

Integrated Molecular and Morphological Studies of the *Daucus guttatus* Complex (Apiaceae)

Carlos I. Arbizu,¹ Philipp W. Simon,¹ Fernando Martínez-Flores,² Holly Ruess,¹
Manuel B. Crespo,² and David M. Spooner^{1,3}

¹U. S. Department of Agriculture, Agricultural Research Service, Vegetable Crops Research Unit, Department of Horticulture, University of Wisconsin, 1575 Linden Drive, Madison, Wisconsin 53706-1590, U. S. A.

²Departamento de Ciencias Ambientales y Recursos Naturales and Centro Iberoamericano de la Biodiversidad (CIBIO), Universidad de Alicante, P. O. Box 99, ES-03080 Alicante, Spain

³Author for correspondence (David.Spooner@ars.usda.gov)

Communicating Editor: Carolyn Ferguson

Abstract—In a previous study using 94 nuclear orthologs, we reported the species status of the *Daucus guttatus* complex to be unresolved, partitioned into three clades. In the present study, a subset of ten of these 94 orthologs was used to infer the phylogeny of the *D. guttatus* complex and related species. A near parallel set of accessions, planted in a common garden, was used for morphological analyses. The molecular trees are highly resolved for most of the clades, grouping accessions of the *D. guttatus* complex into four clades. Bayesian concordance analysis and a coalescent approach gave slightly different topologies. Morphological data likewise support four taxa in the complex. Moreover, herbarium research from a companion study informs nomenclature for taxa of the complex. We identify these four clades as *D. bicolor*, *D. conchitae*, *D. guttatus*, and *D. setulosus*; internested in or among these segregates are the phenetically distinctive species *D. glochidiatus*, *D. involucratus*, *D. littoralis*, and *D. pusillus*. Our research redefines species variation in the *D. guttatus* complex, clarifies species names, interspecific relationships, confirms a useful subset of nuclear orthologs for studies of dominant topologies of *Daucus*, and discovers morphological characters allowing proper identification of the four species of the *D. guttatus* complex and related species.

Keywords—Carrot species, germplasm, nuclear orthologs, phenetics.

Carrot (*Daucus carota* L. subsp. *sativus* Hoffm.) is considered among the ten most important vegetables worldwide and is the highest value crop in the Apiaceae (Rubatzky et al. 1999; Simon 2000; Vilela 2004). The genus *Daucus* is most common in the Mediterranean region and is characterized by dorsally compressed mericarps, hairs on the primary ridges of the mericarp, and spines on the secondary ridges of the mericarps (Okeke 1978).

The latest comprehensive taxonomic monograph of *Daucus* by Sáenz (1981) recognized 21 species divided into five sections: *Daucus* sections *Anisactis* DC., *Chrysodaucus* Thell., *Daucus* L., *Meoides* Lange, and *Platyspermum* DC. Rubatzky et al. (1999) later estimated 25 species of *Daucus*. More than 60 species have been proposed as variants within the “*D. carota* complex” alone (Small 1978), for which there are no or only poorly developed barriers to interbreeding among either the wild forms or their domesticates (Small 1978). Morphological data fail to clearly distinguish many of the taxa within this complex (Spooner et al. 2014).

Because of new species descriptions and the outdated taxonomy of Sáenz (1981; a largely intuitive classification lacking specimen citations, distribution maps and detailed descriptions), taxonomists have relied on local floras for identifying *Daucus* such as those from Algeria (Quézel and Santa 1963), Europe (Heywood 1968), the Iberian Peninsula and Balearic Islands (Pujadas Salvà 2003), Libya (Jafri and El-Gadi 1985), Morocco (Jury 2002), Tunisia (Pottier-Alapetite 1979; Le Floc’h et al. 2010), Palestine (Zohary 1972), Syria (Mouterde 1966), and Turkey and the East Aegean Islands (Cullen 1972).

Delineating species boundaries correctly is crucial to understanding the diversity of life (Dayrat 2005). Historically, delineation of species and their taxonomy has been based mostly on morphological characters (Patterson et al. 1993; Pisani et al. 2007). The utility of DNA sequences for taxonomic studies is also well established (Tautz et al. 2003). Different authors have successfully integrated morphological

and molecular data to delineate and identify species (Emadzade et al. 2010; Martínez-Flores et al. 2012, Cohen 2014). Morphological and molecular studies have been conducted in *Daucus* to distinguish taxa with varying results (Vivek and Simon 1999; Bradeen et al. 2002; Spooner et al. 2013, 2014; Arbizu et al. 2014a, 2014b; Mezghani et al. 2014; Lee and Park 2014).

Single and low-copy nuclear genes provide significant phylogenetically informative characters complementing other types of data in plant systematics (Zimmer and Wen 2013). Neither nuclear orthologs (Arbizu et al. 2014b) nor morphological data (Spooner et al. 2013) have distinguished the subspecies of *D. carota* L., but nuclear orthologs and morphology (Arbizu et al. 2014a) generally provide effective data sets to distinguish the other species in *Daucus*. Relevant to the present study, molecular and morphological studies of multiple accessions of the *D. guttatus* complex collected in Greece, Iran, Israel, Lebanon, Syria, and Turkey demonstrated high diversity in these taxa, which were partitioned into three well-supported clades (Arbizu et al. 2014b), and phenetic groups (Arbizu et al. 2014a), suggesting that these putative species needed reevaluation.

Currently, there are no accepted criteria to clearly distinguish the segregates of the *D. guttatus* complex. They are morphologically similar and difficult to distinguish, causing frequent misidentification of newly acquired germplasm. We here aim to resolve the species boundaries and relationships of these putative species, and other related species as determined by Arbizu et al. (2014b), by integrating molecular and morphological data.

MATERIALS AND METHODS

Accessions Examined—We examined 83 accessions of *Daucus*, and one accession of a non-*Daucus* outgroup *Orlaya daucooides* (84 accessions in total), collected in 12 countries (Suppl. Data 1). Arbizu et al. (2014b) reported that clade B contains samples of the *D. guttatus* complex partitioned into three well supported clades (therein called *D. guttatus* 1,

D. guttatus 2 and *D. guttatus* 3), and other related species that are morphologically distinct from the members of the complex (*D. glochidiatus*, *D. involucreatus*, *D. littoralis*, and *D. pusillus*). Sixty-five of the 84 accessions sampled for the present study correspond to the *D. guttatus* complex and nine accessions are closely related species (sensu Arbizu et al. 2014b; and including additional sampling). Notably, the epitype of *D. guttatus* (PI 652342, designated by Martínez-Flores et al. 2016) is included in this sampling. Our morphological analysis included 40 accessions of six *Daucus* species including most of those examined in the molecular analysis (Suppl. Data 1). All accessions were obtained from the United States National Plant Germplasm System, maintained at the North Central Regional Plant Introduction Station (NCRPIS) in Ames, Iowa. Full details of the accessions examined in this study are available at the Germplasm Resources Information Network — GRIN (http://www.ars-grin.gov/npgs/acc/acc_queries.html).

DNA Extraction and Sequence Generation—DNA was extracted using the CTAB method (Doyle and Doyle 1990) from young leaves of accessions planted in a greenhouse at the University of Wisconsin-Madison. We used the ten nuclear orthologs that were identified by Arbizu et al. (2014b) as a useful subset to recover dominant topologies in *Daucus* (the subset determined by the result obtained by the concatenated data set of all 94 orthologs). DNA sequences of 53 accessions of the *D. guttatus* complex for these ten nuclear orthologs were obtained using Asymmetric PCR Single-Strand Conformation Polymorphism, an efficient alternative technique for isolating allelic variants of *Daucus* (Rodríguez et al. 2011). Sanger sequencing (Sanger et al. 1977) was conducted at the Biotechnology Center of the University of Wisconsin-Madison. DNA sequences for the other accessions (Suppl. Data 1) were obtained from Arbizu et al. (2014b).

Phylogenetic Analyses—DNA sequences were edited with Staden package version 1.7.0 (Staden 1996) and aligned using MUSCLE version 3.5 (Edgar 2004). Subsequent minor manual alignment corrections and shortening of homopolymers to a maximum of six base pairs (to match the data of Arbizu et al. 2014b) were performed using Mesquite version 3.03 (Maddison and Maddison 2015). We rooted our trees on *Orlaya daucooides* based on Arbizu et al. (2014b). Maximum parsimony (MP) analysis was conducted in PAUP* version 4.0a145 (Swofford 2002). Missing data and gaps were all scored as missing data. All characters were treated as unordered and weighted equally (Fitch 1971). The most parsimonious trees were found using a heuristic search (Farris 1970) by generating 100,000 random taxon addition sequence replicates using tree-bisection-reconnection (TBR) and holding one tree for each replicate. Then, we ran a final heuristic search of the most equally parsimonious trees from this analysis using TBR and MULPARS. Bootstrap values (Felsenstein 1985) for the clades were estimated using 1,000 replicates with a heuristic search, TBR and MULPARS, setting MAXTREES to 1000.

Selection of the best-fit evolutionary models for the individual gene data sets was computed using the Akaike information criterion (AIC), using jModelTest version 2.1.4 (Darriba et al. 2012). With these models, we obtained a maximum likelihood (ML) tree with GARLI version 2.1 (Zwickl 2006). Using the same program, 1,000 non-parametric bootstrap inferences were obtained. Both analyses were performed by the GARLI web service hosted at www.molecularevolution.org (Bazin et al. 2014).

In addition, we performed a Bayesian concordance analysis (BCA) (Ané et al. 2007) to obtain a primary concordance tree and clade concordance factors with 95% credibility intervals using the program BUCKY version 1.4.3 (Larget et al. 2010). Based on Arbizu et al. (2014b) for the BCA pruned analysis we chose 23 representative accessions (Suppl. Data 1) from major clades as determined from the MP and ML analyses to explore gene to gene conflict in our data set. All ten genes with their corresponding model of nucleotide substitution (Suppl. Data 2) were analyzed separately in MrBayes version 3.2.3 (Ronquist et al. 2012) using the BEAGLE library (Ayres et al. 2012) with four chains and two searches run simultaneously for ten million generations sampling every 1,000 generations. This analysis was conducted via the CIPRES (Miller et al. 2010) portal at San Diego Supercomputer Center (<http://www.phylo.org>). We summarized the MrBayes results for the ten genes using the program *mbsum* included in BUCKY, removing 1,001 trees from each chain as burn-in. We then performed the BCA with four independent runs with four linked chains for all four different levels of discordance: $\alpha = 0.1, 1, 10$, and infinite (a larger value of α corresponds to greater gene tree incongruence); in each run with 1,100,000 generations, with 100,000 generations discarded as the burn-in period. Default settings were used for all other parameters.

