American Journal of Botany

American Journal of Botany 98(8): 1356–1365. 2011.

A 10-gene phylogeny of *Solanum* section *Herpystichum* (Solanaceae) and a comparison of phylogenetic methods[1]

Eric J. Tepe[2–4], Frank T. Farruggia[2], and Lynn Bohs[2]

[2]Department of Biology, 257 South 1400 East, University of Utah, Salt Lake City, Utah 84112 USA; and [3]Department of Biological Sciences, 613 Rieveschl Hall, University of Cincinnati, Cincinnati, Ohio 45220 USA

- *Premise of the study: Solanum* section *Herpystichum* is a lineage that comprises both widespread and very narrowly distributed species. This study investigates the phylogenetic relationships of sect. *Herpystichum* and evaluates several phylogenetic methods for analysis of multiple sequences.
- *Methods:* Sequence data from seven nuclear (ITS, GBSSI, and five COSII) and three plastid (*psbA–trnH*, *trnT–trnF*, and *trnS–trnG*) regions were concatenated and analyzed under maximum parsimony and Bayesian criteria. In addition, we used two analytical methods that take into account differences in topologies resulting from the analyses of the individual markers: Bayesian Estimation of Species Trees (BEST) and supertree analysis.
- *Key results:* The monophyletic *Solanum* sect. *Herpystichum* was resolved with moderate support in the concatenated maximum parsimony and Bayesian analyses and the supertree analysis, and relationships within the section were well-resolved and strongly supported. The BEST topology, however, was poorly resolved. Also, because of how BEST deals with missing sequences, >25% of our accessions, including two species, had to be excluded from the analyses. Our results indicate a progenitor-descendent relationship with two species nested within the widespread *S. evolvulifolium*.
- *Conclusions:* Analytical methods that consider individual topologies are important for studies based on multiple molecular markers. On the basis of analyses in this study, BEST had the serious shortcoming that taxa with missing sequences must be removed from the analysis or they can produce spurious topologies. Supertree analysis provided a good alternative for our data by allowing the inclusion of all 10 species of sect. *Herpystichum*.

Key words: Bayesian Estimation of Species trees (BEST); concatenation; COSII markers; *Herpystichum*; matrix representation using parsimony (MRP); Solanaceae; *Solanum*; species tree; supertree.

The goal of phylogenetics is to obtain the best possible estimation of the evolutionary relationships of the taxa under study. Rapid radiations often make it difficult to identify regions of DNA that are sufficiently variable to provide resolution among species, resulting in poorly resolved phylogenies. Approaches to this problem include searching for highly variable regions of DNA (Shaw et al., 2005, 2007; Wu et al., 2006), but also using increasingly large numbers of independent markers (Levin et al., 2009; Rodríguez et al., 2009). Phylogenetic analyses based on the concatenation and analysis of numerous independent markers carries the potential risk that the markers may produce misleading results if analyzed together under the assumption that they have the same evolutionary histories. Some authors suggest that concatenation of a sufficient number of markers will overwhelm any misleading signal (e.g., Rokas et al., 2003), whereas other authors have developed methods that take the phylogenetic signal of each of the markers into account when building a consensus tree (e.g., Liu and Pearl, 2007). These methods include supertree analysis (Bininda-Emonds, 2004) and coalescent-based methods such as Bayesian Estimation of Species Trees (BEST; Liu and Pearl, 2007; Liu, 2008). We present here a case study of the pros and cons of these species tree estimation methods using the relatively young and diverse nightshade genus *Solanum* L., a hugely economically important group that includes the potato (*S. tuberosum* L.), tomato (*S. lycopersicum* L.), and eggplant (*S. melongena* L.). Our phylogenetic study focuses on *Solanum* section *Herpystichum* Bitter, a neotropical group of 10 species of ground-trailing and climbing vines. Many species of sect. *Herpystichum* are very narrowly distributed, relatively inconspicuous, and rare in the habitats where they occur. Consequently, they are among the least collected and most poorly known species of *Solanum*.

The explosive radiation of the ca. 1500 species of the genus *Solanum* appears to have occurred within the last ca. 18 million years (Paape et al., 2008). Molecular data identify 12 to 15 major clades within *Solanum*, one of which is known as the potato clade (Bohs, 2005; Weese and Bohs, 2007). With ca. 200 species, the potato clade is a relatively species rich lineage in *Solanum*, with many taxa restricted to small areas in Central America and Andean South America. The potato clade comprises five strongly supported subclades corresponding to *Solanum* sections *Anarrhichomenum* Bitter, *Basarthrum* (Bitter) Bitter,

[1]Manuscript received 19 December 2010; revision accepted 27 April 2011.

The authors thank D. Neill, W. Quizhpe, and the staff at QCNE for assistance with permits and logistics in Ecuador; K. Coral and H. Moya for access to the Estación Biológica Limoncocha; H. Greeny and the staff at Yanayacu Biological Station; F. Villao, the Fundación Wong, and the staff at the Centro Científico Río Palenque; C. Aulestia and the staff at Bilsa Biological Station; S. Stern and F. Villao for assistance in the field; F. Guzmán Díaz, S. Stern, A. Van Wagenen, and A. Crowl for help in the laboratory; K. Lee for maintaining living plants in the greenhouse; F. Rodríguez for the COSII primers; the curators at MO, NY, and US for permission to extract DNA from herbarium specimens; and H. Farrington, J. Howard, T. Sarkinen, and two anonymous reviewers for comments on earlier versions of this manuscript. Most analyses were carried out on the freely available Bioportal (http://www.bioportal.uio.no). This study was supported by NSF through the PBI: *Solanum* grant DEB-0316614 to L.B.

[4]Author for correspondence (e-mail: eric.tepe@gmail.com)

doi:10.3732/ajb.1000516

Herpystichum Bitter, *Pteroidea* Dunal, and a large clade containing sections *Etuberosum* (Bukasov & Kameraz) A. Child, *Juglandifolia* (Rydberg) A. Child, *Lycopersicoides* (A. Child) Perlata, *Lycopersicon* (Miller) Wettstein, and *Petota* Dumort (Spooner et al., 1993; Peralta et al., 2008; E. J. Tepe and L. Bohs, unpublished manuscript). Although the potato clade is strongly supported by molecular data, diagnostic morphological characters for the group are not apparent; however, the group tends to have well-developed compound leaves and a viny habit in some subclades. One lineage within the potato clade corresponds to *Solanum* sect. *Herpystichum*, comprising 10 species of node-rooting vines of wet rainforest and cloud forest habitats (Tepe and Bohs, 2011). Species of the section share characteristically pointed buds, tendencies toward oblique leaf bases and flattened fruits.

Species of sect. *Herpystichum* range from southern Mexico to northern Peru, with their center of diversity in and near the Ecuadorian Andes where all but three species are native. The eastern foothills and Pacific lowlands of Ecuador contain the endemic species, *Solanum loxophyllum* Bitter and *S. pacificum* Tepe, while *S. limoncochaense* Tepe is endemic to the lowlands of eastern Ecuador. *Solanum dolichorhachis* Bitter is also found in the lowlands on both sides of the Andes in Ecuador with the eastern populations extending southward into Peru. *Solanum crassinervium* Tepe is restricted to the coastal hills of northwestern Ecuador and adjacent Colombia. In montane habitats, *S. trifolium* Dunal is a narrow endemic in the Ecuadorian Andes, *S. dalibardiforme* Bitter is endemic to central Colombia, and *S. pentaphyllum* Bitter is found in both Colombia and Venezuela. The remaining two species of the section are widespread. *Solanum evolvulifolium* Greenm. occurs from Costa Rica to Venezuela and Peru, and *S. phaseoloides* Pol. ranges from Mexico to Panama. All species are found in largely undisturbed, very wet habitats, but several species are occasionally found in roadsides, open fields, or other disturbed habitats.

Morphologically, members of sect. *Herpystichum* fall into two groups (Tepe and Bohs, 2011). The more distinct and easily recognized of these groups comprises herbaceous ground-trailing vines and includes *S. dalibardiforme*, *S. limoncochaense*, *S. phaseoloides*, *S. pentaphyllum*, and *S. trifolium*. Species in this ground-trailing group have simple or pinnately compound leaves with three or five leaflets, held on long petioles (>3 cm). Four of these species have markedly flattened fruits that rest on the ground or are pushed into the ground at maturity (Fig. 1A–C; Tepe and Bohs, 2009, 2011). The fruits of *S. dalibardiforme* remain insufficiently known. The second group comprises herbaceous to woody climbing vines and includes *S. crassinervium*, *S. dolichorhachis*, *S. evolvulifolium*, *S. loxophyllum*, and *S. pacificum*. Species in this climbing group have simple leaves on short petioles (<1.5 cm). The fruits of these species tend to be somewhat flattened in cross section, but to a lesser degree than they are in the ground-trailing group (Fig. 1D–F). *Solanum* sect. *Herpystichum* has been one of the most poorly collected and poorly known groups within *Solanum*; thus, the motivation of this study was to examine the relationships among the species of the section using molecular data to complement the revision by Tepe and Bohs (2011).

