DNA Barcoding of the Solanaceae Family in Puerto Rico Including Endangered and Endemic Species

Lumariz Hernandez Rosario, Juan O. Rodríguez Padilla, Desiree Ramos Martínez, Alejandra Morales Grajales, Joel A. Mercado Reyes, Gabriel J. Veintidíos Feliú, Benjamin Van Ee, and Dimuth Siritunga

Department of Biology, University of Puerto Rico Mayaguez Campus, Mayaguez, PR 00680

ABSTRACT. The Solanaceae family is one of the largest and well-distributed plant families in the world. It contains species of agricultural and economical importance, such as Solanum tuberosum, Solanum melongena, Solanum lycopersicum, Nicotiana tabacum, and Capsicum annuum. In Puerto Rico, there are approximately 46 species of Solanaceae of which six are endemic: Brunfelsia densifolia, Brunfelsia lactea, Brunfelsia portoricensis, Goetzea elegans, Solanum ensifolium, and Solanum woodburyi. Our objective was to use DNA barcoding to identify the Solanaceae species in Puerto Rico, including the endemics, and to assess the species relationships between them. To accomplish our objective, two chloroplast regions (psbA-trnH and matK) and a nuclear region [internal transcribed spacer (ITS)] were assessed. Pairwise distance and phylogenetic analysis demonstrate that DNA barcoding can be used to discriminate at the species level among these taxa in Puerto Rico. For all three markers, the genus that showed the highest pairwise distance between represented species was Solanum, whereas the genus that displayed the least was Capsicum. Phylogenetic trees of single and concatenated regions were generated from sequences obtained in this study and from data downloaded from the National Center for Biotechnology Information database. Our results show that this technique can be used to identify species with one, two, or three combinations of DNA barcode markers depending on the taxon. In addition, this is the first study to include the endemic species S. woodburyi in a molecular phylogenetic analysis, and it was found to have a close relationship with S. ensifolium, also endemic to Puerto Rico, and to Solanum bahamense from the Bahamas and Greater Antilles. Therefore, we suggest that S. woodburyi might be part of the Bahamense clade.

The Solanaceae family is one of the major groups of angiosperms with approximately 2500 species and 100 genera (Filipowicz and Renner, 2012). The family contains species with agricultural and economical importance worldwide such as Solanum tuberosum, Solanum melongena, Solanum lycopersicum, Nicotiana tabacum, and Capsicum annuum, as well as numerous toxic or poisonous species. Puerto Rico is considered a biodiversity hotspot due to its tropical location and the high concentration of endemic species that have been threatened by habitat loss (Helmer et al., 2002; Myers et al., 2000). According to Axelrod (2011), in Puerto Rico, there are 46 species of Solanaceae from 14 genera. Of these 46 species, Brunfelsia densifolia, Brunfelsia lactea, Brunfelsia portoricensis, Goetzea elegans, Solanum ensifolium, and Solanum woodburyi are endemic. Some of these species are endangered or rare, with only a few studies conducted on them.

Filipowicz and Renner (2012) included the endemic and nonendemic Brunfelsia species of Puerto Rico in their systematic study. They hypothesized that these species reached the island 3 to 5 million years ago, and their seeds were dispersed by birds from South America. In research by Muscarella et al. (2014), the Puerto Rican endemic Solanaceae species G. elegans and S. ensifolium were included as part of the trees of Puerto Rico that were evaluated. They studied 89% of Puerto Rico’s trees to analyze the community phylogeny based on regions with different climate areas.

In Puerto Rico, there are two endemic and endangered Solanaceae species: G. elegans and S. ensifolium. In 1985, G. elegans was listed by the U.S. Fish and Wildlife Service under the Endangered Species Act. At that time, there were fewer than 50 individuals in three known sites (U.S. Fish and Wildlife Service, 1985). More recently, this species has increased in number of individuals and has been reported to be present in 10 localities, but it still is considered an endangered species (Vargas, 2013). Comparatively, the species distribution and status of S. ensifolium is less known, as it has not been evaluated since 1992. In 1992, 150 individuals were found in the locality of Las Tetas de Cayey (Vargas, 2015). It was hypothesized that the species could be present in other localities, but that has not been confirmed. Unfortunately, during this study, S. ensifolium was not found in Las Tetas de Cayey or at other localities where it was thought to be, such as in Florida, Puerto Rico. Its presence at other localities is currently unknown. Due to Hurricane Maria in 2017, there is a high possibility that S. ensifolium might be even more critically endangered than in 1992, and in the worst-case scenario, it has gone extinct in the wild. More
studies and conservation efforts are critically needed to determine the current status of *S. ensifolium*.