We also performed a Bayesian analysis using BEAST package version 1.8.2 (Drummond et al. 2012) to obtain a species tree estimation using a coalescent approach. First, an XML format file was generated using

BEAUti version 1.8.2 with ten genes with their corresponding model of evolution (Suppl. Data 2) and Yule process as a species tree prior (Gernhard 2008). Based on our MP and ML tree topology, accessions from *Daucus* clades 1, 2, 3 and 4 (Results) were labeled as different species. Even though our MP and ML results show that *D. involucreatus* is embedded within clade 3, we here considered it as a different species based on the molecular and morphological results of Spalik and Downie (2007), Arbizu et al. (2014a, b), and Lee and Park (2014). Markov chain Monte Carlo (MCMC) chains for both analyses were run for one billion generations sampling every 50,000 generations. We imported the log files into Tracer version 1.6 (Rambaut and Drummond 2014) from the BEAST package to analyze the convergence to the stationary distribution and the effective sample size (ESS) of each parameter. The sample files of trees produced by BEAST were summarized discarding 25% of the trees onto a single “target” tree using the program TreeAnnotator version 1.8.2. The resulting trees were viewed in FigTree version 1.4.2 (<http://tree.bio.ed.ac.uk/software/figtree/>).

Morphological Analyses—We considered species that belong to clade B of Arbizu et al. (2014b) for our morphological study. In addition, accessions of *Daucus bicolor*, not characterized by Arbizu et al. (2014a, b), were considered as part of clade B based on MP and ML analyses of the present study (see Results). Each accession was direct seeded by hand in 1m × 6m observation plots at the West Madison Agricultural Research Station, University of Wisconsin, Madison. We watered the rows by hand and weeded the field with a small plot tiller, hoe, and hand weeding. Low germination rates of some accessions prevented us from having plants in every observation plot.

Daucus accessions were scored and analyzed for 41 characters in total: 31 continuous and 10 nominal characters (Table 1). All characters were always recorded by the same individual and from at least one plant per accession. These characters were selected to represent all those previously used in the latest comprehensive monograph of *Daucus* (Sáenz 1981) and in subsequent *Daucus* morphological studies (Small 1978; Arbizu et al. 2014a; Mezghani et al. 2014; Spooner et al. 2014). Plant and leaf characters were evaluated in the field with a ruler or digital calipers, and floral and mericarp characters were measured in the laboratory using a stereoscopic microscope. Forty observation plots lacked plants due to no germination or unintentional elimination during weeding. We obtained one data set including 40 accessions containing all measurements of 82 plants for the 41 characters (plant, leaf, flower and mericarp). Accession PI 652333 (Suppl. Data 1) had neither plant nor fruit data.

We summarized our morphological data using R version 3.2.0 software (R Core Team 2015). For multivariate analyses, means were calculated for the continuous characters using the *ddply* function in the *plyr* package (Wickman 2011). A descriptive statistical analysis was conducted to verify the mean, median, standard deviation and range of values. Box plots were used as a graphic tool to visualize comparisons across accessions and to check for outliers that may represent erroneous entries. Accessions were classified into groups based on our molecular phylogenetic analyses. Then we performed stepwise discriminant analyses (linear, common variance) using all 31 continuous variables in our data set comprising 40 accessions.

Discriminant analyses were performed in JMP® version 11.2 (SAS Institute Inc. 1989–2007) using a backward selection; a model with significant variables in identifying accession structure was obtained by removing characters one at a time until the model *F*-test *p* value ≤ 0.05 . We then performed canonical variate analyses (CVA). Hierarchical cluster analysis (HCA) was conducted using all 10 nominal characters and the significant continuous characters previously obtained. For our data set (31 nominal characters and 40 accessions in total), we considered the following: (1) all taxa with 31 nominal characters (plant, leaf, flower and mericarp); and (2) all taxa with a subset of characters — the mericarp characters only. Analysis using the mericarp characters only was performed to see if mericarps alone could be used to distinguish the species analyzed in this study. In addition, a linear discriminant analysis (LDA) was performed with the data mentioned above using the *lda* function in the *MASS* package (Venables and Ripley 2002) in R. Results were very similar to our CVA and we do not discuss this further.

In concert with our companion study (Martínez-Flores et al. 2016; this issue of *Systematic Botany*), we studied the original descriptions and the floristic literature to assign species names to the four molecular clades (putative species) of the *D. guttatus* complex. Proper names were assigned by (i) analyzing digital pictures from GRIN website and our experimental plots, and (ii) checking names with herbarium specimens. Potential synonyms were inferred through: (1) the latest comprehensive taxonomic treatment of *Daucus* (Sáenz 1981); (2) regional floras outlined in the Introduction; and (3) a database of digitized plant specimens

TABLE 1. The 41 morphological characters, modeling type, *F*-test *p* values of characters retained in a stepwise discriminant analyses for our data set (40 accessions containing all evaluations for the 41 characters): (1) all taxa (*Daucus guttatus* complex, *D. involucratus* and *D. littoralis*); (2) subset with 10 mericarp characters only and all taxa. ^aFurther details of these morphological characters can be obtained from the USDA, ARS, National Genetic Resources Program, Germplasm Resources Information Network (GRIN). [Online Database] National Germplasm Resources Laboratory, Beltsville, Maryland. Available at: <http://www.ars-grin.gov/cgi-bin/npgs/html/crop.pl?70> [Accessed 02 April 2015]. ^bN, nominal; C, continuous. Only continuous characters were used for stepwise discriminant analyses. Nominal characters were used for Hierarchical Cluster Analysis. ^cNS, not significant at *p* = 0.05.

Character ^a	Modeling type ^b	<i>F</i> -test <i>p</i> values ^c	
		<i>p</i> ¹	<i>p</i> ²
Plant			
Plant height (cm)	C	0.0031341	
Stem diameter (mm)	C	0.0093228	
Leaf			
Leaf length including petiole (cm)	C	0.0102793	
Leaf width at maximum (cm)	C	NS	
Petiole length (cm)	C	0.0088217	
Petiole diameter (mm)	C	0.0069283	
Petiole shape (round, 1; semi-round, 2; flat, 3)	N		
Leaf and petiole pubescence (smooth, 1; intermediate, 2; very hairy, 3)	N		
Flower			
Peduncle pubescence (glabrous, 1; soft hairs, 2; scabrous, 3; very scabrous, 4)	N		
Primary umbel shape, full bloom (convex, 1; flat, 2; concave, 3)	N		
Primary umbel shape, mature umbel (convex, 1; flat, 2; concave, 3)	N		
Primary umbel diameter (cm)	C	0.0000344	
Pigmented central umbellules diameter (cm)	C	NS	
Bract length (mm)	C	NS	
Bract width (mm)	C	NS	
Bract posture (deflexed, not deflexed [outward or upward])	N		
Number of bract lobe points	C	0.000071	
Number of bract lobe pairs	C	0.0000028	
Number of umbel rays	C	0.000096	
Pigmented central umbellule (concolorous to outer [uniform color], 1; differently pigmented, 2)	N		
Length of longest peripheral ray (cm)	C	0.0000006	
Length of shortest peripheral ray (cm)	C	NS	
Petal color (white, cream, yellow [only <i>D. carota</i> subsp. <i>capillifolius</i>], pink)	N		
Anther color (white, cream, yellow, pink, purple, brown)	N		
Peripheral petal length (mm)	C	NS	
Central petal length (mm)	C	0.0024263	
Stamen length (mm)	C	NS	
Style length (mm)	C	0.0000022	
Stylopodium length (mm)	C	NS	
Stylopodium width (mm)	C	0.0119646	
Mericarp			
Mericarp length (mm), spines and style excluded	C	NS	0.0000262
Mericarp width (mm), crest and spines excluded	C	0.004501	NS
Confluency of mericarp spines (separate, 1; little confluency, 2; much confluency, 3)	N		
Number of spines on the four secondary mericarp ribs	C	0.0002023	0.007286
Length of spines on the secondary ribs (mm)	C	0.0000021	NS
Vittae length (μm); see Suppl. Figure 4	C	0.0000005	0.0000024
Vittae width (μm); see Suppl. Figure 4	C	NS	NS
Mericarp width/ Mericarp length ratio	C	0.0000002	0
Spine width (μm)	C	0.0000087	NS
Crest width (μm)	C	0.0004024	0.0007935
Length of bristles of primary ribs (μm)	C	0.0001231	NS

available at JSTOR Global Plants (<https://plants.jstor.org/>). With the clade-specific morphological characters determined in our study, and those determined by examination of herbarium sheets by Martínez-Flores et al. (2016), we constructed a key to the identify species in the *D. guttatus* complex and related species in clade B of Arbizu et al. (2014b).

RESULTS

Completeness of Sequencing—Our aligned DNA data set consisted of 10,982 base pairs (bp). The alignment, Supplementary Data 1 and 2, and Supplementary Figs. 1–4 are deposited in Dryad (<http://datadryad.org/review?doi=doi:10.5061/dryad.mp10n>). Sequences of the 53 accessions newly

sequenced here (see Suppl. Data 2 for GenBank accession numbers) resulted in one or two alleles differing only in minor differences in every gene. In previous work, the topologies of the phylogeny of *Daucus* using a single allele vs. two alleles merged into one showed minor topological differences (Arbizu et al. 2014b). As a result, we analyzed the one allele containing the least number of missing data (base pairs). Of these ten markers/53 accessions matrices (530 sequences) there were 48 missing sequences, resulting in 9.1% missing data and we did not observe any ambiguous alignments. The entire data matrix of 840 sequences (ten markers/84 accessions matrices) had 57 missing sequences (6.8% missing data).

Maximum Parsimony Analysis—Concordant with prior molecular studies in *Daucus* (Spalik and Downie 2007;

Spalik et al. 2010; Spooner et al. 2013; Arbizu et al. 2014b; Lee and Park 2014), our MP tree is highly resolved with three main clades, A, A' and B, all with 100% bootstrap support (Suppl. Figure 1). Clade A' contains all species possessing $2n = 18$ chromosomes, while the remaining taxa in clades A and B possess $2n = 16, 20, 22$ and 44 chromosomes, and the cladistic relationships match those of Arbizu et al. (2014b).

Within clade B, our MP tree recovered four major clades, three of which were found by Arbizu et al. (2014b), and one additional clade including samples not previously studied. All four of these *D. guttatus* species complex clades had high bootstrap values of 98–100%. Like the results of Arbizu et al. (2014b), all accessions of *D. involucratus* form a clade, but in the present analysis they are internal to other members of clade 3 rather than sister to it. Also unlike Arbizu et al. (2014b), the single accession of *D. glochidiatus* is sister to *D. littoralis* and other members of *D. guttatus* clade 1 in the present study, whereas *D. glochidiatus* was sister to *D. pusillus* and other members of *D. guttatus* clade 2 in Arbizu et al. (2014b). All other species relationships between this study and Arbizu et al. (2014b) are the same.