The sequence markers derived from the nucleus, GBSSI (*waxy*) and nrDNA ITS (the internal transcribed spacers), have been combined with the chloroplast *trnT–trnF* and *trnS–trnG* regions in a number of phylogenetic studies within *Solanum* (Bohs, 2004, 2007; Levin et al., 2005, 2006; Weese and Bohs, 2007, 2010; Stern et al., 2010). Often these markers do not provide sufficient resolving power for relationships in all parts of the tree (e.g., Levin et al., 2006; Bohs, 2007). Also, typical of large recently derived groups, identifying molecular markers with sufficient variation to resolve species relationships within the potato clade continues to be a challenge (Spooner, 2009). Fortunately, additional more variable molecular markers are continuously being discovered, and among these are COSII markers (Wu et al., 2006; Rodríguez et al., 2009). Wu et al. (2006) designed ca. 3000 primer pairs for COSII markers for the asterid clade, and several of these have been used successfully in studies of closely related species within the Solanaceae (Levin et al., 2009; Rodríguez et al., 2009; Tepe and Bohs, 2010). Within *Solanum*, these have been found to be single-copy, orthologous nuclear markers that range from conserved to highly variable (Rodríguez et al., 2009). The analyses presented here combined the traditionally used markers ITS, GBSSI, *trnT–trnF*, *trnS–trnG*, and *psbA–trnH* with five COSII markers.

We also used our 10-marker data set to evaluate various methods for the analysis of multiple marker sequence data. Increasing the number of markers from different genomes increases the possibility of errors in phylogeny estimation resulting from differing evolutionary histories of the markers. In this study, we used concatenation, a coalescent Bayesian-based method, and supertree analysis to explore our data set, and we discuss the pros and cons of each.

MATERIALS AND METHODS

Taxon sampling—We sampled 24 accessions of *Solanum* (Appendix 1) for the 10 markers used in this study. This sample includes all 10 species of *Solanum* sect. *Herpystichum* (Tepe and Bohs, 2011), and representatives from each of the subclades that make up the potato clade (Spooner et al., 1993; Weese and Bohs, 2007; E. J. Tepe and L. Bohs, unpublished manuscript): sections *Anarrhichomenum* (*S. brevifolium* Dunal), *Basarthrum* (*S. caripense* Dunal), *Lycopersicon* (*S. lycopersicum*), *Petota* (*S. bulbocastanum* Dunal), and *Pteroidea* (*S. anceps* Ruiz & Pav.). We included multiple accessions of species when available and, in the case of *S. evolvulifolium*, included accessions from across the species' range.

Molecular methods and phylogenetic analysis—Leaf material for most accessions was field-collected in silica gel for DNA extraction. DNA samples for five accessions were isolated from herbarium specimens. DNA from silica gel-dried material was extracted with the DNeasy plant mini extraction kit (Qiagen, Valencia, California, USA) following the manufacturer's protocol. For extractions from herbarium specimens or silica gel-dried samples that did not amplify using the DNeasy kit, we used a modified extraction protocol in which a 2× CTAB extraction buffer and a 24-h incubation was used instead of Qiagen's AP1 buffer and 10-min incubation (Green et al., 1999).

PCR amplification followed the procedures described in White et al. (1990) and Vargas et al. (1998) for ITS; Taberlet et al. (1991) for *trnT–trnF*; Levin et al. (2005) for GBSSI; Hamilton (1999) and Levin et al. (2005) for *trnS–trnG*; and Sang et al. (1997) for *psbA–trnH*. The COSII markers (Wu et al., 2006; SOL Genomics Network [http://www.sgn.cornell.edu]) used in this study were specifically selected because they contributed to well-resolved phylogenies in the related sections *Lycopersicon* and *Petota* (Rodríguez et al., 2009) and *Pteroidea* (Tepe and Bohs, 2010). They were also chosen because PCR amplification was easy, they produced single-banded PCR products in most accessions, and they resulted in clean sequences without the need for cloning. In the rare case in which PCR resulted in more than one band, gel isolation and direct sequencing of the target band size resulted in a clean sequence. Primer sequences and PCR conditions are available in Appendices S1 and S2 (see Supplemental Data with the online version of this article). PCR products were cleaned with the Promega Wizard SV Gel and PCR Clean-up system (Promega, Madison, Wisconsin, USA) and sequenced on an ABI automated DNA sequencer at the University of Utah Core Facilities. We sequenced overlapping

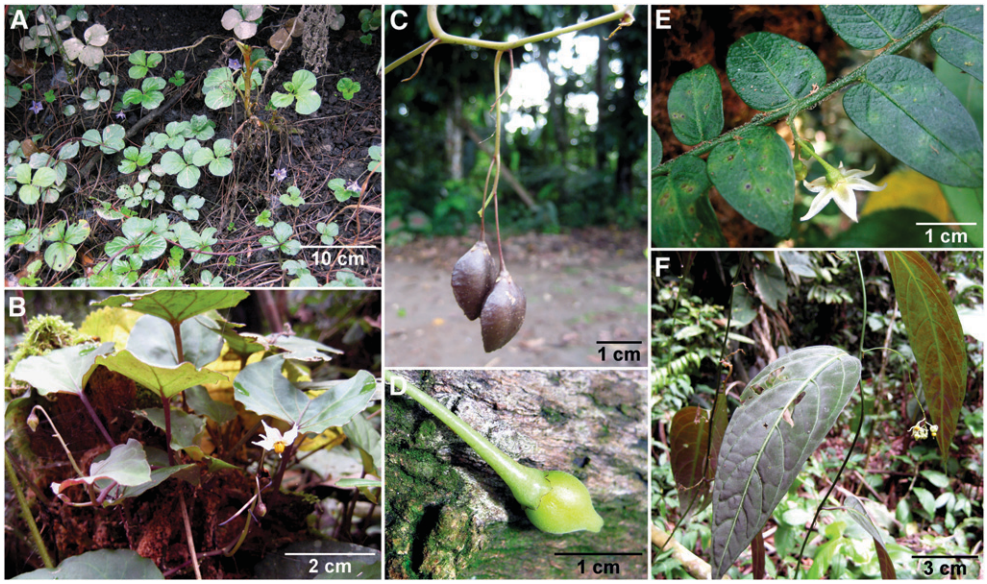

Fig. 1. Representatives of *Solanum* sect. *Herpystichum*: (A–C) ground-trailing species, (D–F) climbing species. (A) *S. trifolium*, habit. (B) *S. limoncochaense,* habit, bud, and flower. (C) *S. limoncochaense*, the flattened fruits of this species are typical of the ground-trailing species. (D) *S. pacificum*, the fruits of the climbing species are typically pointed, but only slightly flattened. (E) *S. evolvulifolium*, habit and flower. (F) *S. pacificum*, habit and inflorescence.

forward and reverse sequences for all samples and assembled and proofread the contigs with the program Sequencher 4.8 (GeneCodes, Ann Arbor, Michigan, USA). We used standard nucleotide ambiguity codes to identify all instances where more than one peak was apparent in the chromatogram. Dubious sequences at the extreme 3′ and 5′ ends of reads were excluded from the analyses. We manually aligned sequences with the program Se-Al v.2.0a11 (Rambaut, 1996). GenBank accession numbers of the DNA sequences used in this study are presented in the Appendix 1. Aligned data sets are available through TreeBASE (http://purl.org/phylo/treebase/phylows/study/TB2:S11112).

Phylogenetic relationships were estimated under Bayesian and maximum parsimony optimality criteria. Maximum parsimony (MP) analyses were performed on each data set separately and on the concatenated matrices using the program PAUP* version 4.0b10 (Swofford, 2003) with all characters weighted equally and gaps treated as missing data in full heuristic analyses with 100 random addition sequence replicates, tree-bisection-reconnection (TBR) swapping, steepest descent, and all other settings kept as the defaults. Bootstrap (BP) values for nodes were estimated from full heuristic searches of 5000 replicates with MaxTrees set at 10 000 and TBR branch swapping.

Bayesian inference (BI) of individual markers and partitioned concatenated data sets was performed using the program MrBayes 3.1.2 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003), with substitution model parameters determined with the program MrModeltest 2.2 (Nylander, 2004) (Table 1). We chose the models estimated under the Akaike information criterion (AIC) because this method has been shown to perform better than the hierarchical likelihood ratio test when comparing nested models (Posada and Buckley, 2004). Using random starting trees, we ran MrBayes for 10 000 000 generations, with one tree sampled every 1000 generations. We ran two independent, parallel runs with 10 chains each using the parallel version of MrBayes 3.1.2 on the freely available Bioportal cluster (http://www.bioportal.uio.no) with all other settings as the defaults. Post analysis was carried out in the serial version of MrBayes 3.1.2 to determine the number of trees to omit as burn-in and to compute the consensus tree and posterior probabilities (PP).

