

Mitochondrial DNA Sequence Phylogeny of *Daucus*

David M. Spooner,^{1,3} Holly Ruess,² Philipp Simon,¹ and Douglas Senalik¹

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

²Bioinformatics Core Facility, University of Texas Southwestern Medical Center, Dallas, Texas 75390 USA

³Author for correspondence (david.spooner@ars.usda.gov)

Communicating Editor: Adriano Stinca

Abstract—We explored the phylogenetic utility of mitochondrial DNA sequences in *Daucus* and compared the results with prior phylogenetic results using the same 36 accessions of *Daucus* (and two additional outgroups) with plastid DNA sequences and with other nuclear results. As in the plastid study we used Illumina HiSeq sequencer to obtain resequencing data of the same accessions of *Daucus* and outgroups, and analyzed the data with maximum parsimony and maximum likelihood. We obtained data from 47 of 71 total mitochondrial genes but only 17 of these 47 genes recovered major clades that were common in prior plastid and nuclear studies. Our phylogenetic trees of the concatenated data set of 47 genes were moderately resolved, with 100% bootstrap support for most of the external and many of the internal clades, except for the clade of *D. carota* and its most closely related species *D. syrticus*. There are areas of hard incongruence with phylogenies using plastid and nuclear data. In agreement with other studies, we conclude that mitochondrial sequences are generally poor phylogenetic markers, at least at the genus level, despite their utility in some other studies.

Keywords—Apiaceae, carrot, next generation sequencing.

The genus *Daucus* is a member of the large and taxonomically complex family Apiaceae with a recent estimate of 466 genera and 3820 species (Plunkett et al. 2019). Identification of members of the Apiaceae to family is generally unambiguous, but generic boundaries are often vague, and recent molecular studies have driven many reassignments of species to other genera. This is most evident in the genus *Daucus* that has benefitted from molecular systematic studies using a variety of plastid and/or nuclear regions (e.g. Downie and Katz-Downie 1996; Downie et al. 1996; Spalik and Downie 2007; Banasiak et al. 2016). Using DNA sequences from nuclear ribosomal internal transcribed spacer regions (ITS) and three plastid markers, Banasiak et al. (2016) investigated many species of *Daucus* and related genera, and transferred 18 species in nine genera into *Daucus*: *Agrocharis* Hochst. (four species), *Melanoselinum* Hoffm. (one species), *Monizia* Lowe (one species), *Pachyctenium* Maire & Pamp. (one species), *Pseudorhiza* (Murb.) Murb. (two species), *Rouya* Coincy (one species), *Tornabenea* Parl. (six species), *Athamanta dellacellae* E.A. Durand & Barratte, and *Cryptotaenia elegans* Webb ex Bolle.

The most recent molecular studies in *Daucus* used next-generation sequencing (NGS) approaches to test these results and add resolution to the clades. Arbizu et al. (2014) investigated the phylogeny of *Daucus* with 94 nuclear orthologs, producing highly resolved phylogenetic trees, and identified ten ortholog markers that provided a phylogeny nearly identical to the entire dataset. Arbizu et al. (2016) demonstrated the utility of these ten nuclear orthologs in a more focused study of the species boundaries of the *D. guttatus* complex. Both of these studies resolved two main clades A and B, and within clade A, subclade A' composed entirely of species with $2n = 18$ chromosomes (*D. carota*, *D. syrticus*), in contrast to all other species with $2n = 16, 20, 22$, and one polyploid of $2n = 66$. In a search for increased taxonomic resolution with a different genome, Spooner et al. (2017) generated resequencing data of 36 accessions of *Daucus* (and two additional outgroups) representing all major clades of *Daucus* with entire plastid DNA sequences showing areas of congruence and incongruence with the nuclear data (Arbizu et al. 2014, 2016). Fortuitously, these same resequencing data

can distinguish not only plastid and single- to low-copy nuclear sequences, but also mitochondrial sequences (Spooner et al. 2017), leading to the present analysis of these same accessions with mitochondrial DNA.

The mitochondrion is one of three DNA-containing genomes in plants, in addition to plastids and nuclei. The mitochondrion encodes for some of the genes necessary to produce the proteins required in the oxidative phosphorylation reaction that produces ATP. Mitochondrial DNA, like plastid DNA, is predominantly maternally inherited, but with exceptions to paternal and bi-parental inheritance (Breton and Stewart 2015).

Mitochondrial DNA has been comparatively underutilized for phylogenetic analyses in plants relative to the nucleus and plastid because of a number of complicating factors. These include dramatic variation in the size of the genome, even among closely related sister taxa, variation in rates of synonymous substitutions, variation in rates of gene loss accompanied by functional transfer to the nucleus, and in rates of genome rearrangement (Cole et al. 2018). Earlier, Rubinoff and Holland (2005) claimed that mtDNA is a rich source of markers for the study of closely related taxa because of the very low rate of recombination, maternal inheritance, simple genetic structure, reduced effective population size, and relatively rapid rates of evolution. This contrasted with data from Wolfe et al. (1987) showing that mtDNA generally has low mutation rates. An alternative perspective was provided by Galtier et al. (2009) who designated mtDNA as “intrinsically the worst population genetic and phylogenetic molecular marker we can think of,” pointing out its very low incidence of recombination, lack of response to selection, and erratic evolutionary rates. With these caveats in mind, the purpose of the present study is to generate a mitochondrial DNA phylogeny of 36 accessions of *Daucus* (and two additional outgroups) and to compare the results with the same accessions using plastid data, and with prior nuclear phylogenies of Arbizu et al. (2014, 2016).