The use of additional accessions of the *D. guttatus* complex in this study revealed a geographical component with the two accessions of *D. guttatus* clade 1 collected in Israel (PI 279763 and PI 295858; 100% bootstrap). Also, clade 2 contains all accessions collected in Greece, except for two accessions collected in Turkey (Ames 25814, PI 652360). Clade 3 contains accessions only collected in Turkey, except PI 652332 (*D. involucratus*).

Maximum Likelihood Analysis—The ML tree using separate models for every single gene with 1,000 bootstrap replicates (Fig. 1) has the same overall topology as the MP tree (Suppl. Figure 1), including the geographic subset in clade 1. It also resolves the same clades, A, A' and B. Furthermore, there are good bootstrap support values in most branches of this tree. As in the MP analysis, *D. involucratus* was embedded in clade 3.

Bayesian Concordance Analysis—Results from our pruned analysis showed good mixing and good results as the standard deviation of concordance factors was less than 0.005. The primary concordance tree (Fig. 2), estimated for ten genes and 23 accessions with Bayesian concordance analyses, showed two differences in the topology compared to the MP and ML trees: (1) *D. glochidiatus* is sister to a clade formed by accessions of clades 1, 3, 4, and *D. littoralis*; (2) *D. pusillus* is sister to all species mentioned in (1). *Daucus pusillus* and *D. glochidiatus* are supported by one gene as sister clades of the species mentioned above. *Daucus involucratus* was also embedded in clade 3. There were no differences among the concordance factors using the four different prior probabilities on gene tree incongruence (α value). Clades A and B are supported by 7.2 and 7 genes, respectively. Less than four genes support clades 1, 2, and 4. There are 5.4 genes supporting clade 3.

Species Tree Estimation—Under the multispecies coalescent approach, the Bayesian analysis using 84 accessions containing 18 species and ten genes with their corresponding model of evolution gave very high ESS values (higher than 1000) for most of the parameters. Concordant with MP, ML, and BCA tree topologies, there are two main clades of *Daucus*, A and B (Fig. 3). Values of support (i.e. posterior probabilities) are slightly higher for many branches (Fig. 3)

compared to the species tree of Arbizu et al. (2014b). The species tree resulted in a similar topology to the MP and ML. However, there are three notable exceptions: (1) *D. glochidiatus* is no longer sister to clade 1 and *D. littoralis*, but rather, with low support, to clade 3 (*D. conchitae* with *D. involucratus*); (2) Clade 2 and *D. pusillus* form a clade with low posterior probability with clade 1, *D. littoralis*, and clade 4; and (3) *D. aureus* and *D. muricatus* form a single clade supported by a high value of support.

Morphological Analyses—Our morphological analyses considered the accessions grouped in the four discussed clades within clade B as distinct taxa as determined by the MP and ML analyses (except that *D. involucratus*, as discussed, was considered separately). Graphical analyses of the 41 character state distributions examined in this study are shown in Suppl. Figure 2. Patterns of variation were very similar to the studies reported by Spooner et al. (2014) and Arbizu et al. (2014a). That is, while a few characters showed little to no variation, most characters showed tremendous variation within all taxa and overlap of ranges across species. However, a combination of morphological characters was useful in identifying species within clade B.

Twenty-two of the 31 continuous variables were significant in the *F*-test, $p \leq 0.05$ in the stepwise discriminant analyses in at least one of the two analyses. The following nine were not significant: leaf width at maximum, pigmented central umbellules diameter, bract length, bract width, length of shortest peripheral ray, peripheral petal length, stamen length, stylopodium length, and vittae width.

Canonical variate analysis was reported to be a useful discriminant method to distinguish taxa of *Daucus* (Spooner et al. 2014; Arbizu et al. 2014a). Furthermore, Arbizu et al. (2014a) showed that HCA produced a tree based on average similarity of all characters of the three subsets of *D. guttatus*. Two different analyses were conducted using discriminant approaches and we present the *F*-test *p* values of the characters retained in all stepwise discriminant analyses (Table 1). Canonical variate analysis separated all taxa easily into six groups: the four *D. guttatus* complex clades, *D. involucratus* and *D. littoralis* (Fig. 4A). Stepwise discriminant analysis identified 21 of the 31 continuous characters as significant discriminators for this data set within all taxa at the *F*-test *p* value ≤ 0.05 (Table 1, column 3). Character state distributions of the six most significant discriminators are displayed in Fig. 5 and as expected, character state overlap even within these best discriminators was common. Eight of the 21 significant characters are mericarp characters. However, an analysis of only mericarp characters failed to clearly separate them into their corresponding groups (Fig. 4B), except accessions from clade 3, *D. involucratus* and *D. littoralis*. Stepwise discriminant analysis identified five of the 10 continuous characters as significant discriminators within all taxa at the *F*-test *p* value ≤ 0.05 (Table 1, column 4). Only one accession of clade 4 (Ames 25833) is intermingled within clade 2. Separation of accessions of clades 1, and 2 is not obvious since there is one accession of these two clades placed in other groups (Fig. 4B). The HCA shows a different grouping pattern. It did not cluster all taxa with clear resolution. *Daucus littoralis* is considered the most distinctive species (Suppl. Figure 3).

Species Names in the *Daucus guttatus* Complex—Based on a comparison of our molecular and morphological results with the original species descriptions, and the

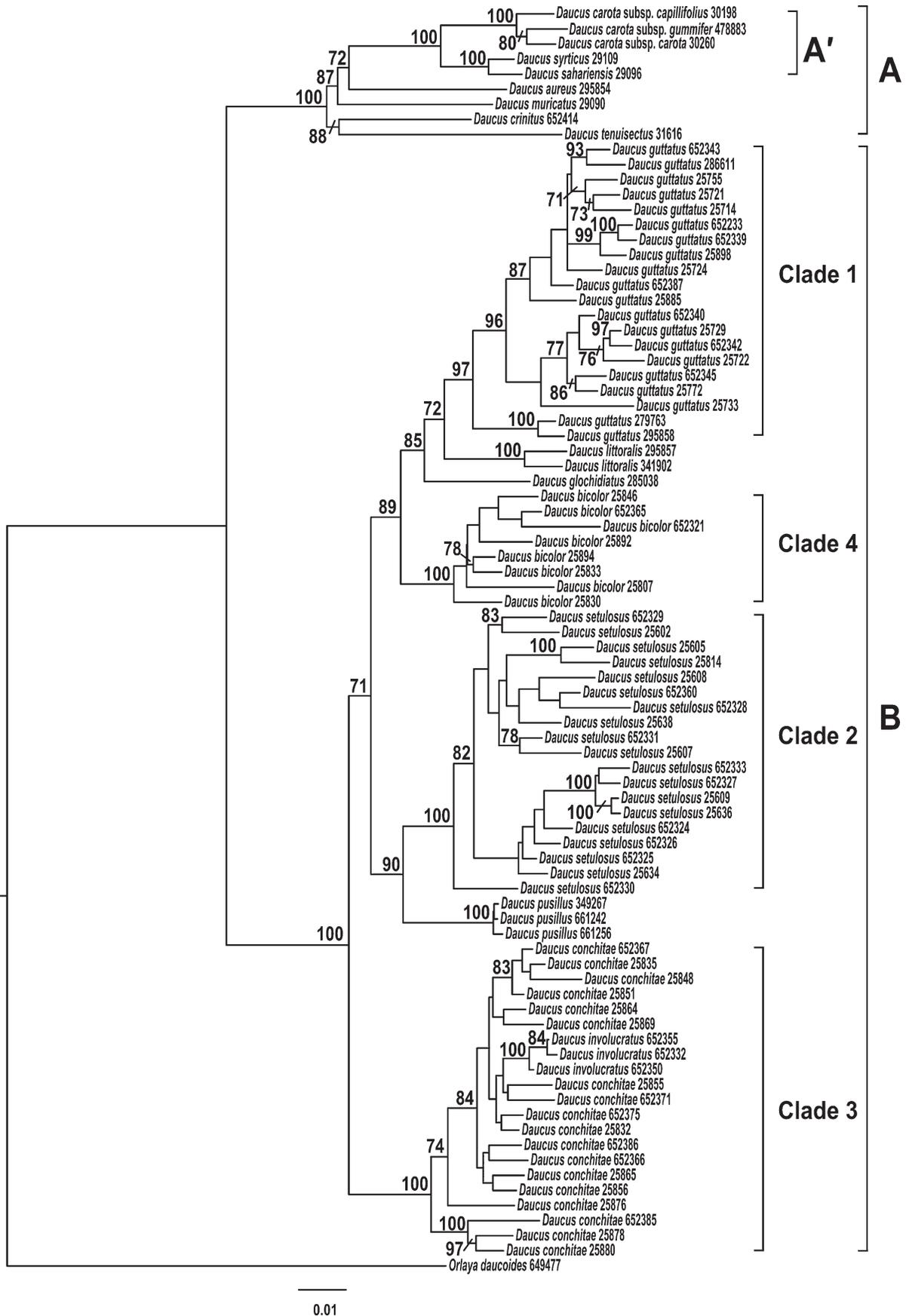


FIG. 1. Maximum likelihood phylogenetic reconstruction of the *Daucus guttatus* complex including species that belong to Clades A and B using 10 nuclear orthologs. Numbers above branches represent bootstrap values. Clades 1, 2, and 3 were identified in Arbizu et al. (2014b). Clade 4 is a new finding in the present study.