In addition, studies of rare species such as the endemic *S. woodburyi* are also needed, especially using molecular tools. This species has not been classified as endangered, but only a few individuals are known to be present in specific localities, such as the Sierra de Luquillo (Axelrod, 2011). Our study is the first to include *S. woodburyi* in a molecular study. Previous research on this species has focused on the description of the species in a synopsis of the endemic species of the West Indies (Howard, 1966; Knapp, 2009). Of the 46 Solanaceae species found in Puerto Rico many, such as *Jaltomata antillana* and *Solanum polygamum*, are not well studied with no molecular data available in public databases such as GeneBank. Therefore, we apply molecular techniques, specifically a DNA barcoding approach, to better understand the species relationships and delimitations for the taxa of Solanaceae in Puerto Rico.

DNA barcoding was proposed as an easy and cost-effective technique for the identification of species. It consists of the amplification of a universal, robust, and standard region. The first region proposed for this analysis was the mitochondrial gene cytochrome c oxidase I (COI) for animals (Hebert et al., 2003). It has been successfully used in studies of taxonomy, population genetics, forensics in wildlife crimes, and conservation and in species identification in cases of fish mislabeling, among others (Cywinska et al., 2006; Di Pinto et al., 2015; Hebert et al., 2004; Rolo et al., 2013). In plants, the use of the COI region is not recommended for DNA barcoding studies due to its low evolutionary rate (Kress et al., 2005). DNA barcoding for plants has been more challenging than for animals, and more than one barcode region is needed for species identification. Currently, the most accepted and widely used regions in DNA barcoding studies in plants are *rbcL*, *matK*, *psbA-trnH*, and *ITS* (Bolson et al., 2015; CBOL Plant Working Group, 2009; Kress, 2017; Li et al., 2014). Different studies have shown successful identification of species using a combination of these barcode markers, but some studies have also found limited ability to discriminate species within groups consisting of recently diverged species (Collins and Cruickshank, 2013; Kress, 2017; Kress et al., 2009; Muscarella et al., 2014; Spooner, 2009).

In this study, the barcode regions used were two chloroplast regions and a nuclear region, *matK*, *psbA-trnH*, and *ITS*, respectively. The *matK* region is 1500 bp long and encodes for maturase-like polypeptide, which is believed to be involved in Group II intron splicing (Selvaraj et al., 2008). It has an evolutionary rate suitable for distinguishing taxa at high taxonomic levels such as order, family, and, in some cases, at the genus levels (Yu et al., 2011). For a coding region, *matK* has higher variability compared with *rbcL*, but it has not been successfully amplified in all plants (such as gymnosperms). However, it has worked well for angiosperms (CBOL Plant Working Group, 2009; Dunning and Savolainen, 2010; González et al., 2009; Hollingsworth et al., 2011; Kress and Erickson, 2007). Because *matK* cannot distinguish at lower taxonomic levels, it is usually combined with other regions such as *psbA-trnH*, which has a higher substitution rate (Hollingsworth et al., 2011). The *psbA-trnH* region is an intergenic spacer of ~400 to 500 bp in length. Because of its variability, it has been demonstrated to discriminate successfully at the species level when combined with other barcodes (Bolson et al., 2015; Kress and Erickson, 2007; Kress et al., 2009; Storchová and Olson, 2007). For discrimination at the species level using DNA barcoding, the nuclear internal transcribed spacer (*ITS*) region been successfully used to complement chloroplast barcodes (Kress et al., 2005; Li et al., 2011). The internal transcribed spacers (*ITS1* and *ITS2*) are between the nuclear ribosomal 18S-5.8S-26S genes, and altogether the marker consists of less than 700 bp (Feliner and Rosselló, 2007). Although *ITS* has helped discriminate at the species level, there are two main concerns when using it in DNA barcoding studies: the possibility of amplifying the *ITS* region from fungal contaminants and the amplification of divergent paralogous copies (Hollingsworth, 2011).

Our overall objectives in this study were 1) to understand the feasibility of using DNA barcoding for identification of Solanaceae species found in Puerto Rico and 2) to study the phylogenetic relationships among the endemic and non endemic Solanaceae species of Puerto Rico.