MATERIALS AND METHODS

DNA Extraction and Sequencing—The complete mitochondrial genome of *Daucus carota* NC_017855.1 can be accessed from NCBI. The remaining

accessions we sequenced are listed in Appendix 1 and are the same ones examined in Spooner et al. (2017).

Mapping Sequences and Assembly—Sequences were quality trimmed with Trimmomatic v. 0.22 (Bolger et al. 2014), using minimum quality 28, minimum length 50, and adapter trim. Cleaned read pairs were mapped to six *Daucus carota* subsp. *sativus* genomes: accession LNRQ01000010 (Iorizzo et al. 2016), NC_017855.1 (Iorizzo et al. 2012), and four unpublished, draft genomes (three male sterile: B493A, B2163A, B10138A; one male fertile: B2163B) using BWA-MEM v. 0.7.10 (Li and Durbin 2009) with parameters -a -M -t 8. The SAM file was then converted to a BAM file (view), sorted (sort), and indexed (index) using SAMtools v. 0.1.19 (Li et al. 2009). All reads that mapped to the mitochondrial genomes were extracted and paired using the standard protocol of Picard SamToFastq v. 1.119 (<https://broadinstitute.github.io/picard/>). Mitochondrial mapped, paired reads were assembled using abyss-pe (Simpson et al. 2009), with a k-mer value of 64. The resulting aligned PAUP file, Supplemental Tables 1, 2, and Supplemental Fig. 1 are available from the Dryad Digital Repository (Spooner et al. 2020). DNA sequences of the 47 completely sequenced genes are available in GenBank.

Gene Length and Coverage—The coverage of each assembled gene was determined by mapping the mapped, paired reads back to the *.8.fa using BWA-MEM, and then converted, sorted, indexed (see above), and duplicates marked (Picard MarkDuplicates, standard protocol). GATK (McKenna et al. 2010) and DepthOfCoverage (DePristo et al. 2011; Van der Auwera et al. 2013) determined the coverage for each contig. Gene locations were determined by comparing the list of gene sequences from NC_017855.1 to the *.8.fa with nucmer -maxmatch, followed by show-coords -rd (v. 3.1) (Delcher et al. 2002). The average coverage of each gene was calculated by averaging the coverage of each gene based on the coordinates generated (Supplemental Table 2, Spooner et al. 2020).

Alignment and Annotation of Genes—Mitochondrial gene sequences were aligned using MUSCLE v. 3.8.31b (Edgar 2004). The aligned sequences were manually corrected to minimize gaps using Mesquite v. 3.03 (Maddison and Maddison 2015). Genes with introns were annotated from the NC_017855.1 reference, and added to the final NEXUS file as a CHARSET.

Phylogenetic Analyses—We rooted our trees on *Oenanthe virgata*, based on Downie et al. (2000). We first performed maximum parsimony (MP) analyses of the entire data set (all taxa and all characters). All MP analyses were conducted in PAUP* v. 4.0a131 (Swofford 2002). Question marks and blank spaces were treated as missing data and gaps, respectively. 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-addition sequence replicates and one tree held for each replicate. Branch swapping used tree-bisection reconnection (TBR) retaining all most parsimonious trees. 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 1000 replicates with simple addition sequence, setting MAXTREES to 1000. Maximum likelihood (ML) phylogenetic analysis was also obtained for the entire data set with the program RAXML v. 8.0.0 (Stamatakis 2014), using GTR + G model and estimating individual alpha-shape parameters, GTR rates, and empirical base frequencies for each individual gene. Using the same program, 1000 nonparametric bootstrap inferences were obtained. Both analyses were conducted via the CIPRES (Miller et al. 2010) portal at the San Diego Supercomputer Center (<http://www.phylo.org>).

RESULTS

Completeness of Sequencing—Forty-seven of the 71 genes we analyzed recovered complete sequences for all but a few of the 36 accessions of *Daucus* (and two additional outgroups) without any early stop codons or no stop codons or pseudogenes; we did not analyze the remaining 24 genes that had these problems (Table 1).

Phylogenetic Analyses—Individual bootstrap consensus trees of the 47 genes are presented in Supplemental Fig. 1 (Spooner et al. 2020). These trees vary greatly in resolution (as assessed by polytomies vs. topological structure and high bootstrap support values) from showing no to very little resolution (e.g. *cox3*, *nad2*, *nad3*, *rps13*), to greatly increased resolution (e.g. *cox2*, *nad4*, *nad7*, *rrn18*, *rrn26*) recovering at

TABLE 1. Topology of the 47 individual gene trees recovering the three main clades and outgroup of *Daucus*.