In addition to the partitioned concatenated analyses, we also ran supertree and Bayesian estimation of species trees (BEST) (Liu and Pearl, 2007; Liu, 2008) analyses. For the supertree, we employed the widely used matrix representation using parsimony analysis method (MRP) (Bininda-Emonds, 2004), which allows for the combination of trees that individually may have limited taxon representation. The method uses phylogenetic relationships (e.g., trees based on morphological and/or molecular characters) to construct a new binary matrix of species relationships based on their presence or absence (character states 0 or 1) at each node (characters) on the input topologies. For input trees, we used the 95% majority rule consensus trees from the BI postburn-in distribution for each of the individual molecular markers (Appendix S3, see online Supplemental Data). The software Mesquite version 2.6 (Maddison and Maddison, 2009) was used to combine each of the new matrices into a new supermatrix. The MRP method requires at least three taxa to be overlapping in any two combined data sets (Bininda-Emonds, 2004). In this study, a minimum of 12 taxa were overlapping between the data sets, including the outgroups. Phylogenetic analysis of the new matrix followed the MP protocol outlined already.

BEST analysis works by estimating phylogenies for all 10 molecular markers separately while simultaneously constructing the "species tree." For each marker, the model of sequence evolution and the estimated model parameters output from Modeltest were used as priors in the analysis. Using MrBayes, we carried out two runs with four chains each to 10 000 000 generations and sampled every 1000 generations. The distribution of the species tree over the separate gene trees was summarized using the BEST software (Liu et al., 2003). One problem with Bayesian analyses (including BEST) is that samples with no sequence data are placed in trees based only on the prior distribution rather than sequence data (Liu et al., 2003). Thus, to avoid possible artifacts caused by missing sequences, those accessions that were missing one or more of the 10 markers were removed from all data sets prior to BEST analysis. Additionally, each individual gene tree distribution was compared to the BI results for each respective region to observe any differences between the two methods.

Prior to concatenating the individual data sets, we ran a partition homogeneity test (PHT) test (Farris et al., 1994) to test for incongruence. The test was performed in PAUP* and implemented using 100 replicates with 10 random addition sequences per replicate and rearrangements limited to 1 000 000 per replicate. Because the PHT test is known to suffer from type I errors when phylogenetic signal is low (Dolphin et al., 2000; Yoder et al., 2001; Darlu and Lecointre, 2002; Dowton and Austin, 2002; Hipp et al., 2004), we compared the topologies of individual markers to each other and to the concatenated and BEST results to identify the presence of well-supported incongruence (i.e., differences supported by high bootstrap values and/or posterior probabilities; Seelanan et al., 1997; Wiens, 1998). Throughout this paper, we conservatively consider well-supported nodes to have both PP ≥ 0.95 and BP ≥ 90. We considered both measures of support together because PP values are often inflated relative to BP (Cummings et al., 2003; Erixon et al., 2003; Simmons et al., 2004).

To test the hypothesis that the ground-trailing *(S. dalibardiforme*, *S. limoncochaense*, *S. phaseoloides, S. pentaphyllum*, *and S. trifolium)* and climbing species *(S. crassinervium*, *S. dolichorhachis, S. evolvulifolium*, *S. loxophyllum*, *and S. pacificum)* formed monophyletic groups, we constructed a tree in which species forming each of the two groups were constrained to monophyly (without specified internal topology) and sister to each other. We used the Shimodaira–Hasegawa test (SH; Shimodaira and Hasegawa, 1999) option in PAUP* to

compare the constrained MP tree to the unconstrained MP tree, the BI consensus tree, and 20 randomly chosen, postburn-in BI trees. The one-tailed SH analysis was run using RELL bootstrap with 1000 pseudoreplicates and the same model parameters as earlier.

RESULTS

The PCR was successful for most markers of most accessions, and *trnT–trnF* and *trnS–trnG* were amplified and sequenced cleanly for all accessions. We were unable to obtain PCR products for one or more accessions of all other markers (Appendix 1). PCR amplification of several accessions produced two bands including cos10B (*S. evolvulifolium* 2671 and 7198, *S. loxophyllum* 2726, and *S. phaseoloides* 3499), GBSSI (primers waxyF–1171R, *S. limoncochaense* 2627), *trnS–trnG* (*S. trifolium* 2682), and *trnT–trnF* (primers tabE–tabF, *S. dalibardiforme*). In all cases, the target bands were much brighter and resulted in clean sequences once gel-isolated. Of the five DNA accessions extracted from herbarium specimens, we achieved a 62% success rate of amplification for the 10 markers used in this study. Success ranged from 100% for *S. dalibardiforme* (collected in 2003) to 30% in *S. dolichorhachis* (collected in 1934). Successful amplification was roughly correlated with the age of the specimen; the most recent collection, *S. evolvulifolium* collected in 2004, amplified for only 80% of the markers, indicating that age of the specimen is only one of several factors affecting the successful sequencing of herbarium material. The extraction protocol that yielded the highest quality DNA from herbarium material used ca. 1 cm^2 of leaf material from the packet (i.e., not previously glued) and the 24-h incubation in 2× CTAB combined with the DNeasy Plant mini kit (determined by trial and error on specimens collected by the senior author for this purpose). Overall, our sequence coverage was 88% for the ingroup, and 91% for all accessions.

Sequence variation in the COSII markers was not strongly correlated to either sequence length or to percent intronic content (Table 1). Of all the nuclear derived sequence regions, ITS had the highest number and percentage of PI characters, but also the highest measure of homoplasy. All the nuclear markers had higher numbers and higher percentages of PI characters than any of the plastid markers. In fact, ITS, GBSSI, and two of the COSII (cos5 and cos9B) markers each had an equal or greater number of PI characters than all three chloroplast markers combined. Nevertheless, the plastid markers were useful in estimating phylogenies for sect. *Herpystichum* because they dramatically increased support for many nodes. Similarly, although the COSII markers combined produced a topology very similar to that produced by the 10-gene data set, support for many nodes was low.

The PHT test comparing all 10 matrices to each other indicated significant conflict among the data sets ($P = 0.01$). Comparisons of the combined nuclear vs. chloroplast data sets were also significantly different ($P = 0.02$). If the PHT test truly measures incongruence among data sets, markers with the same evolutionary history should not differ significantly; however, for a control, we compared the individual chloroplast markers to each other, and they also differed significantly from each other ($P = 0.01$). Comparison of the topologies from the individual markers (online Appendix S3) to the 10-marker combined topologies (Figs. 2, 3) revealed only five well-supported differences. Three of these differences were in cos1C: two involving the relationships of several accessions of *S. evolvulifolium* and *S. loxophyllum*, and one supporting the two accessions of *S. anceps* from sect. *Pteroidea* as sister to the climbing species of sect. *Herpystichum*. Two more supported differences were in cos10B in which *S. anceps* and *S. pentaphyllum* + *S. phaseoloides* formed a grade sister to the climbing species. Because so few well-supported differences were found, we decided to use several methods to combine the data to more fully explore its potential.

The results of our 10-gene concatenated and supertree analyses returned a monophyletic sect. *Herpystichum* (Figs. 2, 3A).

TABLE 1. Summary of sequence data for the data sets analyzed.

Marker	*N*	Align.	No. PI	% PI	% intron	MPT	L	CI	RI	Model	Burn-in	No. nodes resolved	No. nodes supported
GBSSI	21	2233	121	5.4	43 [a]	168	173	0.81	0.88	GTR+G	15 000	14/11	8/3
ITS	21	686	127	18.5	—	9	292	0.64	0.76	GTR+G	5000	12/12	7/7
cos5 [b]	19	1030	101	9.8	77.8	132	140	0.81	0.86	HKY+I	10 000	10/11	8/6
cos11	23	697	94	13.5	80.6	5	151	0.73	0.83	HKY+I	10 000	10/10	10/5
cos9B	21	918	100	10.9	83.3	120	144	0.79	0.87	GTR+G	10 000	10/10	7/4
cos10B	22	716	81	11.3	81.2	96	131	0.74	0.85	HKY+G	10 000	12/13	8/6
cos1C	23	731	87	11.9	71.0	3	123	0.83	0.90	GTR+G	10 000	13/14	10/8
trnT-trnF	24	1736	42	2.4	—	67	51	0.84	0.92	GTR+G	10 000	11/9	6/2
trnS-trnG	24	687	27	3.9	—	15	34	0.85	0.92	GTR+I	10 000	8/9	5/4
psbA-trnH	20	584	31	5.3	—	1	48	0.69	0.86	GTR+I+G	10 000	11/11	4/2
Plastid combined	24	3007	100	3.3	—	1	1249	0.69	0.79	Partitioned [c]	15 000	16/16	11/8
Nuclear combined	24	7011	710	10.1	—	102	149	0.70	0.84	Partitioned	5000	13/13	10/5
Non-COSII combined	24	5926	348	5.9	—	2	631	0.69	0.80	Partitioned	10 000	15/16	9/9
COSII combined	24	4092	462	11.3	—	42	757	0.71	0.80	Partitioned	10 000	16/15	9/5
All combined	24	10 018	810	8.1	—	1	1407	0.69	0.79	Partitioned	10 000	16/16/14	16/10/7

Notes: N = number of included accessions; Align. = aligned sequence length; No. PI = number of parsimony informative characters; % PI = percentage of total characters that are parsimony informative; % intron = percentage of intronic content for COSII markers. MPT = No. of most parsimonious trees; L = length of MPT; CI = consistency index, RI = retention index; Model = model of sequence evolution as determined by MrModeltest 2.2 (Nylander, 2004); Burn-in = number of generations excluded as burn-in from Bayesian analyses when computing the 50% majority rule trees; No. nodes resolved = number of nodes resolved in Bayesian/maximum parsimony/supertree analyses; No. nodes supported = number of ingroup nodes supported by posterior probability ≥ 0.95 and bootstrap ≥ 85) in Bayesian/maximum parsimony/supertree analyses. MPT, CI, and RI are reported with uninformative sites excluded.