**Material and Methods**

**Sample collection.** Permits from the Puerto Rico Department of Natural Resources (2016-EPE-032/O-VS-PVS15-SJ-00854-11072016) and the U.S. Fish and Wildlife Service (41521-2017-08) facilitated the collection of samples from protected areas. Using the location data of previous Solanaceae collections from the University of Puerto Rico Mayaguez’s Department of Biology herbarium (MAPR), 53 samples were collected from different field localities in Puerto Rico. Furthermore, 13 samples were collected from the Jardin Botánico of Puerto Rico, Fundación Luis Muñoz Rivera, Guajataca Forest, Rio Abajo Forest, and Maricao Forest. In addition, six samples from the MAPR herbarium were used in this study.

**DNA extraction.** The DNA was extracted from fresh samples using the Doyle and Doyle (1990) methodology with a few modifications. Approximately 100 mg of leaf tissue was measured from each sample and pulverized in a mortar chilled with liquid nitrogen. Then, 1.2 mL of 3% CTAB buffer was added and the sample was further crushed. After transferring this solution to a 2 mL microcentrifuge tube, the protocol was followed without further modifications (Doyle and Doyle, 1990). For the DNA extraction from herbarium samples, we used a DNeasy Plant Mini Kit (Qiagen, Germantown, MD) following the manufacturer’s recommended procedures.

**DNA amplification.** The polymerase chain reaction (PCR) amplifications of *matK* and *psbA-trnH* regions were conducted using primers suggested by the Barcode of Life Data System (BOLD). For *matK*: MatK-1RKM-f and MatK-3FKIM-r (K.J. Kim, unpublished), and for *psbA-trnH*: trnHf_05 (Tate and Simpson, 2003) and psbA3_f (Sang et al., 1997) were used (Table 1). For the amplification of the *ITS* region, the primers used were: ITS-u1 and ITS-u2 (Cheng et al., 2016), as well as ITS-Leu1 and ITS4 (Baum et al., 1998; White et al., 1990) (Table 1). The reaction for each sample included 5 μL of Promega buffer 1X (Promega, Madison, WI), 5 μL of 5 mM MgCl₂, 1 μL of 0.2 mM dNTPs, 1 μL of 0.2 μm forward and reverse primers, 0.2 μL of Promega GoTaq polymerase (1 unit), and 0.75 μL of dimethyl sulfoxide (3%) in a final volume of 25 μL. The thermo-cycling regimens used in the amplification of each region were: 94 °C for 1 min followed by 45 cycles of 94 °C for 30 s, 52 °C for 1 min, 72 °C for 1 min with a final single cycle of 72 °C for 10 min for *matK*; 94 °C for 45 s, followed by 45 cycles of 94 °C for 30 s, 48 °C for 1 min, 72 °C for 40 s with a final
single cycle of 72 °C for 10 min for psbA-trnH; and 94 °C for 45 s, followed by 45 cycles of 94 °C for 45 s, 48 °C for 40 s, 72 °C for 1 min with a final single cycle of 72 °C for 10 min for ITS. Five microliters of each PCR product was visualized in a 1% agarose gel and the rest of the PCR product was then purified using the QIAquick PCR Purification kit (Qiagen) following manufacturer’s protocol. The purified samples were then sent to the Genomic Science Laboratory at North Carolina State University (Raleigh) for Sanger sequencing following their protocols.

**Data analysis.** Sequences were aligned and manually edited using BioEdit version 7.1.9 (Hall, 1999). The alignment of each region was performed by ClustalX in BioEdit. A pairwise distance and a neighbor joining (NJ) analysis were performed for each marker using MEGA version 7.0.21 (Kumar et al., 2016). The pairwise distance data were used to compare relationships between genera and species. Sequence identifier version 1.8 and SequenceMatrix version 1.8 (Vaidya et al., 2011) were used to concatenate the barcode regions, and covert them to Nexus format. Afterward, the file with the concatenated sequences had a length range between 1984 to 2109 bp. Sequences were aligned and manually edited using BioEdit version 7.1.9 (Hall, 1999). The alignment of each region was performed by ClustalX in BioEdit. A pairwise distance and a neighbor joining (NJ) analysis were performed for each marker using MEGA version 7.0.21 (Kumar et al., 2016). The pairwise distance data were used to compare relationships between genera and species. Sequence identifier version 1.8 and SequenceMatrix version 1.8 (Vaidya et al., 2011) were used to concatenate the barcode regions, and covert them to Nexus format. Afterward, the file with the concatenated sequences was exported to PAUP version 4.0 (Swofford, 2002) in which a NJ analysis with 100 bootstraps was performed.