47 completely sequenced gene trees	<i>Caucalis</i> sister	A'	A	B
<i>atp4</i>				x
<i>atp9</i>				x
<i>ccmB</i>				
<i>ccmC</i>				
<i>ccmFc</i>	x			x
<i>ccmFn</i>	x			x
<i>cob</i>	x			x
<i>cox1</i>				
<i>cox2</i>	x			x
<i>cox3</i>				
<i>matR</i>	x			
<i>mttB</i>	x			
<i>nad1</i>				
<i>nad2</i>	x			
<i>nad3</i>				
<i>nad4</i>	x			x
<i>nad4L</i>				x
<i>nad5</i>	x			
<i>nad6</i>				
<i>nad7</i>	x		x	x
<i>nad9</i>	x			
<i>orf28</i>				
<i>orf31</i>	x		x	x
<i>orf34</i>				
<i>orf39</i>				
<i>orf40</i>				x
<i>orf41</i>				
<i>orf42</i>				
<i>orf46</i>				x
<i>orf47</i>				
<i>orf48</i>				
<i>orf51</i>				x
<i>orf56</i>				
<i>orf57</i>				
<i>orf58</i>	x			
<i>orf59</i>	x			
<i>orf60</i>	x			
<i>rpl5</i>				x
<i>rpl10</i>				
<i>rps1</i>				
<i>rps3</i>				
<i>rps4</i>				
<i>rps7</i>				x
<i>rps12</i>				
<i>rps13</i>				
<i>rrn18</i>	x		x	x
<i>rrn26</i>	x		x	x

least clade B and the outgroup (*Oenanthe virgata*) relative to earlier studies (Appendix 1; Table 1). Maximum parsimony analysis of all 47 genes were performed using the concatenated dataset (Fig. 1A) from 55,730 characters, with 708 variable parsimony uninformative characters, 957 parsimony informative characters, producing 24 equally parsimonious trees of 2121 length, consistency index 0.83, consistency index excluding uninformative characters 0.74, retention index 0.92, and rescaled consistency index 0.76. This concatenated 47 gene mtDNA tree, like the nuclear ortholog data of Arbizu et al. (2014), and the plastid data of Spooner et al. (2017) recovered the same main clades A, B, and outgroup, but with many topological differences within clades A and B. For example, the mtDNA tree failed to separate the 18-chromosome species of clade A' (*D. carota* all subspecies, *D. syrticus*) from other species in clade A: *D. pumilus* and *D. rouyi*. In addition, there were sister group relationship differences in mitochondrial clade A to the plastid (Fig. 1B, C) (Spooner et al. 2017) and nuclear data