[a] Estimation from Levin et al. (2009)

[b] Names used for COSII markers throughout this paper were taken from Rodríguez et al. (2009).

[c] Models for each individual marker were maintained in the partitioned analyses.

The concatenated analyses strongly supported monophyly (1.0 PP, 98 BP), whereas the supertree resolved the group, but with very low support (58 BP). Monophyly of sect. *Herpystichum* was not supported in the BEST tree, in which clades of sect. *Herpystichum* formed a polytomy with accessions of sect. *Pteroidea*. The concatenated (MP and BI) and supertree analyses also provided strong support for sect. *Pteroidea* as sister to sect. *Herpystichum* (1.0 PP, 98–100 BP). All species for which we included multiple accessions were monophyletic except for *S. evolvulifolium*, which was paraphyletic in all analyses; *S. crassinervium* and *S. loxophyllum* were nested within accessions of *S. evolvulifolium*. Two of the Ecuadorian accessions of *S. evolvulifolium* formed a clade, but the third Ecuadorian accession was sister to the two Central American accessions. *Solanum pentaphyllum* was nested within *S. phaseoloides* in the concatenated MP analysis with very low support (59 BP), but these species are sister to each other in the concatenated BI analysis (1 PP, 85 BP). The supertree places *S. pentaphyllum* within a paraphyletic *S. phaseoloides*, but the relationships among these accessions are unsupported.

Topologies of the BI and MP analyses of the concatenated data set were highly congruent (Fig. 2). The trees differed in only two minor areas, the relationships among accessions of *P. pentaphyllum* and *P. phaseoloides* mentioned earlier, and the placement of the two accessions of *S. trifolium* as sister to *S. limoncochaense* in the BI tree (i.e., Fig. 2) vs. sister to *S. pentaphyllum* + *S. phaseoloides* in the MP tree. The nodes supporting these differences have low BP scores (57 and 59, respectively, on the MP trees) and strong PP scores (0.99 in both cases) on the BI trees. Similarly, the supertree is highly congruent with the BI and MP concatenated analyses, but differed in resolution and in the placement of *S. dolichorhachis* and *S. pacificum* as a grade rather than a clade (Fig. 3A). The BEST topology is similar to those of the concatenated and supertree analyses, but with much lower resolution (Fig. 3B). Also, five accessions, including the sole representatives for two species (*S. dolichorhachis*

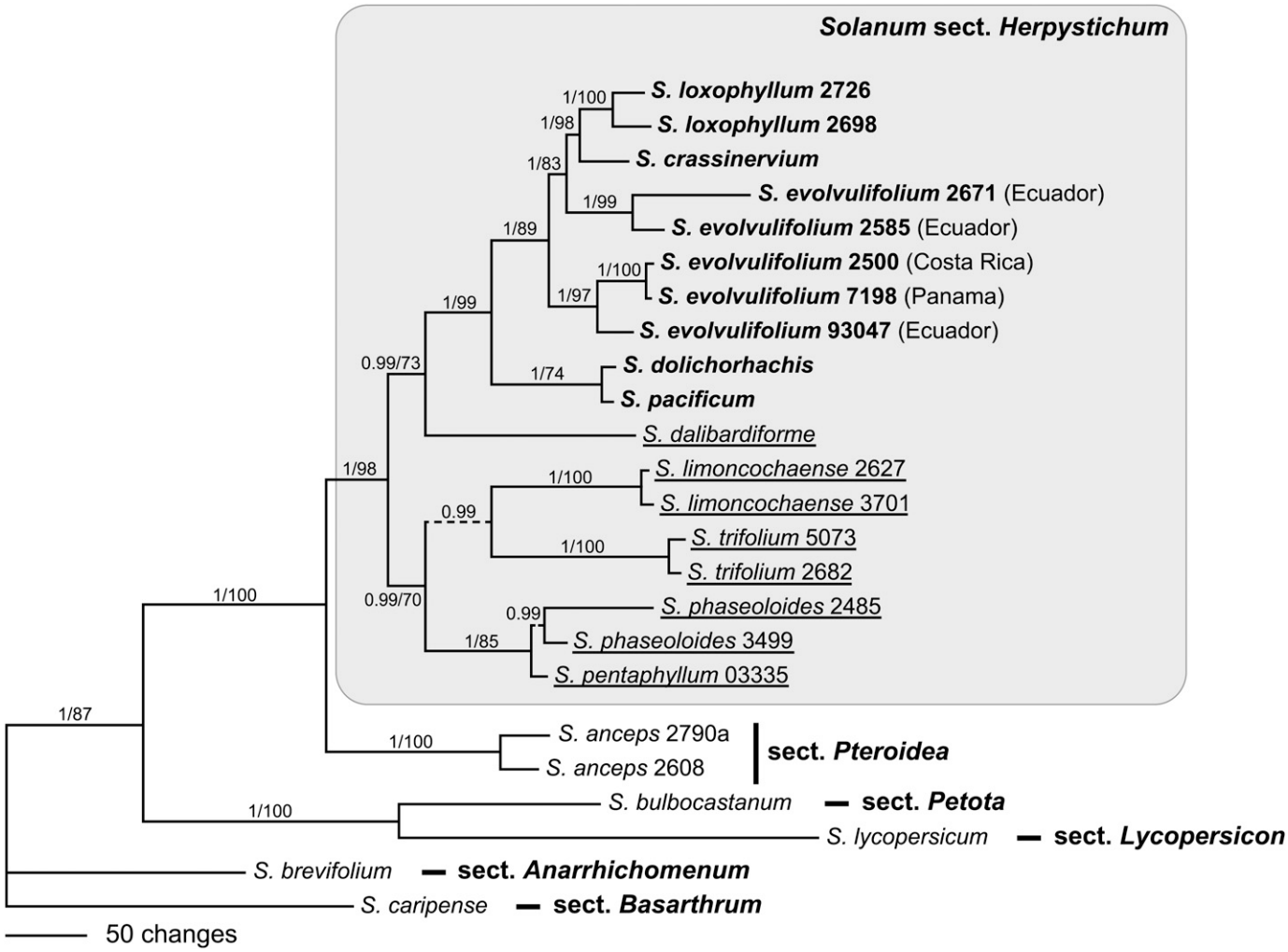

Fig. 2. The 50% majority-rule postburn-in consensus tree from Bayesian analysis of a concatenated data set of GBSSI, ITS, cos5, cos11, cos9B, cos10B, cos1C, *psbA–trnH, trnS–trnG*, and *trnT–trnF* sequences. Branch support values are Bayesian posterior probabilities ≥0.5/maximum parsimony bootstrap ≥50. The dashed branches were present in the Bayesian analysis, but not in the strict consensus tree of the maximum parsimony analysis. Species of *Solanum* sect. *Herpystichum* are indicated by the gray box. Multiple accessions of a species are differentiated by the collection number available in Appendix 1. The geographic sources are given for the accessions of *S. evolvulifolium* since they come from throughout the range of this widespread species. Species with a climbing habit are in bold, and the ground-trailing species are underlined.

and *S. pentaphyllum*), had to be eliminated from the BEST analysis because they were missing sequence data for one or more markers. The BEST tree, however, corroborates the topology recovered in all other analyses (although with limited resolution and poor support), especially the relationships among *S. crassinervium*, *S. evolvulifolium*, and *S. loxophyllum*. Differences include the weakly supported placement of *S. phaseoloides* and *S. dalibardiforme* as sister to the climbing species and the equivocal monophyly of sect. *Herpystichum*.

All analyses agree on relationships among the climbing species: *S. dolichorhachis* and *S. pacificum* sister to a clade containing *S. crassinervium* and *S. loxophyllum*, which are nested within *S. evolvulifolium*. The relationships among the ground-trailing species are less clear. The concatenated analyses suggest that *S. limoncochaense*, *S. pentaphyllum*, *S. phaseoloides*, and *S. trifolium* form a monophyletic group, with relationships among the species differing slightly between the BI and MP analyses (Fig. 2). The supertree resolves a clade that includes the ground-trailing species except for *S. dalibardiforme*, but this relationship is poorly supported (Fig. 3A). The BEST analysis placed *S. phaseoloides* in an unresolved position along with *S. dalibardiforme* sister to the climbing species clade (Fig. 3B). *Solanum phaseoloides* was not placed near the clade with the climbing species in any other analysis. Constraint trees forcing the monophyly of the climbing and ground-trailing groups were significantly different from unconstrained trees under the SH test using the 10-gene concatenated data set (−ln likelihood = 28 669.88409, difference from best tree = 36.91763, $P = 0.014$; constraint trees were 1420 vs. 1407 steps long in the best, unconstrained tree) indicating that the nonmonophyly of the ground-trailing species (i.e., *S. dalibardiforme* as sister to the climbing species rather than with the other ground-trailing species) is likely not an artifact of the analyses.