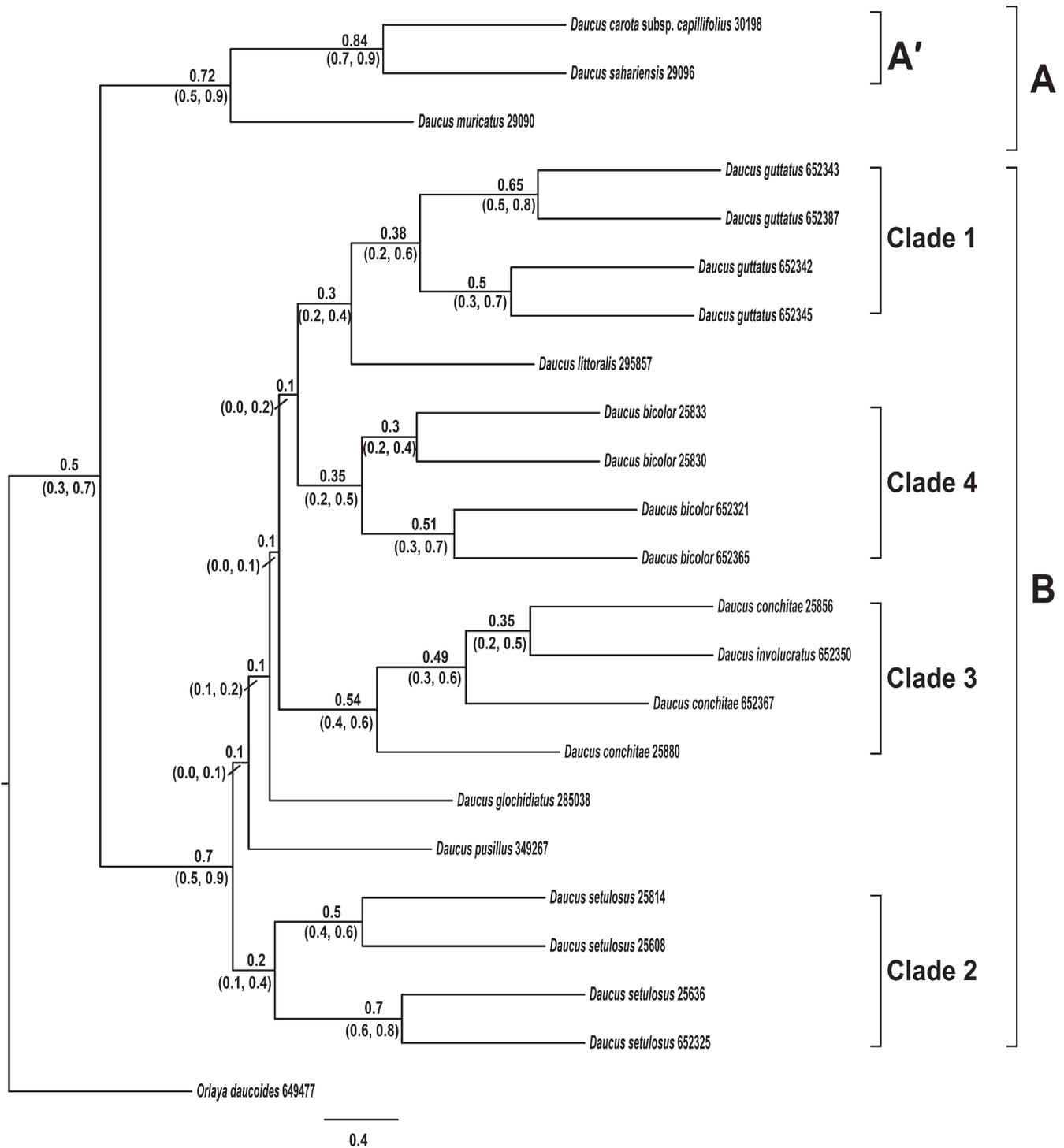


FIG. 2. Primary concordance tree of the *Daucus guttatus* complex including species that belong to Clades A and B obtained with Bayesian concordance analysis using 10 nuclear orthologs and 23 accessions. Numbers above the branches are the concordance factors, which do not show significant differences for different α values (0.1, 1, 10, and infinite). Clades 1, 2, and 3 were identified in Arbizu et al. (2014b). Clade 4 is a new finding in the present study.

companion study of Martínez-Flores et al. (2016), we recognize the following four species (and indented synonyms) for the *D. guttatus* complex: *D. guttatus* (Clade 1), *D. bicolor* (Clade 4), *D. setulosus* (Clade 2), and *D. conchitae* (Clade 3, excluding *D. involucratus*).

DAUCUS GUTTATUS Sm. in Sibth. & Sm., Fl. Graec. Prodr. 1(1): 184. 1806.

Daucus broteri Ten., Syll. Pl. Neapol. App. 3. IV. 1830.—TYPE: ITALY. Aprutti, M. Tenore. (lectotype, designated by Martínez-Flores et al. (2016), NAP Erbario Tenoreano campioni 7, fasc. 32, record number 3906 [digital photograph!]).

Daucus hirsutus DC., Prodr. 4: 213. 1830.—TYPE: Imprecise, including a wide area in the eastern Mediterranean

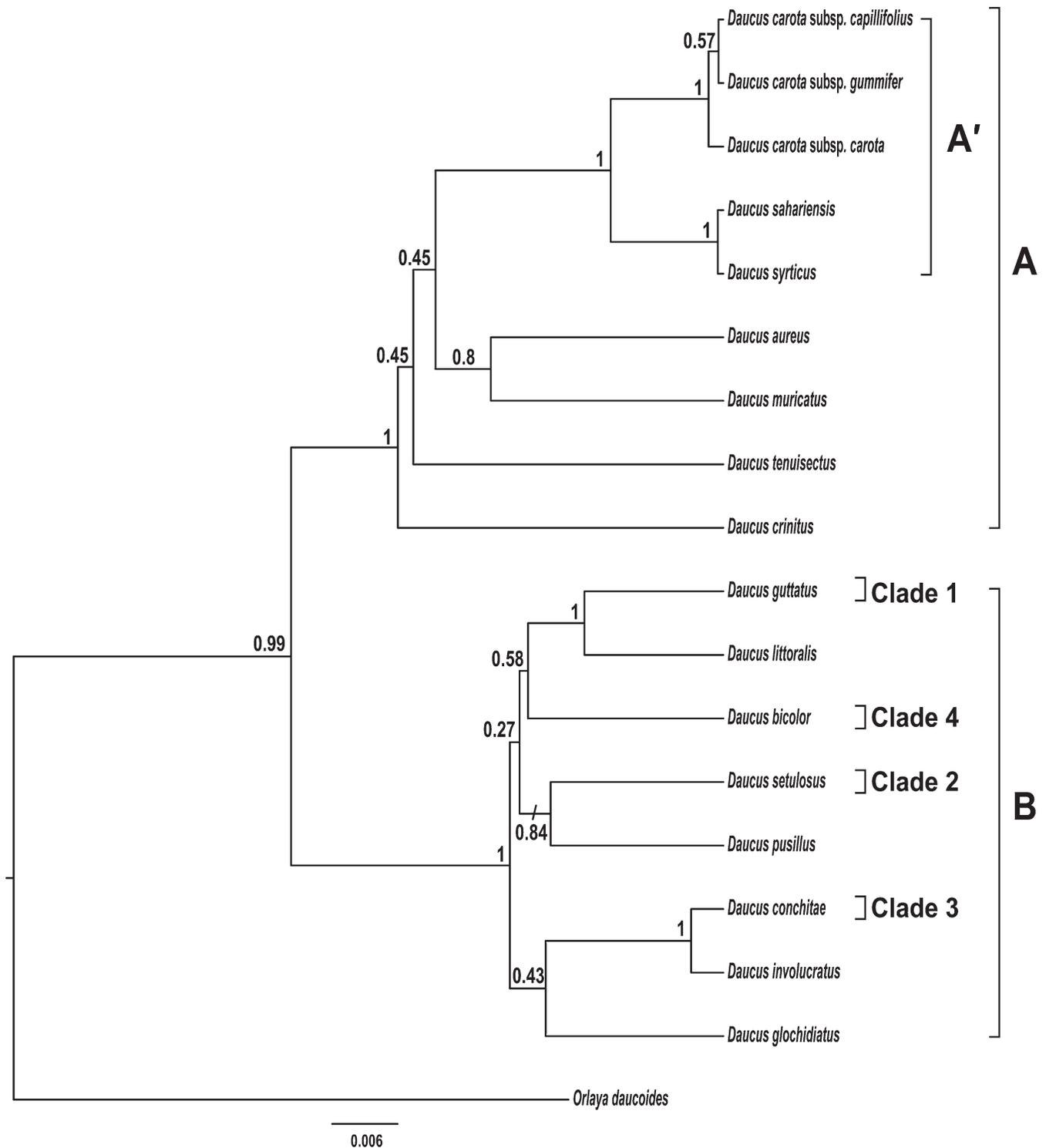


FIG. 3. Species tree of the *Daucus guttatus* complex including species that belong to Clades A and B based on a coalescent analysis using 84 accessions of 18 species of *Daucus* and outgroup using the models of evolution obtained by jModelTest. Numbers above branches are posterior probabilities. Clades 1, 2, and 3 were identified in Arbizu et al. (2014b). Clade 4 is a new finding in the present study.

basin ["frequens in ins. Archipelagi et Asià minore"], specimen not located. (designated by Martínez-Flores et al. 2016, [Icon in] Smith (1806), *Flora Graeca Prodr.*, Tab. 269; excluding explicitly the drawings of the mericarp).

Daucus michelii Caruel, *Prodr. Fl. Tosc.* 292. 1862.—TYPE: ITALY. Florence: Montici, Aug 1855, *T. Caruel.*—Nomen

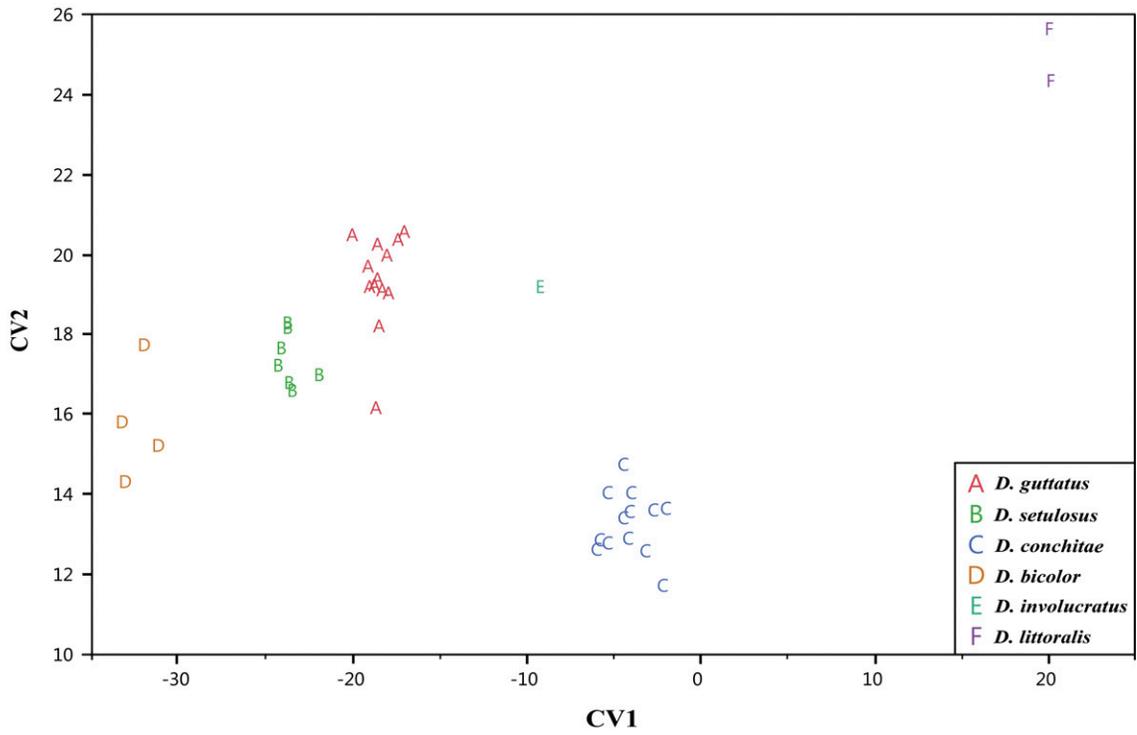
illegitimum without nomenclatural value (Martínez-Flores et al. 2016).

