264, EU440272. *Echinacea laevigata*; Virginia; PI 631316; Rapp et al. 64; ISC; EU423487-9; EU423603-5; EU423758-61; EU440076, EU440084, EU440140, EU440241, EU440250. *Echinacea pallida*; Oklahoma; PI 631275; —; EU423492-98; EU423607-9; EU423762-7; EU440095, EU440181; EU440289, EU440293, EU440297. *Echinacea pallida*; Iowa; PI 631290; —; EU423499-503; EU423610-1; EU423768-72; EU440077, EU440144; EU440208, EU440225, EU440254. *Echinacea pallida*; Arkansas; PI 631293; Rapp et al. 80; ISC; EU423504-7; EU423612-6; EU423773-5; EU440111, EU440148; EU440212, EU440221, EU440281. *Echinacea pallida*; Missouri; PI 631296; Rapp et al. 91; ISC; EU423508-9; EU423617-9; EU423776-82; EU440139, EU440154, EU440157; EU440290, EU440294, EU440298, EU440306. *Echinacea pallida*; North Carolina; PI 631315; —; —; EU423620-3; EU423783-9; EU440159-60, EU440171; EU440309, EU440315, EU440320. *Echinacea paradoxa* var. *neglecta*; Oklahoma; PI 631263; —; EU423490-1; EU423626-8; EU423792-4; EU440097, EU440145, EU440187; —. *Echinacea paradoxa* var. *neglecta*; Oklahoma; PI 631264; —; EU423510-4; EU423629-31; EU423795-8; EU440075, EU440086, EU440156; EU440231, EU440238, EU440246, EU440284. *Echinacea paradoxa* var. *neglecta*; Oklahoma; PI 631265; —; EU423515-9; EU423624-5; EU423799-801; EU440080-1; EU440199, EU440255, EU440300. *Echinacea paradoxa* var. *paradoxa*; Arkansas; PI 631292; —; EU423520-4; EU423633-40; EU423802-4;

- EU440167, EU440178; EU440210, EU440219, EU440268, EU440313, EU440321. *Echinacea paradoxa* var. *paradoxa*; Missouri; *PI 631301*; —; —; EU423641-4; EU423805-8; EU440151, EU440180; —. *Echinacea paradoxa* var. *paradoxa*; Missouri; *PI 631321*; —; EU423525; EU423645-51; EU423809-11; EU440079, EU440088, EU440169; EU440206, EU440216, EU440224. *Echinacea purpurea*; Louisiana; *PI 633669*; —; EU423526-7; EU423652-5; EU423812-4; EU440168, EU440176-7; EU440230, EU440244, EU440287. *Echinacea purpurea*; Missouri; *PI 631307*; *Rapp et al. 74*; ISC; EU423528-32; EU423656-7; EU423815-7; EU440114, EU440182, EU440185; EU440259, EU440269. *Echinacea purpurea*; North Carolina; *PI 631313*; *Rapp et al. 71*; ISC; EU423533-7; EU423658-60; EU423818-20; EU440082, EU440105; EU440234, EU440249. *Echinacea sanguinea*; Louisiana; *PI 631257*; *Rapp et al. 94*; ISC; EU423538-9; EU423661-2; EU423821-2; EU440146-7; EU440243, EU440252, EU440275. *Echinacea sanguinea*; Louisiana; *PI 631258*; —; EU423540-2; EU423663-6; EU423823-7; EU440087, EU440102, EU440112; EU440200, EU440215, EU440260, EU440280. *Echinacea sanguinea*; Louisiana; *PI 633672*; —; EU423543-4; EU423667-71; EU423828-30; —. *Echinacea simulata*; Kentucky; *PI 631249*; *Rapp et al. 92*; ISC; —; EU423672-5; EU423831-2; EU440099, EU440234; EU440257, EU440279, EU440288. *Echinacea simulata*; Missouri; *PI 631304*; —; —; EU423676-80; EU423833-4; EU440091, EU440093, EU440141; EU440235, EU440282. *Echinacea simulata*; Tennessee; *PI 631308*; *Rapp et al. 89*; ISC; EU423545-7; EU423681-87; EU423835-6; EU440113, EU440183-4; EU440261, EU440270. *Echinacea tennesseensis*; Tennessee; *PI 631250*; *Rapp et al. 84*; ISC; EU423548-51; EU423688-91; EU423837-40; EU440135-6, EU440163; EU440195, EU440204, EU440248. *Echinacea tennesseensis*; Tennessee; *PI 631324*; *Rapp et al. 81*; ISC; EU423552-4; EU423692-5; EU423841-2; EU440106, EU440137; EU440226, EU440233, EU440240. *Echinacea tennesseensis*; Tennessee; *PI 631325*; *Rapp et al. 86*; ISC; EU423555-7; EU423696-7; EU423843-8; EU440074, EU440166, EU440175; EU440192, EU440202, EU440245. *Echinacea tennesseensis*; Tennessee; *PI 631326*; *Rapp et al. 85*; ISC; EU423558-61; EU423698-702; EU423849-52; EU440188; EU440220, EU440227, EU440253. *Eclipta prostrata* (L.) L.; Louisiana; —; *Neyland s.n.*; LSU; —; —; —; EU586869; —. *Encelia farinosa* A.Gray ex Torr.; Arizona; —; *Urbatsch 7099*; LSU; —; —; —; EU586875; —.