Topologies of the individual markers were largely congruent with each other, despite the results of the PHT test, and BI and MP analyses produced nearly identical results (online Appendix S3). Section *Herpystichum* was supported as monophyletic in cos9B (1.0 PP, 100 BP), but accessions of sect. *Pteroidea* (i.e., *S. anceps*) were in a polytomy with species of sect. *Herpystichum* in trees resulting from four other markers (cos5, ITS, GBSSI, and *psbA–trnH*) and were nested within sect. *Herpystichum* in analyses of five markers (cos1C, cos10B, cos11, *trnS–trnG*, and *trnT–trnF*). With the exception of cos11 and *psbA–trnH*, all of the individual markers strongly support the monophyly of the climbing species. In cos11, *S. pentaphyllum* and *S. phaseoloides* are in a polytomy with the climbing species and *S. trifolium* is sister to the outgroup members *S. bulbocastanum* and *S. lycopersicum*. In *psbA–trnH*, *S. dalibardiforme*, a ground-trailing species, is nested among the climbing species. *Solanum dalibardiforme* is sister to the climbing species in trees derived from two markers (ITS and GBSSI), nested within the climbing species in one marker (*psbA–trnH*), sister to all other species of sects. *Herpystichum* and *Pteroidea* in two markers (cos1C and cos10B), and in various other positions in the other markers. The remaining ground-trailing species were placed largely in a basal polytomy by ITS, GBSSI, cos5, and *trnT–trnF*, and in various places by the other markers.

Similarly, the topologies produced by concatenated analyses of the chloroplast and nuclear markers separately were largely comparable (Appendix S4, see online Supplemental Data). They were not, however, identical to each other, and neither tree

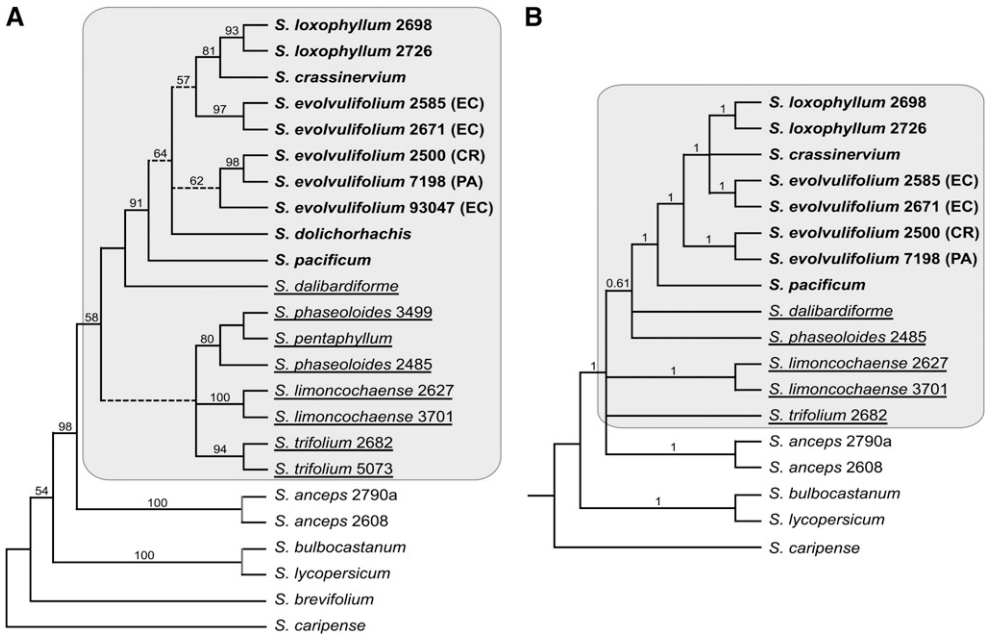

Fig. 3. Phylogenetic relationships of *Solanum* sect. *Herpystichum* from (A) supertree analysis and (B) a coalescent approach using BEST analysis, based on analyses of the 10 markers (GBSSI, ITS, cos5, cos11, cos9B, cos10B, cos1C, *psbA–trnH, trnS–trnG*, and *trnT–trnF*). (A) The 50% majority rule consensus tree from 28 equally most parsimonious supertrees based on the topologies resulting from BI analyses of the individual markers. Dashed branches are not present in the strict consensus tree. Support values above the branches are MP bootstrap >50. (B) The 50% majority-rule postburn-in consensus tree from analysis of a reduced data set using BEST. The data set included only accessions for which we had sequence data for all markers. Bayesian posterior probabilities ≥0.5 are shown above the nodes. Multiple accessions of a species are differentiated by the collection number available in the Appendix 1. Species of sect. *Herpystichum* are indicated by the gray box. Species with a climbing habit are in bold and the ground-trailing species are underlined. The geographic origin for *S. evolvulifolium* accessions in parentheses; abbreviations are CR = Costa Rica, EC = Ecuador, PA = Panama.

matched the tree in Fig. 2. Nevertheless, none of the differences between the two trees compared to each other or to the combined tree in Fig. 2 were supported (i.e., ≥0.95 PP, ≥90 BP).

DISCUSSION

The samples included in this study represent all 10 species of *Solanum* sect. *Herpystichum* in the recent revision by Tepe and Bohs (2011) and is the first phylogenetic study of this previously poorly known and undercollected group of *Solanum*. In this study, sect. *Herpystichum* was supported as monophyletic, as were each of the species for which we included multiple accessions, except for *S. evolvulifolium*. *Solanum* sect. *Pteroidea*, here represented by *S. anceps*, was supported as sister to sect. *Herpystichum* in the 10-gene concatenated and supertree analyses. The 10-gene BEST analysis proved problematic for our data set because we had to eliminate a number of accessions and species from the analyses due to missing data.

Phylogenetic relationships—The BI and MP analyses of the concatenated 10-gene data set provided strong support for the monophyly of sect. *Herpystichum*. The supertree also resolved the section as monophyletic, but with weak support, and its monophyly was unresolved in the BEST analysis. Our data also support a close relationship between sect. *Herpystichum* and *S. anceps* of sect. *Pteroidea*. Section *Pteroidea* is sister to sect. *Herpystichum* in the BI and MP analyses of the concatenated 10-gene data set and in the supertree analysis, but unresolved in the BEST tree. Several of the markers analyzed individually, however, resolve *S. anceps* as nested within sect. *Herpystichum*. This placement could be an artifact, but could also indicate a real relationship. In fact, different preliminary analyses of an incomplete data set with greater sampling across the entire potato clade variously indicate sect. *Herpystichum* as a monophyletic group sister to sect. *Pteroidea* as found in this study, as two monophyletic clades in a polytomy with sect. *Pteroidea*, or as a paraphyletic grade with respect to a nested, monophyletic sect. *Pteroidea* (E. J. Tepe, unpublished data). Despite the ambiguities in their phylogenetic placement, it is clear that the two sections are more closely related to each other than to any other group within the potato clade.

Within sect. *Herpystichum*, our complete data sets (i.e., concatenated and supertree analyses) reveal two subclades: one comprising the climbing species plus *S. dalibardiforme* and one with the ground-trailing species excluding *S. dalibardiforme*. *Solanum dalibardiforme* is resolved as sister to the climbing species in the 10-gene concatenated and supertree analyses, but with moderate to weak support. Despite the low support for this relationship, a SH test indicated that trees in which *S. dalibardiforme* was forced into a clade with the other morphologically similar ground-trailing species were significantly different from the trees in Figs. 2 and 3A. The BEST analysis places *S. dalibardiforme* in a polytomy with *S. phaseoloides* and the climbing species, but not with any of the other ground-trailing species (Fig. 3B). Like the climbing species, *S. dalibardiforme* has simple leaves, but the corollas are rotate-stellate rather than stellate (a character shared only with *S. trifolium*; Tepe and Bohs, 2011). All of the species in the ground-trailing clade have weakly herbaceous stems, long internodes and petioles, adventitious roots at nodes and frequently at inflorescences, inflorescences that are unbranched and few-flowered, and strongly flattened, arrowhead-shaped fruits. *Solanum dalibardiforme* shares the long internodes and long petioles with this group. The fruits are somewhat pointed, but from the herbarium material it is not possible to determine whether they are strongly flattened like the ground-trailing species or only somewhat flattened like the climbing species. While the ground-trailing habit is useful for species identification, it does not appear to characterize a discrete lineage. Strongly flattened fruits may be a synapomorphy for the clade that includes *S. limoncochaense*, *S. pentaphyllum*, *S. phaseoloides*, and *S. trifolium*, but only if the fruits of *S. dalibardiforme* are not markedly flattened.