**Results**

A total of 72 fresh or dried samples were collected (Table 2), of which 66 samples were successfully amplified. PCR amplification success was higher for matK, followed by psbA-trnH and ITS with 91%, 88%, and 87%, respectively. In contrast, sequencing success was higher for psbA-trnH followed by matK and ITS with 89%, 88%, and 85%, respectively. The length of the alignments was 742 bp, 679 bp, and 688 bp for matK, psbA-trnH, and ITS, respectively. Of these regions, ITS provided the highest variability, and matK was the most conserved region. All sequences from this study have been deposited in GenBank (Table 2).

**Pairwise distances**

The pairwise distance range was 0.000 to 0.017, 0.000 to 0.054, and 0.000 to 0.054, and 0.007 to 0.088 for matK, psbA-trnH, and ITS, respectively. The outgroup pairwise distance range was 0.166 to 0.296, 0.350 to 0.670, and 0.172–0.467 for matK, psbA-trnH, and ITS, respectively. In all three markers, the genus that showed the highest pairwise distance between its species was Solanum, with 0.017, 0.054, and 0.088 for matK, trnH-psbA, and ITS, respectively. The genus that displayed the least pairwise distance between its species was Capsicum with 0.000 for the chloroplast barcodes while for ITS the genera that exhibited the least pairwise distance were Datura and Cestrum, both with 0.007 (Supplemental Tables 1–3).

**Neighbor joining analysis**

**MatK.** The matK barcode region was able to distinguish 15 specimens at the species level, and in other cases, it was only able to discriminate at genus level (Fig. 1). In the NJ analysis, matK was able to distinguish individuals at the species level of *G. elegans* and most *Solanum* species. However, species from *Physalis*, *Capsicum*, *Brugmansia*, *Datura*, *Brunfelsia*, and *Cestrum* could not be differentiated.

**psbA-trnH.** The NJ analysis of psbA-trnH showed better discrimination at the species level compared with matK (Fig. 2). Most species were separated into their respective clades except for those of the genera *Brugmansia*, *Capsicum*, *Cestrum*, and *Physalis*. For *Physalis* species, this barcode region showed better resolution than matK but was still not able to discriminate well at the species level. Other samples that could not be grouped into separate clades were those of the species *S. erianthum*, *S. rugosum*, and *S. polygamum*, which were grouped into a single clade. In addition, an unknown *Physalis* species [Table 2 (no. 27)] was recovered out of the *P. angulata* and *P. pubescens* clades.

**ITS.** The ITS barcode region showed higher variability among individuals in comparison with matK and psbA-trnH (Fig. 3). It was able to identify, used alone, even more species than psbA-trnH. Although more resolution is needed to have a more reliable discrimination between *Cestrum*, *Capsicum*, *Datura*, and *Physalis* species. In contrast with the other barcode regions, the specimens of *S. erianthum*, *S. polygamum*, and *S. rugosum* [Table 2 (no. 39, 46, and 47, respectively)] were able to be separated into clades by species. In contrast, *S. polygamum* and *S. erianthum* grouped into the same clade. Meanwhile, *S. erianthum* grouped in the *Solanum rugosum* clade.

**Concatenated analysis.** All three barcode regions were concatenated to increase discrimination (Fig. 4). The concatenated sequences had a length range between 1984 to 2109 bp. A total of 43 Solanaceae were included in the concatenated analysis. The genera *Capsicum*, *Solanum*, *Brugmansia*, *Datura*, *Cestrum*, *Brunfelsia*, and *Physalis* were grouped by clades as shown in Fig. 4. *Solanum* was recovered close to *Cestrum*, and these genera were closely related to *Acistus arborescens* in comparison with the other genera. Similarly, *Brugmansia*, *Datura*, *Solanthera*, and *Lycium* were recovered closely related to each other. A notable relationship was *Brownallia americana* with *Cestrum*, whereas *Brunfelsia* was more closely related to these two genera in comparison with the rest of the species. Moreover, *Physalis* grouped into a single clade separate from the other clades.