DAUCUS BICOLOR Sm. in Sibth. & Sm., *Fl. Graec. Prodr.* 1(1): 184. 1806.

[?] *Daucus microsciadius* Boiss., *Diagn. Pl. Orient.* 10: 47. 1849.

DAUCUS SETULOSUS Guss. ex DC., *Prodr.* 4: 211. 1830.

A



B

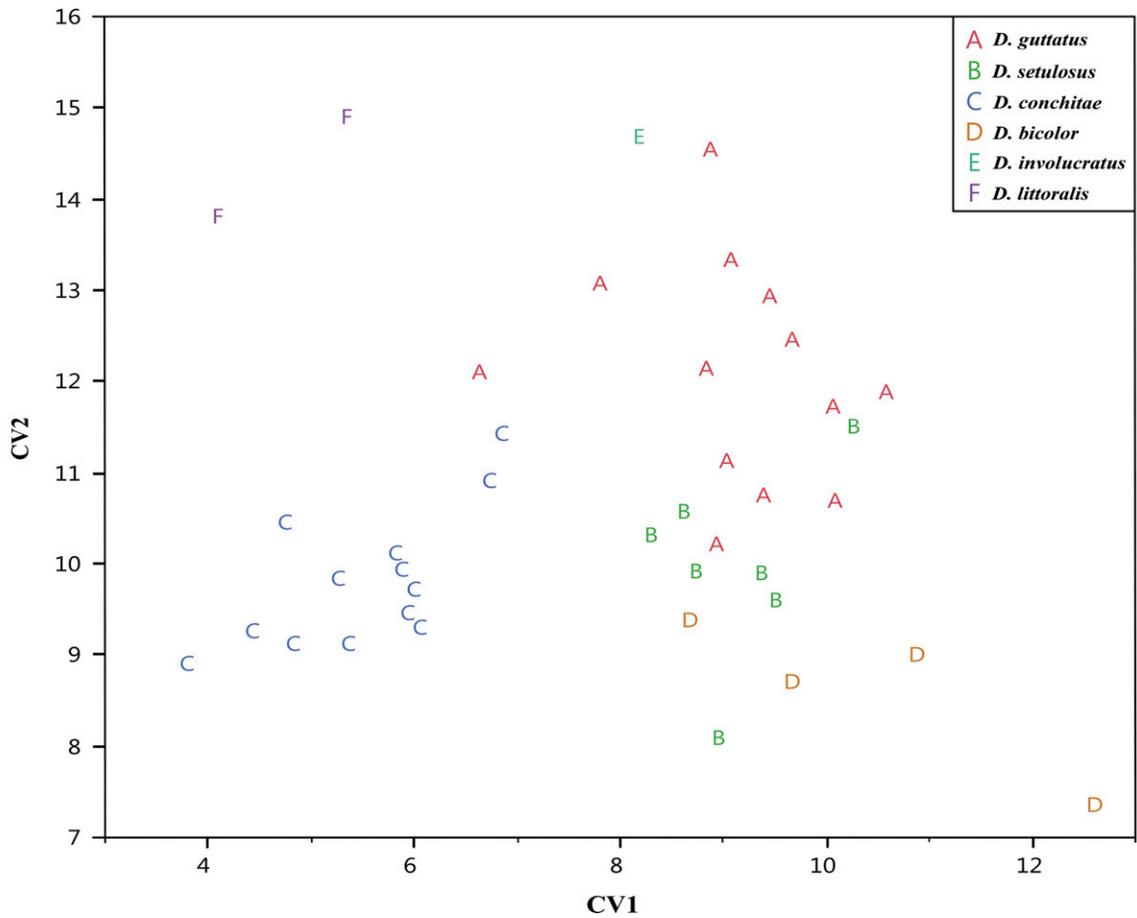


FIG. 4. Canonical variate analyses. A. The *Daucus guttatus* complex, *D. involucratus*, and *D. littoralis* using 31 continuous morphological characters that belong to plant, leaf, flower, and mericarp structures. B. The *Daucus guttatus* complex, *D. involucratus*, and *D. littoralis* using 10 continuous morphological characters that belong to mericarp structure.

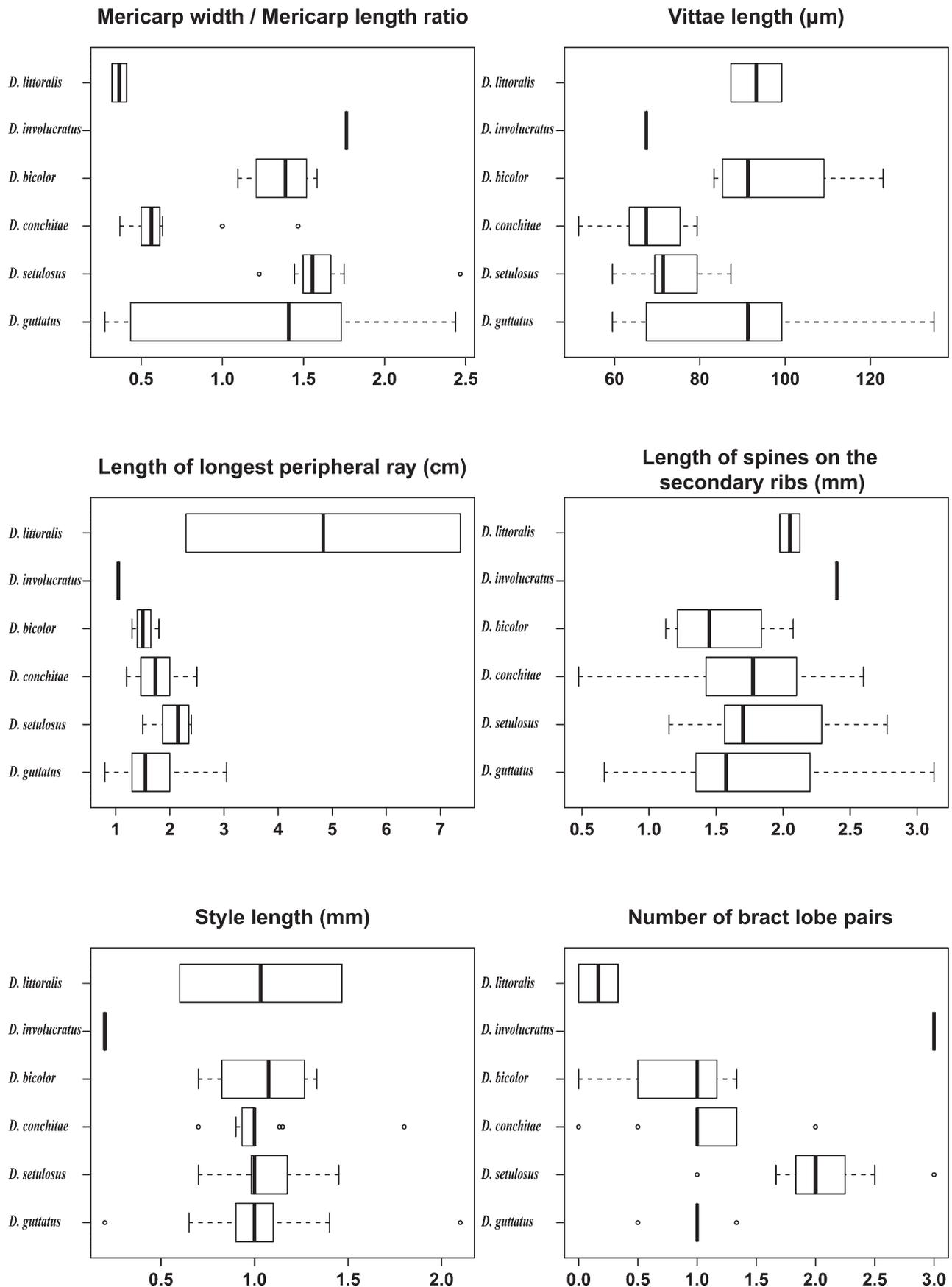


FIG. 5. Box plots of the six most significant morphological characters obtained by canonical variate analysis that best distinguish species of the *Daucus guttatus* complex, *D. involucratus*, and *D. littoralis*. The box plots display individual plant values for median, 25% and 75% percentile, range, and outliers.

Daucus arenicola Pancic ex Boiss., Fl. Orient. 2: 1075. 1872.—Nomen nudum without nomenclatural value (Okeke 1978)].
 [?]*Daucus ponticus* Velen. in Sitz. Boehm. Ges. Wiss. 1889: 34. 1890.
Daucus scabrosus Bert. ex DC., Prodr. 4: 211. 1830.—Nomen nudum without nomenclatural value (Martínez-Flores et al. 2016).

Daucus speciosus Ces. in Linnaea 11: 322. 1837.—TYPE: GREECE: Ex hb. Friedrichsthalii. n. 1186. (lectotype, designated by Martínez-Flores et al. 2016, RO-HC [digital photograph].

DAUCUS CONCHITAE Greuter in Willdenowia 8: 574. 1979.

KEY TO *D. BICOLOR*, *D. CONCHITAE*, *D. GLOCHIDIATUS*, *D. GUTTATUS*, *D. INVOLUCRATUS*, *D. LITTORALIS*,
D. PUSILLUS AND *D. SETULOSUS* (CLADE B)

1. Bracts mostly pinnatifid.
 2. Umbels subsessile *D. glochidiatus*.
 2. Umbels long pedunculate.
 - a. Style length 0.7–1.5 mm, peripheral petal length generally 1.0–2.0 mm, central umbellule concolorous or differently pigmented *D. setulosus*.
 3. Style length 0.3–0.6 mm, peripheral petal length less than 1.0 mm, central umbellule always concolorous.
 4. Peduncles very scabrous, number of rays typically 10–30 *D. pusillus*.
 4. Peduncles with few soft hairs, number of rays less than 11 *D. involucratus*.
1. Bracts mostly trifid or entire.
 5. Involucral bracts distinctly shorter than the mature umbel, mericarp length 4.5–5.5 mm, central umbellule always concolorous *D. littoralis*.
 5. Involucral bracts as long as or distinctly longer than mature umbel, mericarp length generally 2.0–4.0 mm, central umbellule concolorous or differently pigmented.
 6. Bracts mostly entire, mericarp vittae always elliptical *D. conchitae*.
 6. Bracts mostly trifid, mericarp vittae triangular (if mericarp spines strongly widened at bases, mericarp vittae elliptical).
 7. Involucral bracts longer than mature umbel, several central umbellets typically dark-purple, mericarp length 2.0–2.5 mm *D. bicolor*.
 7. Involucral bracts as long as mature umbel, central umbellule and other umbellets typically with few central dark-purple colored flowers, mericarp length 2.0–4.5 mm *D. guttatus*.