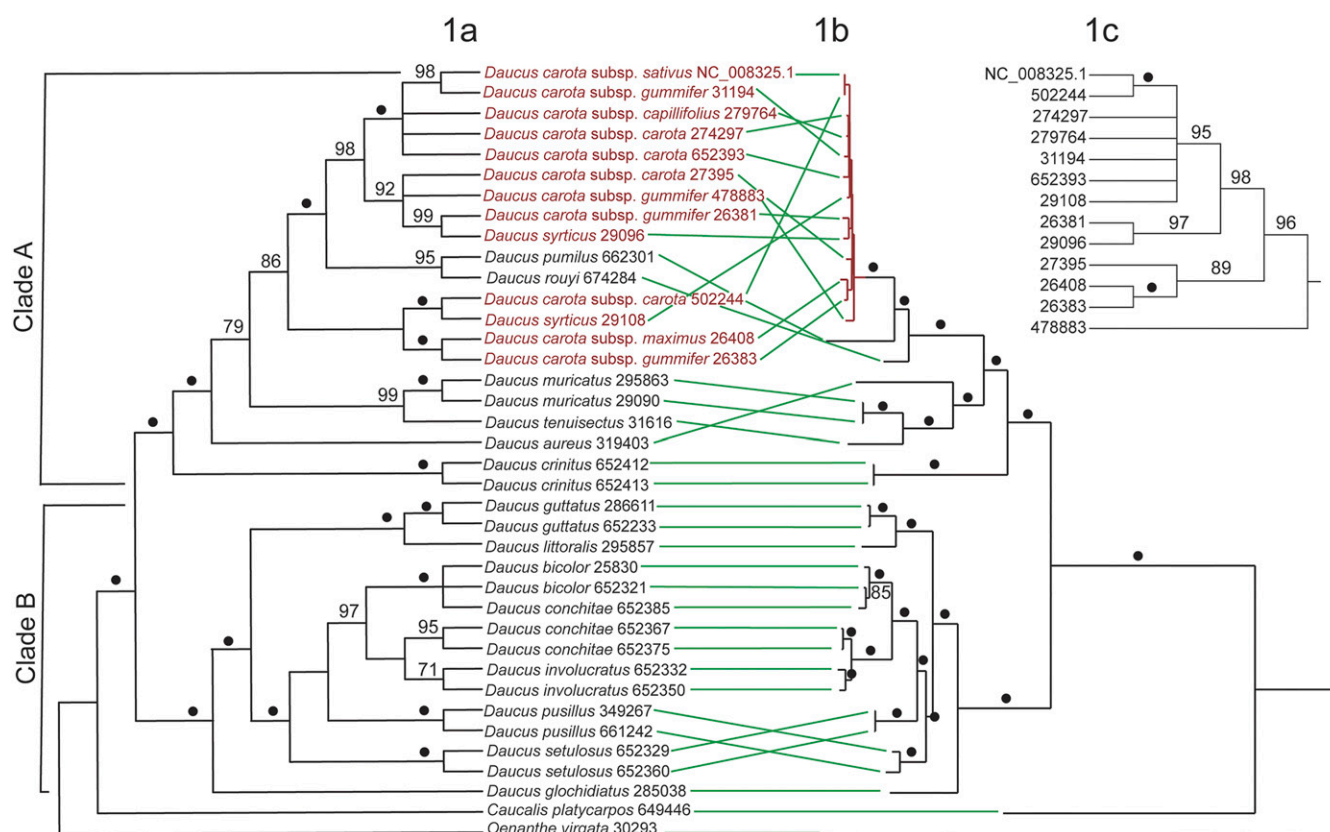


FIG. 1. 1a: Strict consensus of 24 maximum parsimony trees from the 47 completely sequenced mitochondrial genes; clades A and B are outlined at the left; bootstrap values above 70% shown numerically, and dots are 100%. 1b: The entire maximum parsimony plastid tree of Spooner et al. (2017). Green lines show concordance and discordance of the mitochondrial tree to the plastid tree; the red highlighted species and the red highlighted upper portion of the tree in Fig. 1b represent the species and topology of clade A' in the plastid tree. 1c: The expanded upper part of Fig. 1b, highlighting topological structure and bootstrap support of clade A' that is not present in the mitochondrial tree (1a).

(Arbizu et al. 2014, 2016) results regarding *D. aureus*, *D. crinitus*, *D. muricatus*, and *D. tenuisectus* and *D. carota*. Within clade B, both the plastid and mtDNA trees place *D. glochidiatus* sister to all other species of clade B, and *D. conchitae* PI 652385 is in a clade with both accessions of *D. bicolor*, different from the nuclear data (Arbizu et al. 2014). However, there are topological differences in these two datasets to the mitochondrial data regarding *D. pusillus* and *D. setulosus* (Fig. 1B, C).

In an attempt to overcome these differences we examined only the 17 genes as a concatenated dataset that recovered at least clade B and the outgroup, and we then ran a maximum parsimony analysis of these. Maximum parsimony analysis of these 17 genes (Fig. 2) was constructed from 32,462 characters, with 409 parsimony uninformative characters, 638 parsimony informative characters, producing 6 equally parsimonious trees of 1345 length, consistency index 0.82, consistency index excluding uninformative characters 0.74, retention index 0.93, and rescaled consistency index 0.77. This tree had lower bootstrap support values than the 47 gene tree and it retained topological differences from the 47 gene tree making it no better in recovering "expected" topological structure relative to the prior plastid and nuclear results.

Maximum likelihood analysis of 47 genes as a concatenated dataset (Fig. 3) produced a topology essentially equal to the MP analysis of the 47 genes except for minor differences in the apex of the tree containing five accessions of *D. carota*

with both trees containing in this area polytomies or low bootstrap support.

DISCUSSION

In summary, our *Daucus* results support the conclusion of Galtier et al. (2009) that mtDNA is a poor phylogenetic marker, at least at the genus level. We found only 17 of the 47 completely sequenced genes (of 71 total) that recovered clades B and the outgroup, but only four of these (*nad7*, *orf31*, *rrn18*, *rrn26*) recovered clades A, B, and the outgroup, and none of these four provided well-resolved topological or bootstrap support (Table 1) approaching prior results using the plastid or nuclear genes. The concatenated dataset of 47 genes provided better topological structure but only with considerable input of resources and time in bioinformatic analyses, and with less support and incongruence to nuclear and plastid results. These discordant results are part of a pattern showing up in many phylogenetic studies comparing different genomes (plastid, mitochondrial, nuclear) as well as comparing different regions of the same nuclear genome (e.g. Rokas et al. 2003; Baum 2007). There are a variety of possible mechanisms producing the discordance we highlight here, including hybridization, introgression, or other causes (Wendel and Doyle 1998). Hybridization and introgression are perhaps the most frequently cited causes of this discordance, but they are hard to distinguish from each other. Edelman et al. (2019) developed new techniques to analyze assembled genomes to show that