The monophyly of the climbing species is strongly to moderately supported in all 10-gene analyses. Species in this clade share a number of morphological characters including the climbing habit via adventitious roots, a tendency toward distichous leaf arrangement and asymmetrical leaf bases, and pointed fruits that are somewhat flattened perpendicularly to the septum. Despite this list of characters, we have not yet been able to identify a synapomorphy that is unambiguously present in all of the climbing species. Within this clade, *S. dolichorhachis* and *S. pacificum* form either a grade or a clade that is sister to the clade composed of *S. crassinervium*, *S. evolvulifolium*, and *S. loxophyllum*. In the latter clade, it appears that *S. crassinervium* and *S. loxophyllum*, both from lowland Ecuador, evolved from a South American element within the widespread, mid- to high-elevation *S. evolvulifolium*. This pattern results in two monophyletic species and one paraphyletic species. In fact, the nested pattern found among these species is a clear example of a progenitor-derivative species relationship (Gottlieb, 2003; Crawford, 2010). The observed pattern is expected when a population "buds off" from a widespread species that remains unchanged, as in Mayr's (1942) peripatric speciation model. Grant (1981) built on this idea in what he called "quantum speciation" by including the observation that many derivative species appear to experience accelerated morphological evolution relative to the progenitor species. In *Solanum*, *S. crassinervium* and *S. loxophyllum* have indeed appeared to undergo accelerated morphological evolution relative to the widespread *S. evolvulifolium*, which varies little across its range from southern Central through northwestern South America. The accelerated changes observed in these two species are likely due to founder effect and genetic drift associated with small initial populations. Also, the lowlands of western Ecuador, where these two species likely evolved, are a hotspot of diversity and speciation (Myers, 1988; Borchsenius, 1997; Myers et al., 2000; Mutke and Barthlott, 2005), which may have contributed to a higher rate of morphological evolution. We fully agree with other authors that a species concept that requires monophyly is an oversimplification and not biologically realistic (Rieseberg et al., 1990; Rieseberg, 1991; Levin, 1993; Rieseberg and Brouillet, 1994; de Queiroz, 2007; Knowles and Carstens, 2007; Hörandl and Stuessy, 2010), and numerous examples of nested patterns are known from angiosperms (e.g., Baldwin, 2005; Koopman and Baum, 2010; reviewed in Gottlieb, 2003; Crawford, 2010). Progenitor-derivative species groups provide opportunities to study plant speciation in the early stages, especially with respect to barriers to gene flow (Crawford, 2010), and the parallel derivative species found here may present an interesting case study. Throughout its range, *S. evolvulifolium* is found mostly above 1000 m in elevation, whereas *S. loxophyllum* occurs below 850 m and *S. crassinervium* below 500 m, so the derivative species probably do not come into frequent contact with the progenitor. *Solanum crassinervium* and *S. loxophyllum*, however, do co-occur in parts of their ranges, and flower simultaneously (E. J. Tepe, personal observation). We do not have data on the crossability among these three *Solanum* species, but it is likely that these species have evolved some barriers to gene flow.

Relative strengths of analysis methods—The 10-gene concatenated analyses resulted in highly resolved, well-supported topologies that are similar, but not identical, to the 10 markers analyzed individually (but none of the individual markers are identical to each other either; Appendix S3). The topology of the 10-gene concatenated data set agrees, for the most part, with morphology and indicated close relationships among the climbing species, and among most of the ground-trailing species. Although analyses of the concatenated data set resulted in highly resolved and highly supported phylogenies, nevertheless, concatenating data poses potential problems if the various markers are products of different evolutionary histories (reviewed in Degnan and Rosenberg, 2009). The differences in the individual topologies found in this study could be artifacts of analysis due to methodological errors (e.g., long branch attraction, incorrect alignments or models of sequence evolution, undetected paralogy) or biological factors that are not the result of different evolutionary histories (e.g., homoplasy, variation in PI sites or sequence length, nonindependence of nucleotide substitutions, etc.) and are not problematic if analyzed in concatenation (Sullivan, 1996; Averof et al., 2000; Sanderson and Shaffer, 2002; Huson and Bryant, 2006; Baum, 2007; Rodríguez et al., 2009). The different topologies, however, could be due to biological factors and the result of varying genetic histories (e.g., hybridization, allopolyploidy, incomplete lineage sorting, differential modes of inheritance or rates of evolution of different parts of the genome, etc.), in which case concatenated analyses could be misleading (Mason-Gamer and Kellogg, 1996; Comes and Abbott, 2001; Anderson et al., 2006; Prohens et al., 2006; Liu and Pearl, 2007; Spooner et al., 2008; Degnan and Rosenberg, 2009; Rodríguez and Spooner, 2009).

To avoid the possible confounding effects of combining data with different evolutionary histories, we used a coalescent approach implemented in BEST and MRP supertree analysis. These both use combined data, but in ways that consider the topologies of each of the independent markers in constructing the consensus tree. BEST uses a Bayesian algorithm to simultaneously search tree space of each individual marker while concurrently constructing a "species tree" in which the independent topologies are combined (Liu, 2008). One of the biggest drawbacks of using BEST, however, is that if one or more sequences are not available for a given accession, either that accession or the entire marker must be excluded from the BEST analyses. In BEST, samples with missing data are placed seemingly randomly, based solely on the prior distribution, which can lead to potentially misleading results (Liu et al., 2003). In our data set, we were unable to obtain sequences of all 10 markers for six accessions (five from the ingroup, and one from the outgroup; see Appendix 1). Two of these accessions, *S. dolichorhachis* and *S. pentaphyllum*, were the sole representatives of their species, and their elimination from the data set represents a serious problem with BEST analyses or any other method that analyzes each marker independently prior to combining them into a consensus tree. BEST outperformed other similar methods of analysis in a recent study of simulated datasets (Leaché and Rannala, 2011); however, the inability to deal with missing sequences limits its utility in the real world where many data sets are likely to have gaps in coverage. Nevertheless, except for the placement of *S. phaseoloides* in a polytomy with *S. dalibardiforme* and the clade formed by the climbing species rather than sister to *S. pentaphyllum* (which had to be excluded from the analysis), the BEST results are highly congruent with the concatenated analyses.

All three analytical methods used in this study have advantages and disadvantages; however, because of the incomplete sequence coverage, the concatenated and supertree analyses were more suitable for this particular data set. MRP supertree analysis has the advantage that it is allows for the combination of multiple markers while maintaining the hierarchical information of the individual markers. This method builds a new data matrix based on relationships from input trees (Bininda-Emonds, 2004). In our study, the supertree is less resolved, and most nodes have lower support relative to the concatenated analyses and is thus a more conservative estimate of relationships. Like the concatenated analysis, the supertree relationships are highly concordant with the morphology of these plants, with the aforementioned possible exception of *S. dalibardiforme*. The primary disadvantage of supertrees is that the resulting consensus topology is further removed from the raw data matrix. As shown in the supertree (Fig. 3A), the resolution is relatively low, but it is fully congruent with the concatenated analyses and highly congruent with the BEST analysis. The advantage of the supertree over BEST was that it allowed us to combine incomplete data sets, without the need to remove accessions and species from the analysis, thereby fulfilling the ultimate goal of phylogenetics, allowing us obtain the best possible estimation of the evolutionary relationships of these taxa. Thus, supertree analysis represents a good substitute for BEST in cases where multiple markers may have different evolutionary histories, but where sequences are not available for all markers from all accessions. The comparison of analysis methods presented here is based on our experience with the data set presented in this study and is not intended to be a definitive evaluation of the methods; however, this study illustrates the importance of analyzing data with more than one tool and of not relying disproportionally on the latest methods until they have been employed and evaluated comparatively in a number of practical studies.

LITERATURE CITED

ANDERSON, G. J., C. MARTINE, J. PROHENS, AND F. NUEZ. 2006. *Solanum perlongistylum* and *S. catilliflorum*, new endemic Peruvian species of *Solanum*, section *Basarthrum*, are close relatives of the domesticated pepino, *S. muricatum. Novon* 16: 161–167.

AVEROF, M., A. ROKAS, K. H. WOLFE, AND P. M. SHARP. 2000. Evidence for a high frequency of double-nucleotide substitutions. *Science* 287: 1283–1286.

BALDWIN, B. G. 2005. Origin of the serpentine-endemic herb *Layia discoidea* from the widespread *L. glandulosa* (Compositae). *Evolution* 59: 2473–2479.

BAUM, D. A. 2007. Concordance trees, concordance factors, and the exploration of reticulate genealogy. *Taxon* 56: 417–426.

BININDA-EMONDS, O. R. P. 2004. Phylogenetic supertrees: Combining information to reveal the tree of life. Kluwer, Dordrecht, Netherlands.

BOHS, L. 2004. A chloroplast DNA phylogeny of *Solanum* section *Lasiocarpa. Systematic Botany* 29: 177–187.

BOHS, L. 2005. Major clades in *Solanum* based on *ndhF* sequence data. *In* R. C. Keating, V. C., Hollowell, and T. B. Croat [eds.], A festschrift for William G. D'Arcy: The legacy of a taxonomist, 27–49. Missouri Botanical Garden Press, St. Louis, Missouri, USA.

BOHS, L. 2007. Phylogeny of the Cyphomandra clade of the genus *Solanum* (Solanaceae) based on ITS sequence data. *Taxon* 56: 1012–1026.

BORCHSENIUS, F. 1997. Patterns of plant species endemism in Ecuador. *Biodiversity and Conservation* 6: 379–399.

COMES, H. P., AND R. J. ABBOTT. 2001. Molecular phylogeography, reticulation, and lineage sorting in Mediterranean *Senecio* sect. *Senecio* (Asteraceae). *Evolution* 55: 1943–1962.

CRAWFORD, D. J. 2010. Progenitor-derivative species pairs and plant speciation. *Taxon* 59: 1413–1423.