---

**Table 1. Primers used for this study.** Primers where selected from the Barcode of Life Data System (BOLD) primer database (Ratnasingham and Hebert, 2007).

| Primer   | Sequence 5’–3’ | Primer length (bp) | Reference                |
|----------|----------------|--------------------|--------------------------|
| ITS-u1(f) | GGAAGKARAAGTGCCTAACAAGG | 22 | Cheng et al., 2016 |
| ITS-u4(r) | GTTGGTTTTACCTCCGGCTTA | 20 | Cheng et al., 2016 |
| ITS   | GTCCTCAGAATCTATTTAG | 23 | Urbatsch et al., 2000 |
| ITS 4 | CCTTCCGGTATTGATAGC | 20 | White et al., 1990 |
| PsbA(f) | GGTATTGCAATGCAATGCTGTC | 22 | Sang et al., 1997 |
| trnH(r) | CGCCGATGGGATTTGACATCC | 23 | Tate and Simpson, 2003 |
| MatK-1RKIM(f) | ACCCAGTCCATCTGGAAACATTCTGGTTC | 27 | K.J. Kim, unpublished data |
| MatK-3FKIM(r) | CGTACAGTACTTTTGTGTACGAG | 25 | K.J. Kim, unpublished data |
Table 2. GenBank accession numbers of samples used for individual region analysis. Sequences obtained in this study are in bold.

| Scientific name | Nonendemic (N) or endemic (E) | **matK** | **psbA-trnH** | **ITS** |
|-----------------|-------------------------------|----------|---------------|---------|
| 1 *Acnistus arborescens* | N | KU568472.1, MH717175 | KJ426590.1, MK381182 | DQ314183.1 (1), DQ314181.1 (2), DQ314173.1 (3), MK412103 |
| 2 *Browallia americana* | N | EF430905.1, MH717176 | MF663708.1, MK381183 | MK412104 |
| 3 *Brugmansia × candida* | N | KP756823.1 (1), KC146591.1 (2), KC146590.1 (3), MH717177 | JX467616.1 (1), JX467615.1 (2), JX467614.1 (3), KC146626.1 (4), KC146625.1 (5) | JX467597.1 (1), MG693010.1 (2), LC076493.1 (3), HG738853.1 (4), MK412106, MK412107, MK412108 |
| 4 *Brugmansia suaveolens* | N | MF348991.1 (1), KP756824.1 (2), HM851090.1 (3), KC146589.1 (4), MH717178, MH717179, MH717180 | KC146624.1, MK381184, MK381185, MK381186 | JQ081178.1 (1), JQ081177.1 (2), MK412110, MK412111, MK412112, MK412113 |
| 5 *Brunfelsia sp.* | N | KP756837.1, MH717185 | KJ426625.1, MK381187, MK381188 | JQ081203.1 (1), JQ081202.1 (2), MK412114 |
| 6 *Brunfelsia Americana* | N | KP756837.1, MH717185 | KJ426625.1, MK381187, MK381188 | JQ081203.1 (1), JQ081202.1 (2), MK412114 |
| 7 *Brunfelsia densifolia* | E | MH717181, MH717182 | — | MK412105 |
| 8 *Brunfelsia lactea* | E | KJ012480.1, MH717183 | KJ426626.1, MK381192 | MK412116 |
| 9 *Brunfelsia nitida* | N | KP756841.1 | MF663707.1 (1), KJ426627.1 (2), HM446890.1 (3), MK381193 | JQ081211.1 (1), JQ081210.1 (2), MK412115 |
| 10 *Capsicum chinense* | E | HM446595.2 (1), KJ012481.1 (2)*, MH717186 | MF663707.1 (1), KJ426627.1 (2), HM446890.1 (3), MK381193 | MH717188 |
| 11 *Capsicum annuum* | N | EF537310.1 (1), EF537309.1 (2), EF537308.1 (3) | EF537250.1 (1), EF537249.1 (2), EF537248.1 (3) | JQ885438.1 (1), JQ885437.1 (2), JQ885436.1 (3), JQ885435.1 (4) |
| 12 *Capsicum frutescens* | N | AB721831.1 (1), AB721832.1 (2), AB721832.1 (3), HQ705990.1 (4), MH717188 | AB721831.1 (1), AB721832.1 (2), AB721832.1 (3), HQ705990.1 (4), JQ087871.1 (5), MK381195 | JQ081217.1, MK412115 |
| 13 *Cestrum sp.* | N | KP756837.1, MH717185 | KJ426649.1 (1), KJ426649.1 (2), MK381196 | MK412118, MK412119, MK412120 |
| 14 *Cestrum citrifolium* | N | JU012506.1 (1), JU012505.1 (2), MH717189, MH717190 | JU012506.1 (1), JU012505.1 (2), MH717189, MH717190 | MK381197, MK381200 |
| 15 *Cestrum diurnum* | N | GH54071.1 (1), KH65185.1 (2), MH717191, MH717192 | GH54071.1 (1), KH65185.1 (2), MH717191, MH717192 | MH717193 |
| 16 *Cestrum nocturnum* | N | MH446657.2, MH717193 | GH54071.1, MH381201 | MH381203 |
| 17 *Datura inoxia* | N | MF350104.1 (1), KJ738546.1 (2), JU52181.1 (3), MH717194 | MF348491.1 (1), HG63496.1 (2), JX856331.1 (3), MK381202 | JQ30970.1 (3), MK412125 |
| 18 *Datura metel* | N | JU243436.1 (1), JQ434221.1 (2), JQ434220.1 (3), JU114747.1 (4) | JQ434221.1 (1), JQ434220.1 (2), JQ434220.1 (3), JU114747.1 (4) | JQ30970.1 (3), MK412125 |
| 19 *Datura stramonium* | N | KJ738546.1 (1), KJ738546.1 (2), KJ738546.1 (3), MH717194 | JX467623.1 (1), JN244737.1 (2), JN244737.1 (3), JN244375.1 (4), MK381203 |
| 20 *Datura stramonium* | N | JU243436.1 (1), JQ434221.1 (2), JQ434220.1 (3), JU114747.1 (4) | JQ434221.1 (1), JQ434220.1 (2), JQ434220.1 (3), JU114747.1 (4) | JQ30970.1 (3), MK412125 |
| 21 *Datura stramonium* | N | KJ738546.1 (1), KJ738546.1 (2), KJ738546.1 (3), MH717194 | JX467623.1 (1), JN244737.1 (2), JN244375.1 (4), MK381203 |