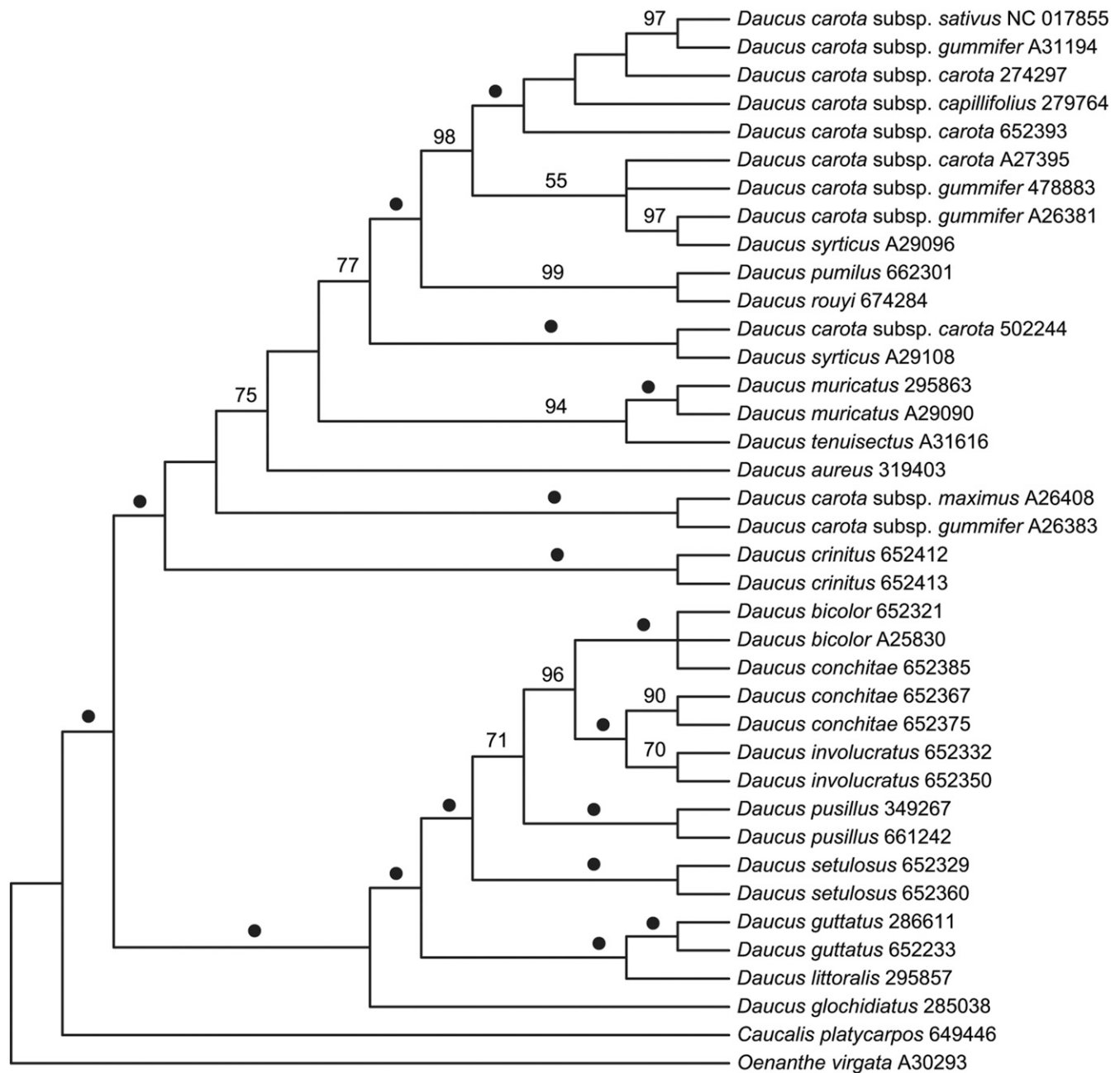


FIG. 2. Strict consensus of the 16 maximum parsimony mitochondrial trees from the 17 genes with similar topologies. Bootstrap values above 70% shown; dots are 100%.

hybridization was operative in discordant phylogenies in *Heliconius* butterflies, and such techniques surely will be used in other groups.

Despite these problems in *Daucus*, mtDNA has been shown to be useful in addressing important evolutionary and phylogenetic questions at higher taxonomic levels. For example, Beckert et al. (2001) found the *nad2* and *nad5* genes to be useful for phylogenetic analysis at the ordinal level (but not below the family level) in mosses. Adams et al. (2002) surveyed mitochondrial gene losses in 280 genera of flowering plants and found that the oldest groups of angiosperms contain nearly the same set of genes as their algal ancestors relative to more advanced families that experienced more losses. Mitochondrial DNA has been used in systematic studies at the ordinal

level in the Asteraceae (e.g. Wang et al. 2018). Park et al. (2015) constructed a phylogeny of 17 species of *Geranium*, along with representatives of parasitic plants in other plant families, to infer sources of intracellular gene transfer and horizontal gene transfer in *Geranium*. In summary, mtDNA appears to provide well-resolved phylogenetic results in some cases, but generally only at higher taxonomic levels, and not well resolved in *Daucus*.

AUTHOR CONTRIBUTIONS

DSp conceived of the study and conducted the phylogenetic analyses; HR and DSe performed bioinformatic analyses; all authors wrote and approved of the manuscript.