CUMMINGS, M. P., S. A. HANDLEY, D. S. MYERS, D. L. REED, A. ROKAS, AND K. WINKA. 2003. Comparing bootstrap and posterior probability values in the four-taxon case. *Systematic Biology* 52: 477–487.

DARLU, P., AND G. LECOINTRE. 2002. When does the incongruence length difference test fail? *Molecular Biology and Evolution* 19: 432–437.

DEGNAN, J. H., AND N. A. ROSENBERG. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution* 24: 332–340.

DE QUEIROZ, K. 2007. Species concepts and species delimitation. *Systematic Biology* 56: 879–886.

DOLPHIN, K., R. BELSHAW, C. D. L. ORME, AND D. L. J. QUICKE. 2000. Noise and incongruence: Interpreting results of the incongruence length difference test. *Molecular Phylogenetics and Evolution* 17: 401–406.

DOWTON, M., AND A. D. AUSTIN. 2002. Increased congruence does not necessarily indicate increased phylogenetic accuracy: The behavior of the incongruence length difference test in mixed-model analyses. *Systematic Biology* 51: 19–31.

ERIXON, P., B. SVENNBLAD, T. BRITTON, AND B. OXELMAN. 2003. Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. *Systematic Biology* 52: 665–673.

FARRIS, S. J., M. KÄLLERSJÖ, A. G. KLUGE, AND C. BULT. 1994. Testing significance of incongruence. *Cladistics* 10: 315–319.

GOTTLIEB, L. D. 2003. Rethinking classic examples of recent speciation in plants. *New Phytologist* 161: 71–82.

GRANT, V. 1981. Plant speciation, 2nd ed. Columbia University Press, New York, New York, USA.

GREEN, M. J., D. A. THOMPSON, AND D. J. MACKENZIE. 1999. Easy and efficient DNA extraction from woody plants for the detection of phytoplasmas by polymerase chain reaction. *Plant Disease* 83: 482–485.

HAMILTON, M. B. 1999. Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular Ecology* 8: 521–523.

HIPP, A. L., J. C. HALL, AND K. J. SYTSMA. 2004. Congruence versus phylogenetic accuracy: Revisiting the incongruence length difference test. *Systematic Biology* 53: 81–89.

HÖRANDL, E., AND T. F. STUESSY. 2010. Paraphyletic groups as natural units of biological classification. *Taxon* 59: 1641–1653.

HUELSENBECK, J. P., AND F. RONQUIST. 2001. MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* 17: 754–755.

HUSON, D. H., AND D. BRYANT. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23: 254–267.

KNOWLES, L. L., AND B. C. CARSTENS. 2007. Delimiting species without monophyletic gene trees. *Systematic Biology* 56: 887–895.

KOOPMAN, M. M., AND D. A. BAUM. 2010. Isolating nuclear genes and identifying lineages without monophyly: An example of closely related species from southern Madagascar. *International Journal of Plant Sciences* 171: 761–771.

LEACHÉ, A. D., AND B. RANNALA. 2011. The accuracy of species tree estimation under simulation: A comparison of methods. *Systematic Biology* 60: 126–137.

LEVIN, D. A. 1993. Local speciation in plants: the rule, not the exception. *Systematic Botany* 18: 197–208.

LEVIN, R. A., N. R. MYERS, AND L. BOHS. 2006. Phylogenetic relationships among the "spiny solanums" (*Solanum* subgenus *Leptostemonum*, Solanaceae). *American Journal of Botany* 93: 157–169.

LEVIN, R. A., K. WATSON, AND L. BOHS. 2005. A four-gene study of evolutionary relationships in *Solanum* section *Acanthophora*. *American Journal of Botany* 92: 603–612.

LEVIN, R. A., A. WHELAN, AND J. S. MILLER. 2009. The utility of nuclear conserved ortholog set II (COSII) genomic regions for species-level phylogenetic inference in *Lycium* (Solanaceae). *Molecular Phylogenetics and Evolution* 53: 881–890.

LIU, L. 2008. BEST: Bayesian estimation of species trees under the coalescent model. *Bioinformatics* 24: 2542–2543 www.stat.osu.edu/~dkp/BEST/.

LIU, L., AND D. K. PEARL. 2007. Species trees from gene trees: Reconstructing Bayesian posterior distributions of a species phylogeny using estimated gene tree distributions. *Systematic Biology* 56: 504–514.

LIU, L., D. K. PEARL, AND S. V. EDWARDS. 2003. BEST 2.3 manual. Website http://www.stat.osu.edu/~dkp/BEST/help/.

MADDISON, W. P., AND D. R. MADDISON. 2009. Mesquite: A modular system for evolutionary analysis, version 2.6. Website http://mesquiteproject.org.

MASON-GAMER, R. J., AND E. A. KELLOGG. 1996. Testing for phylogenetic conflict among molecular data sets in the tribe Triticeae (Gramineae). *Systematic Biology* 45: 524–545.

MAYR, E. 1942. Systematics and the origin of species. Columbia University Press, New York, New York, USA.

MUTKE, J., AND W. BARTHLOTT. 2005. Patterns of vascular plant diversity at continental to global scales. *Biologiske Skrifter* 55: 521–531.

MYERS, N. 1988. Threatened biotas: "Hot spots" in tropical forests. *Environmentalist* 8: 187–208.

MYERS, N., R. A. MITTERMEIER, C. G. MITTERMEIER, G. A. B. DA FONSECA, AND J. KENT. 2000. Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858.

NYLANDER, J. A. A. 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University. Website http://www.csit.fsu.edu/~nylander/.

PAAPE, T., B. IGIC, S. D. SMITH, R. OLMSTEAD, L. BOHS, AND J. R. KOHN. 2008. A 15-Myr-old genetic bottleneck. *Molecular Biology and Evolution* 25: 655–663.

PERALTA, I. E., S. KNAPP, AND D. M. SPOONER. 2008. Taxonomy of wild tomatoes and their relatives (*Solanum* sect. *Lycopersicoides*, sect. *Juglandifolia*, sect. *Lycopersicon*; Solanaceae). *Systematic Botany Monographs* 84: 1–186.

POSADA, D., AND T. R. BUCKLEY. 2004. Model selection and model averaging in phylogenetics: Advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology* 53: 793–808.

PROHENS, J., G. J. ANDERSON, J. M. BLANCA, J. CAÑIZARES, E. ZURIAGA, AND F. NUEZ. 2006. The implications of AFLP data for the systematics of the wild species of *Solanum* section *Basarthrum. Systematic Botany* 31: 208–216.

RAMBAUT, A. 1996. Se-Al: Sequence Alignment Editor. Computer program and documentation distributed by the author, website http://evolve.zoo.ox.ac.uk/.

RIESEBERG, L. H. 1991. Homoploid reticulate evolution in *Helianthus* (Asteraceae): Evidence from ribosomal genes. *American Journal of Botany* 78: 1218–1237.

RIESEBERG, L. H., AND L. BROUILLET. 1994. Are many plant species polyphyletic? *Taxon* 43: 21–32.

RIESEBERG, L. H., R. CARTER, AND S. ZONA. 1990. Molecular tests of the hypothesized hybrid origin of two diploid *Helianthus* species (Asteraceae). *Evolution* 44: 1498–1511.

RODRÍGUEZ, F., AND D. M. SPOONER. 2009. Nitrate reductase phylogeny of potato (*Solanum* sect. *Petota*) genomes with emphasis on the origins of the polyploid species. *Systematic Botany* 34: 207–219.

RODRÍGUEZ, F., F. WU, C. ANÉ, S. TANKSLEY, AND D. M. SPOONER. 2009. Do potatoes and tomatoes have a single evolutionary history, and what proportion of the genome supports this history? *BMC Evolutionary Biology* 9: 191 .

RONQUIST, F., AND J. P. HUELSENBECK. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.

ROKAS, A., B. L. WILLIAMS, N. KING, AND S. B. CARROLL. 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature* 425: 798–804.

SANDERSON, M. J., AND H. B. SHAFFER. 2002. Troubleshooting molecular phylogenetic analyses. *Annual Review of Ecology and Systematics* 33: 49–72.

SANG, T., D. J. CRAWFORD, AND T. F. STUESSY. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *American Journal of Botany* 84: 1120–1136.

SEELANAN, T., A. SCHNABEL, AND J. F. WENDEL. 1997. Congruence and consensus in the cotton tribe (Malvaceae). *Systematic Botany* 22: 259–290.

SHAW, J., E. B. LICKEY, J. T. BECK, S. B. FARMER, W. LIU, J. MILLER, K. C. SIRPUN, 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.

SHAW, J., E. B. LICKEY, E. E. SCHILLING, AND R. L. SMALL. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: The tortoise and the hare III. *American Journal of Botany* 94: 275–288.

SHIMODAIRA, H., AND M. HASEGAWA. 1999. Multiple comparisons of log-likelihoods with applications for phylogenetic inference. *Molecular Biology and Evolution* 16: 1114–1116.

SIMMONS, M. P., K. M. PICKETT, AND M. MIYA. 2004. How meaningful are Bayesian support values? *Molecular Biology and Evolution* 21: 188–199.

SPOONER, D. M. 2009. DNA barcoding will frequently fail in complicated groups: An example in wild potatoes. *American Journal of Botany* 96: 1177–1189.