DISCUSSION

Phylogenetic Analyses—Previous phylogenetic studies (Spalik and Downie 2007; Spalik et al. 2010; Spooner et al. 2013; Arbizu et al. 2014b; Lee and Park 2014) demonstrated two main clades of *Daucus*, here labeled as clades A and B. Arbizu et al. (2014b) showed that clade A consists of (i) clade A', which contains all subspecies of *Daucus* and *D. sahariensis* and *D. syrticus*, all with $2n = 18$ chromosomes; and (ii) other *Daucus* species and related genera with $2n = 16, 20, 22$ chromosomes. Clade B contains species of the *D. guttatus* complex and close relatives with $2n = 20, 22, 44$. Arbizu et al. (2014b) demonstrated that a subset of ten of 94 nuclear orthologs provided effective discriminators for species boundaries and relationships of *Daucus* but highlighted unresolved species boundaries in the *D. guttatus* complex that were partitioned into three subclades within clade B. Our present study explored this problem through the use of additional accessions of the *D. guttatus* complex and discovered them to now be partitioned into the original three clades and one additional clade (now four). Morphological analyses of these accessions provided support for these as four species of the *D. guttatus* complex (*D. guttatus*, *D. bicolor*, *D. setulosus* and *D. conchitae*), but with character-state overlap among the species. Concatenated analyses and BCA placed all accessions of *D. involucratus* together, but internested entirely within *D. conchitae* (clade 3 of the *D. guttatus* complex). This can be explained by the use of a reduced number of nuclear orthologs (10 vs. 94) compared to the study of Arbizu et al. (2014b).

The primary concordance tree (PCT, Fig. 2) produced a topology slightly different from MP and ML analyses. BCA does not assume that genes have the same topology (Ané et al. 2007). Concordance factors (i.e. genomic support) are

less than 50% for the clades 1, 2, and 4, showing discordant gene histories in the genomes of *Daucus*, as reported by Arbizu et al. (2014b). On the other hand, the PCT analysis shows very high concordance factors for the clade containing *D. carota* subsp. *capillifolius* and *D. sahariensis*. In addition, the clade grouping two accessions of clade 1 (PI 652387 and PI 652343), and clade 2 (Ames 25636 and PI 652325) also have high concordance factors indicating less discordant histories in these genomes. The most likely factors that may be causing discordance are recombination, hybridization and introgression (Rieseberg et al. 2000), and incomplete lineage sorting (Pamilo and Nei 1998). The species tree topology is similar to the PCT.

The multispecies coalescent approach (Fig. 3) assumes that genealogical discordance is entirely due to incomplete lineage sorting, the most common cause of discordance among gene trees (Maddison and Knowles 2006; Degnan and Rosenberg 2009; Baum and Smith 2013). We place greater weight on the BCA results since it makes no assumption about the reason of genealogical discordance (Larget et al. 2010). The species tree obtained by Arbizu et al. (2014b) included 104 accessions of 27 species of *Daucus* and outgroups and showed a topology that was not congruent with the concatenated analyses of that study. In the present study, we report a species tree different from the species tree of Arbizu et al. (2014b). This can be explained by the following factors: (1) ESS values for all parameters are very high in our present study (higher than 500) indicating good mixing; and (2) we used ten selected genes that resemble the dominant topology of *Daucus*.

Morphological Analyses—Discriminant analysis was used by Spooner et al. (2014) and Arbizu et al. (2014a) to explore the validity of species and subspecies within *Daucus*, showing good discrimination except for the subspecies of

D. carota. The taxonomic value of fruits in the Apiales have been reported by Lee et al. (2001), Liu (2004), Baranski et al. (2006), Liu et al. (2006) Akalin Uruşak and Kizilarıslan (2013), Mezghani et al. (2014), and Tavares et al. (2014). Sáenz (1981), Pujadas Salvà (2003), Baranski et al. (2006), and Mezghani et al. (2014) indicated that fruits provided critical taxonomic characters distinguishing taxa within *Daucus* since fruits tend to show less phenotypic plasticity in comparison with other plant organs. Other studies questioned the utility of fruit characters alone as a practical tool to classify taxa in the Apiales (Downie and Katz-Downie 1999; Katz-Downie et al. 1999; Spalik et al. 2001; Vallejo-Roman et al. 2002; Lowry et al. 2004). We investigated this issue in the present study by examining only the 10 continuously variable fruit characters. Our discriminant analyses identified a subset of five mericarp characters that are most useful to distinguish *D. conchitae*, *D. involuocratus* and *D. littoralis* (Table 1, Fig. 4B), but not the other species. Thus, a combination of mericarp and other vegetative and floral characters are needed to distinguish members of the *D. guttatus* complex (but with some character overlap). Using 11 accessions from the *D. guttatus* complex, Arbizu et al. (2014a) could separate three species of this complex using HCA. On the other hand, in the present study our HCA could not separate the four species of the *Daucus guttatus* complex (Suppl. Figure 3). This difference can be explained by the greater number of accessions (37 in total for the *D. guttatus* complex) used in the present study.

Association of Species Names to the *Daucus guttatus* Complex (clades 1, 2, 3, and 4)—To date, the species boundaries and application of existing names to members of the *D. guttatus* complex have remained unclear. Our molecular analyses identified four well-supported clades of members of this complex and showed that they had morphological support, but only with multiple characters that exhibited some overlapping sets of character states (Suppl. Figure 2). Members of clade B possess chromosome numbers of $2n = 20, 22$ and 44 , but many of them are from few counts. The accuracy of $2n = 20$ and 22 counts and their possible association with clades in the *D. guttatus* complex need further investigation with additional accessions.

The original descriptions of *D. bicolor* and *D. guttatus* by Smith (1819) discriminated these two species by differences in plant height, degree of branching, pubescence, bract size, and color pattern of central umbellets. *Daucus broteri* was originally described by Tenore (1830) as a plant with hairy leaves and tripinnate, compact umbels, multiple umbel rays, almost all bracts 3-fid, wide fruit spines confluent at the base, and 6–8 spines per ridge. Greuter (1979) described *D. conchitae* based on plants without fruits. He described it as a plant branched at the base, with five to seven simple bracts with a length of 15 mm, exceeding the umbel perimeter. He mentioned a distinguishing character of *D. conchitae* as simple or entire bracts. *Daucus setulosus* (de Candolle 1830) was described as a plant with stems branched and setulose, leaves pinnatisect with various segments, and bracts with multiple segments. Different authors have considered *D. setulosus* as a synonym of *D. guttatus* and/or *D. bicolor*. We here quantify tremendous variation in all of the above characters. Our analysis showed that these species are similar, with sometimes overlapping character states, but with the best discriminating characters as shown in our key to

be involuocral bract length and shape, primary umbel diameter, coloring of the central umbellets, mericarp size, vittae shape, and degree of confluency of the spines. Clearly, discrimination of species in clade B is difficult and relies on the use of a suite of characters that overlap in range (a polythetic approach). Ultimately, we associate species names with the clades by examining characters of the majority of accessions in these clades and comparing to type specimens (Martínez-Flores et al. 2016).

Flower color variation is a model for the experimental study of evolution. Clegg and Durbin (2000) found that almost all mutations that determine phenotypic differences in *Ipomoea purpurea* (L.) Roth flower color variations are the result of transposon insertions. This may be a plausible explanation for the case of dark-colored pattern present in *D. guttatus* clades 1, 2, 4. Westmoreland and Muntan (1996), and Goulson et al. (2009) favored the hypothesis that the dark flowers in umbels of *Daucus* significantly increase both insect attraction and fruit production. Conversely, Polte and Reinhold (2013) demonstrated that the dark central floret in wild *Daucus* has no role in pollination, but may play a role in reducing parasite infestation by the gall midge *Kiefferia pericarpicola* (Bremi, 1847). This issue needs further research to clarify the role of purple-colored flowers in *Daucus* species.

Most of the accessions examined in clade 1 were collected in Syria very close to the collection conducted by Al-Safadi (2008) (Slanfah, Salma and Rabiaa, and near Tartus). Interestingly, accessions that matched the original description of *D. guttatus* were placed in clade 1, including the epitype of *D. guttatus* (PI 652342). Accessions of clade 4 match the original description of *D. bicolor*. They look similar to *D. bicolor* due to the dark-colored pattern in the umbels, but bracts do not always exceed the umbel perimeter. However, bracts are 3-fid. Peripheral petal length is longer than in *D. guttatus*, *D. setulosus* and *D. conchitae*. Leaf, petiole, and peduncle pubescence is smoother. Mericarp length and width is shorter and has a triangular vittae shape. As a consequence, we identified clades 1 and 4 as *D. guttatus* and *D. bicolor*, respectively.

There are accessions in Italy that have been reported as *D. broteri* and were originally misidentified as *D. muricatus* because of the wide mericarp spine confluent at the base (Martínez-Flores et al. 2016). Based on morphological evidence only, it is possible that those accessions reported in Italy may be *D. guttatus* or yet another species in this complex. Sáenz (1981) classified *D. muricatus* and *D. bicolor*, and *D. guttatus* within sections *Platyspermum* and *Daucus*, respectively. She considered *D. broteri* and *D. setulosus* as synonyms of *D. bicolor* and *D. guttatus*, respectively. We suggest that samples examined by Sáenz (1981) for *D. bicolor* are one form of *D. guttatus* with spines strongly widened at the base, confluent into a crest (the same as Ames 25733, PI 652340, among others). In addition, Sáenz (1981) described *D. guttatus* as a species with bracts divided and exceeding the umbel perimeter, spines slightly confluent at the base, and large triangular shaped vittae. However, we propose samples examined are *D. setulosus* as they match the bract description reported by de Candolle (1830) and our mericarp description for clade 2. Therefore, we conclude accessions of clade 2 correspond to the species *D. setulosus*.

Daucus involuocratus is embedded within clade 3 and was previously reported as a close relative to *D. guttatus* 3

(Arbizu et al. 2014b; here named as clade 3; chiefly *D. conchitae*). Greuter (1979) indicated that *D. conchitae* resembles *D. involucratus* in general habit. On the contrary, Sáenz (1981) indicated *D. conchitae* appeared to be a taxon close to *D. littoralis* because of their anatomical characters. Aytaç and Duman (2013) collected individuals in Antalya, Turkey, with simple bracts and identified them as *D. conchitae*. Constantinidis (2013) evaluated fruit morphology of all plants collected in the Kastellorizo Island group and concluded *D. conchitae* is simply a form of *D. guttatus*. He reported his collection has individuals with simple and pinnatifid bracts, usually in the same umbel, and fruit spines weakly confluent at the base. His descriptions are very similar to the accessions examined here placed in clade 3. Surprisingly, molecular analyses placed *D. conchitae* as a close relative of *D. involucratus* (Spalik and Downie 2007; Lee and Park 2014). We here infer that most accessions of clade 3 belong to *D. conchitae* (with specimens we identify as *D. involucratus* also nested within this clade, as discussed). Vittae shape has been considered as a discriminant character within taxa of *Daucus* (Thellung 1926; Sáenz 1981; Heywood 1982). We documented the mericarp of *D. conchitae* to have elliptical vittae. This characteristic was not present in any other species examined here, except the accessions of *D. guttatus* with large spines confluent at the base forming a wide crest, and *D. littoralis*.