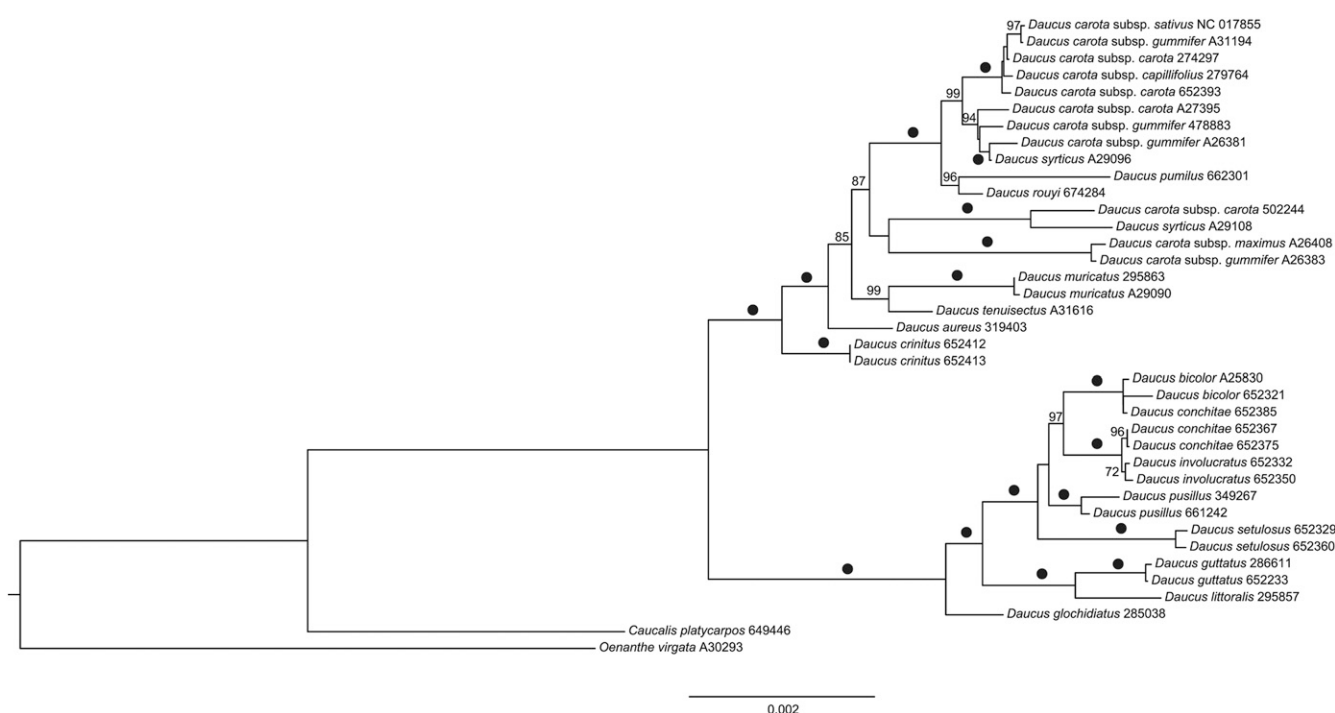


FIG. 3. Maximum likelihood tree of the 47 completely sequenced genes. Bootstrap values above 70% shown; dots are 100%.

LITERATURE CITED

- Adams, K. L., Y.-L. Qiu, M. Stouthamer, and J. D. Palmer. 2002. Punctuated evolution of mitochondrial gene content: High and variable rates of mitochondrial gene loss and transfer to the nucleus during angiosperm evolution. *Proceedings of the National Academy of Sciences USA* 99: 9905–9912.
- Arbizu, C., H. Ruess, D. Senalik, P. W. Simon, and D. M. Spooner. 2014. Phylogenomics of the carrot genus (*Daucus*, Apiaceae). *American Journal of Botany* 101: 1666–1685.
- Arbizu, C., P. W. Simon, F. Martínez-Flores, H. Ruess, M. B. Crespo, and D. M. Spooner. 2016. Integrated molecular and morphological studies of the *Daucus guttatus* complex (Apiaceae). *Systematic Botany* 41: 479–492.
- Banasiak, L., A. Wojewódzka, J. Baczyński, J.-P. Reduron, M. Piwczyński, R. Kurzyńska-Młynik, R. Gutaker, A. Czarnocka-Cieciura, S. Kosmala-Grzechnik, and K. Spalik. 2016. Phylogeny of Apiaceae subtribe Daucinae and the taxonomic delineation of its genera. *Taxon* 65: 563–585.
- Baum, D. A. 2007. Concordance trees, concordance factors, and the exploration of reticulate genealogy. *Taxon* 56: 417–426.
- Beckert, S., H. Muhle, D. Pruchner, and V. Knoop. 2001. The mitochondrial *nad2* gene as a novel marker locus for phylogenetic analysis of early land plants: A comparative analysis in mosses. *Molecular Phylogenetics and Evolution* 18: 117–126.
- Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120.
- Breton, S. and D. T. Stewart. 2015. Atypical mitochondrial inheritance patterns in eukaryotes. *Genome* 58: 423–431.
- Cole, L. W., W. Guo, J. D. Mower, and J. D. Palmer. 2018. High and variable rates of repeat-mediated mitochondrial genome rearrangement in a genus of plants. *Molecular Biology and Evolution* 35: 2773–2785.
- Delcher, A. L., A. Phillippy, J. Carlton, and S. L. Salzberg. 2002. Fast algorithms for large-scale genome alignment and comparison. *Nucleic Acids Research* 30: 2478–2483.
- DePristo, M., E. Banks, R. Poplin, K. Garimella, J. Maguire, C. Hartl, A. Philippakis, G. del Angel, M. A. Rivas, M. Hanna, A. McKenna, T. J. Fennell, A. M. Kernysky, A. Y. Sivachenko, K. Cibulskis, S. B. Gabriel, D. Altshuler, and M. J. Daly. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics* 43: 491–498.
- Downie, S. R. and D. S. Katz-Downie. 1996. A molecular phylogeny of Apiaceae subfamily Apioideae: Evidence from nuclear ribosomal DNA internal transcribed spacer sequences. *American Journal of Botany* 83: 234–251.
- Downie, S. R., D. S. Katz-Downie, and K. J. Cho. 1996. Phylogenetic analysis of Apiaceae subfamily Apioideae using nucleotide sequences from the chloroplast *rpoC1* intron. *Molecular Phylogenetics and Evolution* 6: 1–18.
- Downie, S. R., D. S. Katz-Downie, and M. F. Watson. 2000. A phylogeny of the flowering plant family Apiaceae based on chloroplast DNA *rpl16* and *rpoC1* intron sequences: Towards a suprageneric classification of the subfamily Apioideae. *American Journal of Botany* 87: 273–292.
- Edgar, R. C. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32: 1792–1797.
- Edelman, N. B., P. B. Frandsen, M. Miyagi, B. Clavijo, J. Davey, R. B. Dikow, G. Garcia-Accinelli, S. M. Van Belleghem, N. Patterson, D. E. Neafsey, R. Challis, S. Kumar, G. R. P. Moreira, C. Salazar, M. Chouteau, B. A. Counterman, R. Papa, M. Blaxter, R. D. Reed, K. K. Dasmahapatra, M. Kronforst, M. Joron, C. D. Jiggins, W. O. McMillan, F. Di Palma, A. J. Blumberg, J. Wakeley, D. Jaffe, and J. Mallet. 2019. Genomic architecture and introgression shape a butterfly radiation. *Science* 366: 594–599.
- 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.
- Galtier, N., B. Nabholz, S. Glémin, and G. D. D. Hurst. 2009. Mitochondrial DNA as a marker of molecular diversity: A reappraisal. *Molecular Ecology* 18: 4541–4550.
- Iorizzo, M., D. Senalik, M. Szklarczyk, D. Grzebelus, D. Spooner, and P. Simon. 2012. *De novo* assembly of the carrot mitochondrial genome using next generation sequencing of whole genomic DNA provides first evidence of DNA transfer into an angiosperm plastid genome. *BMC Plant Biology* 12: 61.
- Iorizzo, M., S. Ellison, D. Senalik, P. Zeng, P. Satapoomin, M. Bowman, M. Iovene, W. Sanseverino, P. Cavagnaro, M. Yildiz, A. Macko-Podgórn, E. Moranska, E. Grzebelus, D. Grzebelus, H. Ashrafi, Z. Zheng, S. Cheng, D. Spooner, A. Van Deynze, and P. Simon. 2016. A high-quality carrot genome assembly reveals new insights into carotenoid accumulation and Asterid genome evolution. *Nature Genetics* 48: 657–666.
- Li, H. and R. Durbin. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25: 1754–1760.

- Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, and R. Durbin. the 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25: 2078–2079.
- Maddison, W. P. and D. R. Maddison. 2015. Mesquite: A modular system for evolutionary analysis. Version 3.03. <http://mesquiteproject.org>.
- McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, and M. A. DePristo. 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* 20: 1297–1303.
- Miller, M. A., W. Pfeiffer, and T. Schwartz. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *Proceedings of the Gateway Computing Environments Workshop (GCE)*. New Orleans: Gateway Computing.
- Park, S., F. Grewe, A. Zhu, T. A. Ruhlman, J. Sabir, J. P. Mower, and R. K. Jansen. 2015. Dynamic evolution of *Geranium* mitochondrial genomes through multiple horizontal and intracellular gene transfers. *The New Phytologist* 208: 570–583.
- Plunkett, G. M., M. G. Pimenov, J.-P. Reduron, E. V. Kluykov, B.-E. van Wyk, T. A. Ostroumova, M. J. Henwood, P. M. Tilney, K. Spalik, M. F. Watson, B.-Y. Lee, F.-D. Pu, C. J. Webb, J. M. Hart, A. D. Mitchell, and B. Muckensturm. 2019. Apiaceae. Pp. 9–206 in *The Families and Genera of Vascular Plants*, vol. 15, eds. J. W. Kadereit and V. Bittrich. Berlin: Springer-Verlag.
- 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.
- Rubinoff, D. and B. S. Holland. 2005. Between two extremes: Mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. *Systematic Biology* 54: 952–961.
- Simpson, J. T., K. Wong, S. D. Jackman, J. E. Schein, S. J. Jones, and I. Birol. 2009. ABySS: A parallel assembler for short read sequence data. *Genome Research* 19: 1117–1123.
- Spalik, K. and S. R. Downie. 2007. Intercontinental disjunctions in *Cryptotaenia* (Apiaceae, Oenantheae): An appraisal using molecular data. *Journal of Biogeography* 34: 2039–2054.
- Spooner, D. M., H. Ruess, M. Iorizzo, D. Senalik, and P. Simon. 2017. Entire plastid phylogeny of the carrot genus (*Daucus*, Apiaceae): Concordance with nuclear data and mitochondrial and nuclear DNA insertions to the plastid. *American Journal of Botany* 104: 296–312.
- Spooner, D. M., H. Ruess, P. Simon, and D. Senalik. 2020. Data from: Mitochondrial DNA sequence phylogeny of *Daucus*. Dryad Digital Repository. <https://doi.org/10.5061/dryad.p80j893>.
- Stamatakis, A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30: 1312–1313.
- Swofford, D. 2002. PAUP*: Phylogenetic analysis using parsimony (*and other methods), v. 4.0a131. Sunderland: Sinauer Associates.
- Van der Auwera, G. A., M. Carneiro, C. Hartl, R. Poplin, G. del Angel, A. Levy-Moonshine, T. Jordan, K. Shakir, D. Roazen, J. Thibault, E. Banks, K. V. Garimella, D. Altshuler, S. Gabriel, and M. A. DePristo. 2013. From FastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. *Current Protocols in Bioinformatics* 43: 11.10.1–11.10.33.
- Wang, S., Q. Song, S. Li, Z. Hu, G. Dong, C. Song, H. Huang, and Y. Liu. 2018. Assembly of a complete mitogenome of *Chrysanthemum nankingense* using Oxford nanopore long reads and the diversity and evolution of Asteraceae mitogenomes. *Genes* 9: 547.
- Wendel, J. F. and J. J. Doyle. 1998. Phylogenetic incongruence: Window into genome history and molecular evolution, Pp. 265–296 in *Molecular Systematics of Plants II: DNA Sequencing*, eds. D. E. Soltis, P. S. Soltis, and J. J. Doyle. Boston: Springer.
- Wolfe, K. H., W. H. Li, and P. M. Sharp. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences USA* 84: 9054–9058.