Continued next page
| Scientific name         | Nonendemic (N) or Endemic (E) | matK                      | psbA-trnH                  | ITS                          |
|-------------------------|--------------------------------|---------------------------|----------------------------|------------------------------|
| **Goetzea elegans**     | E                              | KJ012611.1 (1), KJ012610.1 (2), HQ384563.1 (3), MH717197, MH717198, MH717199 | MF663711.1 (1), KJ426751.1 (2), KJ426750.1 (3), MK381204, MK381205 | AY206736.1, MK412127, MK412128, MK412129 |
| **Jaltomata antillana** | N                              |                           |                            |                              |
| **Lycianthes virgata**  | N                              | AB036627.1                |                            |                              |
| **Lyctium americanum**  | N                              |                           |                            |                              |
| **Physalis angulata**   | N                              |                           |                            |                              |
| **Physalis sp. (1)**    |                                |                           |                            |                              |
| **Physalis sp. (2)**    |                                |                           |                            |                              |
| **Physalis cordata**    | N                              | JQ589235.1 (1), QJ589234.1 (2), QJ589233.1 (3) |                            |                              |
| **Physalis ignota**     | N                              |                           |                            |                              |
| **Physalis pubescens**  | N                              | EF438977.1 (1), EF438975.1 (2), EF438974.1 (3), EF438944.1 (4) | KY263837.1 (1), KY263836.1 (2), KY263835.1 (3), GQ435283.1 (4) |                              |
| **Solanum grandiiflora**| N                              | KP756859.1, MH717204, MH717205, MH717206 | MK381210, MK381211, MK381212 | KT366097.1, MK412133, MK412134, MK412135 |
| **Solanum sp.**         |                                | MH717220                  |                            |                              |
| **Solanum americanum**  | N                              | HQ235334.1 (1), KJ773158.1 (2), MH717228, MH717229 | MK381213, MK381214 | KC540786.1 (1), KC540784.1 (2), GQ478109.1 (3), GU323361.1 (4), MK412136, MK412137, MK412138, MK412139, AF244728.1* |
| **Solanum bahamense**   | N                              | MH717207, MH717208        | MK381215, MK381216 | KY699687.1, MK412138, MK412139, AF244728.1* |
| **Solanum campechiense**| N                              |                           |                              |                              |
| **Solanum capsicoides** | N                              | MH717209, MH717210        | MK381217, MK381218 |                              |
| **Solanum elaeagnifolium** | N                           | EU983576.1, MH717213, MH717214 | HM016411.1, MK381221, MK381222 |                              |
| **Solanum ensifolium**  | E                              | KJ012782.1, MH717211, MH717212 | KJ426942.1 (1), KJ426941.1 (2), MK381219, MK381220 |                              |
| **Solanum erianthum**   | N                              | KP093367.1 (1), KP093366.1 (2), KJ012783.1 (3)*, MH717215, MH717216 | HG963512.1 (1), KJ426943.1 (2), KP095977.1 (3), KP095978.1 (4), MK381223 |                              |
| **Solanum jamaicense**  | N                              | GU135067.1 (1), KC312939.1 (2), JQ589259.1 (3), MH717221 |                              |                              |
| **Solanum lycopersicum**| N                              | QJ412261.1 (1), EF438904.1 (2) | HG963710.1 (1), HQ586112.1 (2) |                              |
| **Solanum mammosum**    | N                              | KC312940.1                | JX856342.1 |                              |