The use of wild crop relatives in plant breeding is likely to intensify as breeding techniques improve and adaptation to climate change becomes more pressing (Khoury et al. 2013). There is a constant demand for development of new hybrid carrot cultivars (Rubatzky et al. 1999; Frese and Nothnagel 2008) and wild *Daucus* species have utility for breeding against pest and disease tolerance or resistance, yield increase, male sterility, and for nutraceutical and culinary traits (Camadro et al. 2008). To better use germplasm maintained at different germplasm banks, it is imperative to determine the species boundaries and phylogenetic relationships of *Daucus*. Wild *Daucus* species possessing $2n = 18$ chromosomes are not the only source of genetic resources for carrot. *Daucus pusillus* Michx. ($2n = 22$), a wild carrot species from the American continent, was tested to establish the feasibility of hybridization with the cultivated carrot under controlled pollination, resulting in the possibility of obtaining interspecific hybrids (Camadro et al. 2008). *Daucus guttatus*, with $2n = 20$ chromosomes, may also be used as a potential breeding source. It has been reported to have prominent antibacterial activity against *Corynebacterium pyogenes* (Radulović et al. 2011), and (as *D. broteri*) to have resistance to carrot fly, *Psila rosae* (Fabricius, 1794) (Hardman et al. 1990). *Daucus littoralis* ($2n = 20$), a species closely related to *D. guttatus*, shows anti-mycobacterial activity against one of the most lethal infections worldwide, tuberculosis (Başer et al. 2008). Indeed, this species can be considered a good candidate for the potential production of new carrot hybrids with new useful bioactive components.

ACKNOWLEDGMENTS. We thank Kathleen Reitsma at the North Central Regional Plant Introduction Station in Ames, Iowa for providing germplasm. This paper represents partial fulfillment of a Ph.D. degree for CA in Plant Breeding and Plant Genetics at the University of Wisconsin-Madison. CA is partly funded by the National Council of Science and Technology of Perú (Concytec, by its initials in Spanish). We appreciate the comments of two anonymous reviewers and Editor-in Chief James Smith that greatly improved our manuscript.

LITERATURE CITED

- Akalin Uruşak, E. and C. Kizilarlan. 2013. Fruit anatomy of some *Ferulago* (Apiaceae) species in Turkey. *Turkish Journal of Botany* 37: 434–445.
- Al-Safadi, B. 2008. Characterization and distribution of *Daucus* species in Syria. *Biologia* 63: 177–182.
- Ané, C., B. Larget, D. A. Baum, S. D. Smith, and A. Rokas. 2007. Bayesian estimation of concordance among gene trees. *Molecular Biology and Evolution* 24: 412–426.
- Arbizu, C., K. R. Reitsma, P. W. Simon, and D. M. Spooner. 2014a. Morphometrics of *Daucus* (Umbelliferae): A counterpart to a phylogenomics study. *American Journal of Botany* 101: 2005–2016.
- Arbizu, C., H. Ruess, D. Senalik, P. W. Simon, and D. M. Spooner. 2014b. Phylogenomics of the carrot genus (*Daucus*, Apiaceae). *American Journal of Botany* 101: 1666–1685.
- Ayres, D., A. Darling, D. Zwickl, P. Beerli, M. T. Holder, P. O. Lewis, J. Huelsenbeck, F. Ronquist, D. L. Swofford, M. P. Cummings, A. Rambaut, and M. A. Suchard. 2012. Beagle: An application programming interface and high-performance computing library for statistical phylogenetics. *Systematic Biology* 61: 170–173.
- Aytaç, Z. and H. Duman. 2013. A new species and 2 new records from Turkey. *Turkish Journal of Botany* 37: 1055–1060.
- Baranski, R., M. Baranska, H. Schulz, P. W. Simon, and T. Nothnagel. 2006. Single seed raman measurements allow taxonomical discrimination of Apiaceae accessions collected in gene banks. *Biopolymers* 81: 497–505.
- Başer, K. H. C., M. Kurkcuoğlu, T. Askun, and G. Tumen. 2008. Anti-tuberculosis activity of *Daucus littoralis* Sibth. et Sm. (Apiaceae) from Turkey. *The Journal of Essential Oil Research* 21: 572–575.
- Baum, D. A. and S. D. Smith. 2013. *Tree thinking: An introduction to phylogenetic biology*. Ed. 1. Greenwood Village: Roberts and Co.
- Bazinnet, A. L., D. J. Zwickl, and M. P. Cummings. 2014. A gateway for phylogenetic analysis powered by grid computing featuring GARLI 2.0. *Systematic Biology* 63: 812–818.
- Bradeen, J. M., I. C. Bach, M. Briard, V. le Clerc, D. Grzebelus, D. A. Senalik, and P. W. Simon. 2002. Molecular diversity analysis of cultivated carrot (*Daucus carota* L.) and wild *Daucus* populations reveals a genetically nonstructured composition. *Journal of the American Society for Horticultural Science* 127: 383–391.
- Camadro, E. L., M. A. Cauhépe, and P. W. Simon. 2008. Compatibility relations between the edible carrot *Daucus carota* and *D. pusillus*, a related wild species from the Argentinian Pampas. *Euphytica* 159: 103–109.
- Clegg, M. T. and M. L. Durbin. 2000. Flower color variation: A model for the experimental study of evolution. *Proceedings of the National Academy of Sciences USA* 97: 7016–7023.
- Cohen, J. I. 2014. A phylogenetic analysis of morphological and molecular characters of Boraginaceae: evolutionary relationships, taxonomy, and patterns of character evolution. *Cladistics* 39: 139–169.
- Constantinidis, T. 2013. The flora of the Kastellorizo island group (East Aegean Islands, Greece): New records and comments. *Flora Mediterranea* 23: 69–86.
- Cullen, J. 1972. *Daucus*. Pp 531–536 in *Flora of Turkey and the East Aegean Islands* vol. 4, ed. P. H. Davis. Edinburgh: Edinburgh University Press.
- Darriba, D., G. L. Taboada, R. Doallo, and D. Posada. 2012. jModelTest 2: Mode models, new heuristics and parallel computing. *Nature Methods* 9: 772.
- Dayrat, B. 2005. Towards integrative taxonomy. *Biological Journal of the Linnean Society. Linnean Society of London* 85: 407–415.
- de Candolle, A. P. de 1830. *Prodromus systematis naturalis regni vegetabilis, sive enumeratio contracta ordinum, generum, specierumque plantarum hucusque cognitarum, juxta methodi naturalis normas digesta* vol. 6. Paris: Treuttel et Würtz.
- Degnan, J. H. and N. A. Rosenberg. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution* 24: 332–340.
- Downie, S. R. and D. S. Katz-Downie. 1999. Phylogenetic analysis of chloroplast *rps16* intron sequences reveals relationships within the woody southern African Apiaceae subfamily Apioideae. *Canadian Journal of Botany* 77: 1120–1135.
- Doyle, J. J. and J. L. Doyle. 1990. Isolation of plant DNA from fresh tissue. *Focus (San Francisco, Calif.)* 12: 13–15.
- Drummond, A. J., M. A. Suchard, D. Xie, and A. Rambaut. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* 29: 1969–1973.
- Edgar, R. C. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32: 1792–1797.