SPOONER, D. M., G. J. ANDERSON, AND R. K. JANSEN. 1993. Chloroplast DNA evidence for the interrelationships of tomatoes, potatoes, and pepinos. *American Journal of Botany* 80: 676–688.

SPOONER, D. M., F. RODRÍGUEZ, Z. POLGÁR, H. E. BALLARD JR., AND S. H. JANSKY. 2008. Genomic origins of potato polyploids: GBSSI gene sequencing data. *Plant Genome* 48: S27–S36.

STERN, S., T. WEESE, AND L. BOHS. 2010. A three-gene evolutionary study of *Solanum* section *Androceras. Systematic Botany* 35: 885–893.

SULLIVAN, J. 1996. Combining data with different distributions of among-site rate variation. *Systematic Biology* 45: 375–380.

SWOFFORD, D. L. 2003. PAUP*: Phylogenetic analysis using parsimony (* and other methods), version 4.0b10. Sinauer, Sunderland, Massachussets, USA.

TABERLET, P., L. GIELLY, G. PAUTOU, AND J. BOUVET. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* 17: 1105–1110.

TEPE, E. J., AND L. BOHS. 2009. Three new species of *Solanum* section *Herpystichum* (Solanaceae) from Ecuador. *Journal of the Botanical Research Institute of Texas* 3: 511–519.

TEPE, E. J., AND L. BOHS. 2010. A molecular phylogeny of *Solanum* sect. *Pteroidea* (Solanaceae) and the utility of COSII markers in resolving relationships among closely related species. *Taxon* 59: 733–743.

TEPE, E. J., AND L. BOHS. 2011. A revision of *Solanum* section *Herpystichum. Systematic Botany* 36(4): in press.

VARGAS, P., B. G. BALDWIN, AND L. CONSTANCE. 1998. Nuclear ribosomal DNA evidence for a western North American origin of Hawaiian and South American species of *Sanicula* (Apiaceae). *Proceedings of the National Academy of Sciences, USA* 95: 235–240.

WEESE, T. L., AND L. BOHS. 2007. A three-gene phylogeny of the genus *Solanum* (Solanaceae). *Systematic Botany* 32: 445–463.

WEESE, T. L., AND L. BOHS. 2010. Eggplant origins: Out of Africa, into the Orient. *Taxon* 59: 49–56.

WHITE, T. J., T. BRUNS, S. LEE, AND J. W. TAYLOR. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *In* M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White. [eds.], PCR protocols: A guide to methods and applications, 315–322. Academic Press, San Diego, California, USA.

WIENS, J. J. 1998. Combining data sets with different phylogenetic histories. *Systematic Biology* 47: 568–581.

WU, F., L. A. MUELLER, D. CROUZILLAT, V. PÉTIARD, AND S. D. TANKSLEY. 2006. Combining bioinformatics and phylogenetics to identify large sets of single-copy orthologous genes (COSII) for comparative, evolutionary and systematic studies: A test case in the euasterid plant clade. *Genetics* 174: 1407–1420.

YODER, A. D., J. A. IRWIN, AND B. A. PAYSEUR. 2001. Failure of the PHT to determine data combinability for slow loris phylogeny. *Systematic Biology* 50: 408–424.

APPENDIX 1. List of accessions used in this study, with voucher information and GenBank accession numbers for the 10 markers studied. Missing data are indicated by dashes (—). The following abbreviations are used for herbaria: BM = The Natural History Museum, London; COL = Herbario Nacional Colombiano; MO = Missouri Botanical Garden; NY = New York Botanical Garden; PTIS = Potato Introduction Station Herbarium; QCNE = Herbario Nacional del Ecuador; US = United States National Herbarium; UT = Garrett Herbarium, University of Utah.

Taxon; *voucher* (herbarium); Collection location; GenBank accessions: GBSSI; ITS; cos5; cos11; cos9B; cos10B; cos1C; *psbA–trnH*; *trnS–trnG*; *trnT–trnF*.

***Solanum anceps* Ruiz & Pav.**; *Bohs 2790a* (UT); Cultivated (cutting from Bolivia); GQ221593; GQ221541; HQ856154; GQ221513; GQ221486; HQ856188; GQ221458; HQ856107; HQ856087; GQ221568. *Tepe & Stern 2608* (QCNE); Ecuador; GQ221597; GQ221545; HQ856155; GQ221517; GQ221489; HQ856189; GQ221462; HQ856108; HQ856088; GQ221572. ***Solanum brevifolium* Dunal**; *Bohs 3112* (UT); Ecuador; GQ221614; GQ221562; —; GQ221535; GQ221507; HQ856190; GQ221480; HQ856109; HQ856109; GQ221589. ***Solanum bulbocastanum* Dunal**; *Tarn 153* (PTIS); Mexico; DQ169020; GQ221564; HQ856157; GQ221536; GQ221508; HQ856157; GQ221481; HQ856111; HQ856111; DQ180444. ***Solanum caripense* Dunal**; *Bohs 3149* (UT); Ecuador; GQ221615; GQ221563; HQ856156; GQ221537; GQ221509; HQ856191; GQ221482; HQ856191; HQ856090; GQ221590. ***Solanum crassinervium* Tepe**; *Tepe 2729* (QCNE); Ecuador; HQ856216; HQ856119; HQ856119; HQ856202; HQ856166; HQ856182; HQ856133; HQ856133; HQ856080; HQ856062. ***Solanum dalibardiforme* Bitter**; *Mora 924* (COL); Colombia; HQ856220; HQ856123; HQ856152; HQ856206; HQ856206; HQ856186; HQ856137; HQ856105; HQ856084; HQ856066. ***Solanum dolichorhachis* Bitter**; *Mexia 6617* (US); Ecuador; HQ856212; —; —; —; —; —; —; —; HQ856074; HQ856074. ***Solanum evolvulifolium* Greenm.**; *Knapp & Mallett 7198* (BM); Panama; DQ169028; HQ856113; HQ856113; HQ856194; HQ856159; HQ856173; HQ856126; HQ856094; HQ856070; DQ180464. *Bohs 2500* (UT); Costa Rica; GQ221616; GQ221565; HQ856140; Q221538; GQ221510; HQ856140; GQ221483; HQ856093; HQ856069; GQ221591. *Croat 93047* (MO); Ecuador; HQ856209; HQ856114; —; HQ856195; HQ856195; HQ856174; HQ856127; —; HQ856071; HQ856054. *Tepe 2585* (QCNE); Ecuador; HQ856211; HQ856211; HQ856143; HQ856197; HQ856162; HQ856176; HQ856129; HQ856096; HQ856073; HQ856056. *Tepe 2671* (QCNE); Ecuador; HQ856210; HQ856115; HQ856115; HQ856196; HQ856161; HQ856175; HQ856128; HQ856095; HQ856072; HQ856055. ***Solanum limoncochaense* Tepe**; *Tepe 2627* (QCNE); Ecuador; HQ856217; HQ856120; HQ856149; HQ856203; HQ856167; HQ856183; HQ856134; HQ856102; HQ856081; HQ856063. *Bohs 3701* (UT); Cultivated (seeds: Tepe 2627); Ecuador; HQ856218; HQ856121; HQ856150; HQ856204; HQ856168; HQ856184; HQ856135; HQ856103; HQ856082; HQ856064. ***Solanum loxophyllum* Bitter**; *Tepe 2726* (QCNE); Ecuador; HQ856213; HQ856117; HQ856146; HQ856200; HQ856164; HQ856180; HQ856131; HQ856099; HQ856075; HQ856058. *Tepe 2698* (QCNE); Ecuador; HQ856214; HQ856118; HQ856147; HQ856201; HQ856165; HQ856181; HQ856132; HQ856100; HQ856076; HQ856059. ***Solanum lycopersicum* L.**; No voucher; USA (cultivated); DQ169036; GQ221566; HQ856158; GQ221539; GQ221511; HQ856193; GQ221484; HQ856112; HQ856092; DQ180450. ***Solanum pacificum* Tepe**; *Tepe 2696* (QCNE); Ecuador; HQ856219; HQ856122; HQ856151; HQ856205; HQ856169; HQ856185; HQ856136; HQ856104; HQ856083; HQ856065. ***Solanum pentaphyllum* Bitter**; *Grant 99-03335* (US); Venezuela; HQ856215;—; —; HQ856199; —; HQ856179; —; —; HQ856079; HQ856061. ***Solanum phaseoloides* Pol.**; *Bohs 2485* (UT); Costa Rica; GQ221617; GQ221567; HQ856144; GQ221540; GQ221512; HQ856177; GQ221485; HQ856097; HQ856077; GQ221592. *Bohs 3499* (UT); Costa Rica; HQ856222; —; HQ856145; HQ856198; HQ856163; HQ856178; HQ856130; HQ856098; HQ856078; HQ856060. ***Solanum trifolium* Dunal**; *Spooner 5073* (NY); Ecuador; —; HQ856124; —; HQ856207; —; —; HQ856138; —; HQ856085; HQ856067. *Tepe 2682* (QCNE); Ecuador; HQ856221; HQ856125; HQ856153; HQ856208; HQ856171; HQ856187; HQ856139; HQ856106; HQ856086; HQ856068.