- Helianthella quinquenervis* (Hook.) A. Gray; Colorado; —; *Burkhalter 2674*; LSU; —; —; —; EU586870; —. *Helianthella uniflora* Torr. & Gray; Wyoming; —; *Urbatsch 7064*; LSU; —; —; —; EU440130; —. *Heliopsis filifolia* S. Watson; Mexico; —; *Wendt et al. 1896*; TEX; —; —; —; EU586846; —. *Heliopsis helianthoides* Sweet; Missouri; —; *Urbatsch s.n.*; LSU; —; —; —; EU586878; —. *Heliopsis oppositifolia* (L.) Druce; Texas; —; *Butterwick 3833*; TEX; —; —; —; EU586845; —. *Heliopsis parvifolia* A.Gray; Mexico; —; *Wendt et al. 1662*; TEX; —; —; —; EU586848; —.
- Parthenium hysterophorus* L.; Louisiana; —; *Urbatsch s.n.*; LSU; —; —; —; EU440122; —. *Podachaenium pachyphyllum* (Klotzsch) R.K.Jansen, N.A.Harriman & Urbatsch; Mexico; —; *Panero 2760*; TEX; —; —; —; EU586854; —. *Philactis nelsonii* S.F.Blake; Mexico; —; *Breedlove 28267*; TEX; —; —; —; EU586885; EU440323. *Philactis zinniioides* Scrad.; Mexico; —; *Hartman & Funk 4184*; TEX; —; —; —; EU586853; —.
- Ratibida columnifera* (Nutt.) Wootton & Standl.; Louisiana; —; *Richardson 60*; LSU; —; —; —; EU440161; —. *Ratibida mexicana* (S.Watson) Sharp; Mexico; —; *Panero 2257*; TEX; —; —; —; EU440117; —. *Ratibida pinnata* Barnh.; Tennessee; —; *Urbatsch s.n.*; LSU; —; —; —; EU586880; —. *Rudbeckia alpicola* Piper; Washington; —; *Urbatsch s.n.*; LSU; —; —; —; EU586881; —. *Rudbeckia grandiflora* (Sweet) C.C.Gmel. ex DC.; Louisiana; —; *Urbatsch s.n.*; LSU; —; —; —; EU586877; EU440299, EU440307. *Rudbeckia laciniata* L.; Texas; —; *Urbatsch s.n.*; LSU; —; —; —; EU440118; —. *Rudbeckia nitida* Nutt.; Florida; —; *Urbatsch s.n.*; LSU; —; —; —; EU586879; —. *Rudbeckia triloba* L.; Virginia; —; *Urbatsch s.n.*; LSU; —; —; —; EU440142; —.
- Salmea oligocephala* Hemsl.; Mexico; —; *Escobedo 1861*; TEX; —; —; —; EU586860; —. *Salmea palmeri* S.Watson; Mexico; —; *Solis 644*; TEX; —; —; —; EU586858; —. *Salmea scandens* DC.; Mexico; —; *Diaz-Luna 19824*; TEX; —; —; —; EU586861; —. *Sanvitalia angustifolia* Engelm. ex A.Gray; Mexico; —; *Chiang et al. 8820*; TEX; —; —; —; EU586841; —. *Sanvitalia fruticosa* Hemsl.; Mexico; —; *Medrano 1446*; TEX; —; —; —; EU586843; —. *Sanvitalia fruticosa*; Mexico; —; *Valient-Banueta 701*; TEX; EU423445; —; EU423703-4; EU440108; EU440301. *Sclerocarpus divaricatus* Benth. & Hook.f. ex Hemsl.; Mexico; —; *Feddema 1728*; TEX; —; —; —; EU586844; —. *Silphium compositum* Michx.; Florida; —; *Urbatsch s.n.*; LSU; —; —; —; EU440133; —. *Spilanthes nervosa* Chodat; Brazil; —; *Irwin 1532*; TEX; —; —; —; EU586857; —. *Spilanthes urens* Jacq.; Mexico; —; *Keil 15103*; TEX; —; —; —; EU586855; —.
- Wyethia angustifolia* Nutt.; California; —; *Baldwin s.n.*; UC/JEPS; —; —; —; EU440119; —.
- Zaluzania megacephala* Sch.Bip.; Mexico; —; *Patterson 6378*; TEX; —; —; —; EU586888; EU440311. *Zaluzania mollissima* A.Gray; Mexico; —; *Hendrickson 17544*; TEX; —; —; —; EU586889; EU440295, EU440317. *Zaluzania pringlei* Greenm.; Mexico; —; *Koch et al. 79121*; TEX; —; —; —; EU586890; EU440312, EU440318. *Zaluzania triloba* (Ortega) Pers.; Mexico; —; *Johnson 5857*; TEX; —; —; —; EU586891; EU440304. *Zinnia acerosa* A.Gray; Mexico; —; *Nesom 6204*; TEX; —; —; —; EU586830; —. *Zinnia acerosa*; Arizona; —; *Urbatsch 7100*; LSU; —; —; —; EU586876; —. *Zinnia flavicoma* (DC.) Olorode & A.M.Torres; Mexico; —; *Hartman & Funk 4192*; TEX; —; —; —; EU586831; —. *Zinnia grandiflora* Nutt.; Texas; —; *Sikes & Smith 647*; TEX; —; —; —; EU586829; —. *Zinnia grandiflora*; Arizona; —; *Urbatsch 7102*; LSU; —; —; —; EU440124; —. *Zinnia maritima* Kunth; Mexico; —; *Spooner et al. 2551*; TEX; —; —; —; EU586832; —. *Zinnia oligantha* I.M.Johnst.; Mexico; —; *Hendrickson 14163*; TEX; —; —; —; EU586833; EU440291, EU440316. *Zinnia peruviana* L.; Mexico; —; *Mayfield et al. s.n.*; TEX; —; —; —; EU586886; —. *Zinnia peruviana*; Bolivia; —; *Spooner 6549*; TEX; —; —; —; EU586887; EU440303.