- Emadzade, K., C. Lehnebach, P. Lockhart, and E. Hörandl. 2010. A molecular phylogeny, morphology and classification of genera of Ranunculaceae (Ranunculaceae). *Taxon* 59: 809–828.
- Farris, J. S. 1970. Methods for computing Wagner trees. *Systematic Zoology* 19: 83–92.
- Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39: 783–791.
- Fitch, W. 1971. Toward defining the course of evolution: minimum change for a specific tree topology. *Systematic Zoology* 20: 406–416.
- Frese, L. and T. Nothnagel. 2008. *Daucus* species promise a glossy future in carrot production. Pp. 4–6 in *Crop wild relative* vol. 6, ed. N. Maxted and E. Dulloo. Birmingham: IUCN/SSC.
- Gernhard, T. 2008. The conditioned reconstructed process. *Journal of Theoretical Biology* 253: 769–778.
- Goulson, D., K. Mcguire, E. E. Munro, S. Adamson, L. Colliar, K. J. Park, M. C. Tinsley, and A. S. Gilburn. 2009. Functional significance of the dark central floret of *Daucus carota* (Apiaceae) L.; is it an insect mimic? *Plant Species Biology* 24: 77–82.
- Greuter, W. 1979. The flora and phytogeography of Kastellorizo (Dhodonkanisos, Greece). *Willdenowia* 8: 531–611.
- Hardman, J. A., P. R. Ellis, and P. L. Saw. 1990. Further investigations of the host range of the carrot fly, *Psila rosae* (F). *Annals of Applied Biology* 117: 495–506.
- Heywood, V. H. 1968. *Daucus*. *Feddes Repertorium* 79: 1–68.
- Heywood, V. H. 1982. Multivariate taxonomic synthesis of the tribe Caucaleidae. Pp. 727–736 in *Les Ombellifères "Contributions Pluridisciplinaires a la Systematique"* vol. 6, ed. A. M. Cauwet-Marc and J. Carbonnier. Monographs in systematic botany from the Missouri Botanical Garden. Ann Arbor: Brumfield.
- Jafri, S. M. H. and A. El-Gadi. 1985. *Daucus*. Pp. 130–144 in *Flora of Libya* vol. 117, ed. S. M. H. Jafri, and A. El-Gadi. Tripoli: Al-Faateh University, Faculty of Science, Department of Botany.
- SAS Institute Inc. 1989–2007. JMP[®], v. 11.2. Cary, North Carolina.
- Jury, J. L. 2002. *Daucus*. Pp. 467–469 in *Catalogue des plantes vasculaires du Nord de Maroc, incluant des clés d'identification* vol. 1, ed. B. Valdés, M. Rejdali, A. Achhal El Kadmiri, J. L. Jury, and J. M. Montserrat. Madrid: Consejo Superior de Investigaciones Científicas.
- Katz-Downie, D. S., C. M. Vallejo-Roman, E. I. Terentjeva, A. V. Troitsky, M. G. Pimenov, B. Lee, and S. R. Downie. 1999. Towards a molecular phylogeny of *Apiaceae* subfamily *Apioideae*: Additional information from nuclear ribosomal DNA ITS sequences. *Plant Systematics and Evolution* 216: 167–195.
- Khoury, C. K., S. Greene, J. Wiersema, N. Maxted, A. Jarvis, and P. C. Struik. 2013. An inventory of crop wild relatives of the United States. *Crop Science* 53: 1–13.
- Larget, B., S. K. Kotha, C. N. Dewey, and C. Ané. 2010. BUCKY: Gene tree/species tree reconciliation with the Bayesian concordance analysis. *Bioinformatics* 26: 2910–2911.
- Le Floch, É., L. Boulos, and E. Vela. 2010. *Catalogue synonymique commenté de la flore de Tunisie*. Ed. 1. Tunis: République Tunisienne, Ministère de l'Environnement et du Développement Durable Banque Nationale de Gènes.
- Lee, B. Y. and C. Park. 2014. Molecular phylogeny of *Daucus* (Apiaceae): Evidence from nuclear ribosomal DNA ITS sequences. *Journal of Species Research* 3: 39–52.
- Lee, B. Y., G. A. Levin, and S. R. Downie. 2001. Relationships within the spiny-fruited umbellifers (Scandiceae subtribes Daucinae and Torilidinae) as assessed by phylogenetic analysis of morphological characters. *Systematic Botany* 26: 622–642.
- Liu, M. 2004. *A taxonomic evaluation of fruit structure in the family Apiaceae*. Ph. D. thesis. Johannesburg: Rand Afrikaans University.
- Liu, M., G. M. Plunkett, P. P. Lowry II, B. Van Wyk, and P. M. Tilney. 2006. The taxonomic value of fruit wing types in the order Apiales. *American Journal of Botany* 93: 1357–1368.
- Lowry, P. P., G. M. Plunkett, and J. Wen. 2004. Genetic relationships in Araliaceae: looking into the crystal ball. *South African Journal of Botany* 70: 382–392.
- Maddison, W. P. and D. R. Maddison. 2015. Mesquite: a modular system for evolutionary analysis, v. 3.03. <http://mesquiteproject.org>.
- Maddison, W. P. and L. L. Knowles. 2006. Inferring phylogeny despite incomplete lineage sorting. *Systematic Biology* 55: 21–30.
- Martínez-Flores, F., C. I. Arbizu, K. Reitsma, A. Juan, P. W. Simon, D. M. Spooner, and M. B. Crespo. 2016. Lectotype designation for seven species names in the *Daucus guttatus* complex (Apiaceae) from the Central and Eastern Mediterranean Basin. *Systematic Botany* 41: 464–478.
- Martínez-Flores, F., A. Juan, and M. B. Crespo. 2012. Molecular and morphological evidence supports the resurrection of a neglected species of *Daucus* (Apiaceae) from North Africa. *Botanical Journal of the Linnean Society* 169: 320–337.
- Mezghani, N., I. Zaouali, W. Amri, S. Rouz, P. W. Simon, C. Hannachi, Z. Ghrabi, M. Neffati, B. Bouzbidia, and D. M. Spooner. 2014. Fruit morphological descriptors as a tool discrimination of *Daucus* L. germplasm. *Genetic Resources and Crop Evolution* 61: 499–510.
- Miller, M. A., W. Pfeiffer, and T. Schwartz. 2010. Creating the CIPRES science gateway for inference of large phylogenetic trees. New Orleans: Proceedings of the Gateway Computing Environments Workshop (GCE).
- Mouterde, P. 1966. *Nouvelle Flore de Liban et de la Syrie*. Ed. 2. Beirut: Dar El-Machreq.
- Okeke, S. E. 1978. *Systematic studies in Daucus* L. (Umbelliferae). Ph. D. thesis. Reading: University of Reading.
- Pamilo, P. and M. Nei. 1998. Relationships between gene trees and species trees. *Molecular Biology and Evolution* 7: 231–240.
- Patterson, C., D. M. Williams, and C. J. Humphries. 1993. Congruence between molecular and morphological phylogenies. *Annual Review of Ecology, Evolution, and Systematics* 24: 153–188.
- Pisani, D., M. J. Benton, and M. Wilkinson. 2007. Congruence of morphological and molecular phylogenies. *Acta Biotheoretica* 55: 269–281.
- Polte, S. and K. Reinhold. 2013. The function of the wild carrot's dark central floret: Attract, guide or deter? *Plant Species Biology* 28: 81–86.
- Pottier-Alapetite, G. 1979. *Daucus*. Pp. 615–621 in *Flore de la Tunisie: Angiospermes-Dicotyledones Apetales – Dialypetales* vol. 2, ed. G. Pottier-Alapetite. Tunis: Ministère de l'Enseignement Supérieur et de la Recherche Scientifique et le Ministère de l'Agriculture.
- Pujadas Salvà, A. J. 2003. *Daucus*. Pp. 97–125 in *Flora Iberica: Plantas vasculares de la Península Ibérica e islas Baleares* vol. 10, ed. G. Nieto-Feliner, S. L. Jury, and A. Herrero. Madrid: Real Jardín Botánico, CSIC.
- Quézel, P. and S. Santa. 1963. *Daucus*. Pp. 659–663 in *Nouvelle Flore de l'Algérie et des Régions Désertiques Méridionales* vol. 2, ed. P. Quézel and S. Santa. Montpellier: Editions du Centre National de la Recherche Scientifique.
- R Core Team. 2015. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.
- Radulović, N., N. Dordević, and Z. Stojanović-Radić. 2011. Volatiles of the Balkan endemic *Daucus guttatus* ssp. *zahariadii* and cultivated and wild-growing *D. carota* – A comparison study. *Food Chemistry* 125: 35–43.
- Rambaut, A. and A. J. Drummond. 2014. Tracer, v. 1.6. <http://beast.bio.ed.ac.uk/Tracer>.
- Rieseberg, L. H., S. J. Baird, and K. A. Gardner. 2000. Hybridization, introgression, and linkage evolution. *Plant Molecular Biology* 42: 205–224.
- Rodríguez, F., D. Cai, Y. Teng, and D. Spooner. 2011. Asymmetric single-strand conformation polymorphism: An accurate and cost-effective method to amplify and sequence allelic variants. *American Journal of Botany* 98: 1061–1067.
- Ronquist, F., M. Teslenko, P. van der Mark, D. L. Ayres, A. Darling, S. Höhna, B. Larget, L. Liu, M. A. Suchard, and J. P. Huelsenbeck. 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61: 539–542.
- Rubatzky, V. E., C. F. Quiroz, and P. W. Simon. 1999. *Carrots and related vegetable Umbelliferae*. Ed. 1. New York: CABI.
- Sáenz, C. 1981. Research on *Daucus* L. (Umbelliferae). *Anales del Jardín Botánico de Madrid* 37: 481–533.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences USA* 74: 5463–5467.
- Simon, P. W. 2000. Domestication, historical development, and modern breeding of carrot. Pp. 157–190 in *Plant Breeding Reviews* vol. 19, ed. J. Janick. Oxford: John Wiley & Sons, Inc.
- Small, E. 1978. A numerical taxonomic analysis of the *Daucus carota* complex. *Canadian Journal of Botany* 56: 248–276.
- Smith, J. E. 1819. *Daucus*. Pp. 63–66 in *Flora Graeca: sive Plantarum rariorum historia, quas in provinciis aut insulis Graeciae* vol. 3, eds. J. Sibthorp and J. E. Smith. London: R. and A. Taylor.
- Spalik, K. and S. R. Downie. 2007. Intercontinental disjunctions in *Cryptotaenia* (Apiaceae, Oenantheae): An appraisal using molecular data. *Journal of Biogeography* 34: 2039–2057.
- Spalik, K., M. Piwczynski, C. A. Danderson, R. Kurzydina-Młynik, T. S. Bone, and S. R. Downie. 2010. Amphitropic amphiantarctic disjunctions in Apiaceae subfamily Apioideae. *Journal of Biogeography* 37: 1977–1994.

- Spalik, K., A. Wojewódzka, and S. R. Downie. 2001. The evolution of fruit in Scandiceae subtribe Scandicinae (Apiaceae). *Canadian Journal of Botany* 79: 1358–1374.
- Spooner, D. M., P. Rojas, M. Bonierbale, L. A. Mueller, M. Srivastav, D. Senalik, and P. W. Simon. 2013. Molecular phylogeny of *Daucus* (Apiaceae). *Systematic Botany* 38: 850–857.
- Spooner, D. M., M. P. Widrlechner, K. R. Reitsma, D. E. Palmquist, S. Rouz, Z. Ghrabi-Gammar, M. Neffati, B. Bouzbida, H. Ouabbou, M. El Koudrin, and P. W. Simon. 2014. Reassessment of practical subspecies identifications of the USDA *Daucus carota* L. germplasm collection: Morphological data. *Crop Science* 54: 706–718.
- Staden, R. 1996. The Staden sequence analysis package. *Molecular Biotechnology* 5: 233–241.
- Swofford, D. 2002. PAUP*: Phylogenetic analysis using parsimony (*and other methods), version 4.0a145. Sunderland: Sinauer Associates.
- Tautz, D., P. Arctander, A. Minelli, R. H. Thomas, and A. P. Vogler. 2003. A plea for DNA taxonomy. *Trends in Ecology & Evolution* 18: 70–74.
- Tavares, A. C., J. Loureiro, S. Castro, A. P. Coutinho, J. Paiva, C. Cavaleiro, L. Salgueiro, and J. M. Canhoto. 2014. Assessment of *Daucus carota* L. (Apiaceae) subspecies by chemotaxonomic and DNA content analyses. *Biochemical Systematics and Ecology* 55: 222–230.
- Tenore, M. 1830. *Ad Florae Neapolitanae plantarum vascularium Syllogem, Appendix tertia: Emendationes, atque Additamenta novissima complectens*. Napoli: Francese.
- Thellung, A. 1926. *Daucus*—studien. *Repertorium Specierum Novarum Regni Vegetabilis* 23: 147–159.
- Vallejo-Roman, C. M., E. I. Terentieva, T. H. Samigullin, and M. G. Pimenov. 2002. Relationships among genera in Saniculoideae and selected Apioideae (Umbelliferae) inferred from nrITS sequences. *Taxon* 51: 91–101.
- Venables, W. N. and B. D. Ripley. 2002. *Modern applied statistics with S*. Ed. 4. New York: Springer.
- Vilela, N. J. 2004. Cenoura: Um alimento nobre na mesa popular. *Horticultura Brasileira* 22: 1–2.
- Vivek, B. S. and P. W. Simon. 1999. Phylogeny and relationships in *Daucus* based on restriction fragment length polymorphisms (RFLPs) of the chloroplast and mitochondrial genomes. *Euphytica* 105: 183–189.
- Westmoreland, D. and C. Muntan. 1996. The influence of dark central florets on insect attraction and fruit production in Queen Anne's Lace (*Daucus carota* L.). *American Midland Naturalist* 135: 122–129.
- Wickman, H. 2011. The split-apply-combine strategy for data analysis. *Journal of Statistical Software* 40: 1–29.
- Zimmer, E. A. and J. Wen. 2013. Using nuclear gene data for plant phylogenetics: Progress and prospects. *Molecular Phylogenetics and Evolution* 66: 539–550.
- Zwickl, D. J. 2006. *Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion*. Ph. D. thesis. Austin: University of Texas.