*Continued next page*
Table 2. Continued.

| Scientific name            | matK-psbA-trnH ITS                                                                 |
|----------------------------|-----------------------------------------------------------------------------------|
| 368
Solanum melongena          | KJ873108.1 (1), KJ652176.1 (2), MF663709.1 (3), HF572794.1 (4), KC312941.1 (3), KX258752.1 (4), MK381226, MK381227, MK381228, MK381229, MK381230, MK381231, MK381232, MK381233, MK381234 |
| 370
Solanum nigrescens         |                                                                                   |
| 372
Solanum nudum              |                                                                                   |
| 374
Solanum polygamum          | MK412147, MK412148                                                               |
| 376
Solanum rugosum            | JQ589269.1 (1), JQ589268.1 (2), JN542598.1, MK381234                              |
| 378
Solanum seaforthianum      |                                                                                   |
| 380
Solanum torvum             | MF444846.1 (1), GU135088.1 (2), HG963489.1 (3), HG963494.1 (3), GU135408.2 (2), MK381235, MK381236, MK381237, MK381238, MK381239 |
| 382
Solanum wendlandii         |                                                                                   |
| 384
Solanum woodburyi          |                                                                                   |
| Ipomoea obscura            |                                                                                   |
| 386
Buxus microphylla          |                                                                                   |

Discussion

The main purpose of this study was to obtain DNA sequence information from DNA barcoding loci, and to understand the feasibility of using DNA barcoding to identify Solanaceae specimens in Puerto Rico, including endemic and endangered species. An additional objective was to ascertain which DNA barcode marker, individually or in combination, can provide greater discrimination within Solanaceae, which then can help in future studies or conservation efforts.

**Endemic species discrimination.** Puerto Rico has six endemic Solanaceae species: *B. densifolia*, *B. lactea*, *B. portoricensis*, *G. elegans*, *S. ensifolium*, and *S. woodburyi* [Table 2 (no. 7, 8, 10, 22, 39, and 53)]. DNA barcoding of *Brunfelsia* species has shown discrimination at the species level for both chloroplast and nuclear barcodes, especially when the barcodes are combined (Filipowicz and Renner, 2012). *Brunfelsia* consists of 50 species, high endemism, and a wide distribution in South America and the Antilles. It has been suggested that the adaptation of endemic species of *Brunfelsia* to certain habitats will contribute to the successful species identification within this genus (Filipowicz and Renner, 2012; Filipowicz et al., 2012). Our study also showed good discrimination of the genus *Brunfelsia* using all three barcodes combined (Fig. 4) or using only two barcodes, either ITS-matK or ITS-psbA-trnH (Supplemental Figs. 1 and 2). In fact, species of *Brunfelsia* could be discriminated only using the ITS barcode (Fig. 3) thus do not require assessment of other DNA barcodes.

Friedrich von Wettstein (1895–1945) proposed that *Goetzea* was part of the Solanaceae. Subsequently, *Goetzea* was classified by different authors in its own family, the Goetzeaceae (Carlquist, 1988; Cronquist, 1981; Dahlgren, 1980). However, morphological and molecular studies have placed *Goetzea* once again as part of the Solanaceae (Fay et al., 1998; Filipowicz and Renner, 2012). Our study also showed good discrimination of the genus *Brunfelsia* using all three barcodes combined (Fig. 4) or using only two barcodes, either ITS-matK or ITS-psbA-trnH (Supplemental Figs. 1 and 2). In fact, species of *Brunfelsia* could be discriminated only using the ITS barcode (Fig. 3) thus do not require assessment of other DNA barcodes.