APPENDIX 1. List of taxa. The 6-digit Plant Introduction (PI) numbers are permanent numbers assigned to germplasm accessions in the National Plant Germplasm System (NPGS). Germplasm centers in the NPGS assign temporary site-specific numbers to newly acquired germplasm (the 5-digit Ames numbers for carrots and other Apiaceae maintained at the North Central Regional Plant Introduction Station in Ames, Iowa, USA) until an accession's passport data and taxonomy are verified, it is determined not to be a duplicate accession, and it has been determined the accession can be successfully maintained. These accessions may or may not be assigned a PI number after the assessment period.

Ingroup: *D. aureus* Desf.; 319403; Israel. Mediterranean Region. *D. bicolor* Sm. in Sibth.; 25830; Turkey. Mugla: 10 km northeast. *D. bicolor*; 652321; Greece. Macedonia: 1 km from Sarakini, toward Edessa. *D. carota* subsp. *capillifolius* (Gilli) C. Arbizu; 279764; Libya. Near Jefren. *D. carota* L. subsp. *carota*; 27395; Uzbekistan. Northeast of Gazelkent on main road to Tashken. *D. carota* subsp. *carota*; 652393; Turkey. Konya: 10–15 km to Seydisehir, between Yarpuz and Konya. *D. carota* subsp. *carota*; 502244; Portugal. Coimbra: Lousa. *D. carota* subsp. *carota*; 274297; Pakistan. Northern areas. *D. carota* subsp. *gummifer*; 26381; Portugal. Faro: Near Portunao. *D. carota* subsp. *gummifer*; 26383; Portugal. Faro: Near Aljezur. *D. carota* subsp. *gummifer*; 31194; Source: France. Parc Zoologique et Botanique de la Ville de Mulhouse. *D. carota* subsp. *gummifer*; 478883; France. Finistere: maritime turf, Le Conquet. *D. carota* subsp. *maximus* (Desf.) Ball; 26408; Portugal. Beja. *D. carota* L. subsp. *sativus* (Hoffm.) Schübl. & G. Martens; —; —; Source: Sequence obtained from NCBI accession NC_008325.1 from plant purchased in grocery store in US, cultivar 'Half-long'. *D. conchitae* Greuter; 652367; Turkey. Mugla. *D. conchitae*; 652375; Turkey. Mugla: between Dalaman-Gocik and Fethiye. *D. conchitae*; 652385; Turkey. Antalya: Olympos. *D. crinitus* Desf.; 652412; Portugal. Braganca: near Zava. *D. crinitus*; 652413; Portugal. Guarda: near Barca de Alva. *D. glochidiatus* (Labill.) Fisch., C.A. Mey. & Avé-Lall.; 285038; Source: CSIRO, Australia. Capital Territory. *D. guttatus* Sibth. and Sm.; 286611; Source: Lebanon. Faculty of Agricultural Sciences. *D. guttatus*; 652233; Iran. Mazandaran: Dhalus Road, Dasht-e Nazir, Kandalus. *D. involucratus* Sm.; 652332; Greece. Peloponnese: village of Loutra Agias Elenis, 17 km south of Korinthos, Korinthia Prefecture. *D. involucratus*; 652350; Turkey. Izmir. *D. littoralis* Sibth. & Sm.; 295857; Israel. Beit Alpha. *D. muricatus* L.; 295863; Spain. Cordoba. From Villa del Rio (Cordoba). *D. muricatus*; 29090; Tunisia. South of Tunis along Hwy. 3 toward Zaghouan. *D. pumilus* (L.) Hoffmanns and Link; 662301; Tunisia. South of Medenine toward Tataouine, near Bir Lahmer. *D. pusillus* Michx.; 349267; Uruguay. Montevideo. Near La Colorado Beach. *D. pusillus*; 661242; United States. Oregon: east side of road, 1 mile north of Pistol River Road, about 1.5 miles north of Pistol River, near Hunters River Cove, Curry County. *D. rouyi* Spalik and Reduron; 674284; Tunisia. Jendouba: road to Tabarka, near Tabarka airport. *D. setulosus* Guss. ex DC.; 652329; Greece. Peloponnese: 4 km from Skoura, toward Leonidion, Laconia Prefecture. *D. setulosus*; 652360; Turkey. Mugla: between Soke and Milas. *D. syrticus* Murb.; 29096; Tunisia. Between Tataouine and Bir Lahmer. *D. syrticus*; 29108; Tunisia. Between Medenine and Matnatas. *D. tenuisectus* Coss. ex Batt.; 31616; Morocco. Al Haouz: 25.7 km north of center of Ijoukak, 29 km south of Asni, Nfiss River Valley, Imgdal Region.

Outgroup: *Caucalis platycarpus* L.; 649446; Germany. Saxony-Anhalt: Mannsdorf. *Oenanthe virgata* Poir.; 30293; Tunisia. Beja: Route 11, 41 km from Eudiana, 254 km from Beja.