*Goetzea* is an endangered species endemic to the Dominican Republic. The critically endangered Puerto Rican endemic species *S. woodburyi* was found in neither Cayey nor in Puerto Rico, thus the status of this plant in the wild is unknown (Vargas, 2015). Further work is needed to propagate the plant and to assess if the species is present in other parts of the island. In the wild, this plant can be confused with *S. bahamense* when there are no flowers or fruit (Strickland-Constable et al., 2010), which makes DNA barcoding important for identifying sterile specimens of this species. In our study, it was demonstrated that *matK* or *psbA-trnH*, individually or in combination, can discriminate between these morphologically similar species (Figs. 1 and 2, Supplemental Fig. 3). Unfortunately, amplification and sequencing of the nuclear region *ITS* did not work for *S. ensifolium*, nor for *Lyciantes virginia*, *S. jamaiicensis*, *Solanium* sp. 49, and *S. woodburyi*, even with the use of different primer sets.

**Nonendemic species discrimination.** Although *S. woodburyi* [Table 2 (no. 53)] is not considered an endangered...
species, it is rare and its known distribution is limited to Sierra de Luquillo (Axelrod, 2011). This is the first time *S. woodburyi* has been sampled in a molecular study. With the successful analysis of the chloroplast barcodes *matK* and *psbA-trnH*, our analysis shows a close phylogenetic relationship between *S. woodburyi*, *S. ensifolium*, and *S. bahamense* (Figs. 1 and 2). We suggest that *S. woodburyi* might be part of the Bahamense clade. Using either *matK* or *psbA-trnH*, specimens of *S. woodburyi* can be successfully discriminated at the species level. The ITS region of *S. woodburyi* was not able to be amplified.

Fig. 1. Phylogenetic neighbor joining analysis based on the *matK* barcode. Bootstraps scores are shown (100 replicates, ≥50%) for each branch. Samples with sequences obtained in this study are depicted with an asterisk (*). Discrimination at genus level can be observed using *matK* barcode.
Similar results were obtained with the species *S. erianthum*, *S. polygamum*, and *S. rugosum* [Table 2 (no. 40, 47, and 48)]. These three species were grouped in the same clade with the chloroplast barcodes, but the nuclear barcode ITS displayed more variability (Fig. 3). This information, interestingly, was used to rectify an error, which occurred during collection of the field samples. A leaf sample collected by a collaborator (employee of the Department of Natural Resources) was first labeled as Tabac/C19, which is the common name for both *S. rugosum* and *S. erianthum*. It was subsequently labeled as *S. rugosum*.
erianthum for our study. Upon analysis of the ITS region, we are able to identify the sample as *S. rugosum*, which is found in wet or moist areas at high elevations. In contrast, *S. erianthum* is found in dry limestone areas. Because the sample was obtained from Cayey, which is an area of high elevation and moisture, we were able to reliably identify the sample correctly as *S. rugosum*.

**DISCRIMINATION OF AGRICULTURALLY VALUABLE CROPS.** This study includes six agriculturally valuable crops, which are *Capsicum annuum*, *C. frutescens*, *C. chinense*, *S. lycopersicum*, *Lyium americanum*, and *Physalis angulata*.
S. melongena, and Physalis pubescens. The genus that showed least species discrimination in our study was Capsicum, which is consistent with previous studies (Olmstead et al., 2008; Walsh and Hoot, 2001). However, the concatenated phylogenetic analysis showed better resolution in species discrimination of the Capsicum species compared with the single and double DNA barcode analysis (Fig. 4). Capsicum is considered to be native to the new world and domesticated by Native Americans (Moscone et al., 2007). As a result, in comparison with other Solanaceae species, Capsicum species have been more recently diverged, thus making species identification with only chloroplast barcodes not feasible. In Puerto Rico, the most cultivated and consumed Capsicum species are C. chinense, C. annuum, and C. frutescens [Table 2 (no. 11, 12, and 13)]. These three species are some of the most economically important crops of this genus (Carrizo García et al., 2016). In the present study, ITS showed more discrimination between C. chinense, C. annuum, and C. frutescens, with less variability observed between C. chinense and C. frutescens (Fig. 3). Previously, Walsh and Hoot (2001) studied the species relationships of Capsicum using other chloroplasts regions and the nuclear region, and obtained similar results. Although the chloroplast
region was not able to discriminate between species, the nuclear region was able to provide significant variation to discriminate at species level.

*S. lycopersicum* is easily discriminated with individual and concatenated DNA barcodes in this study. A close relationship between *S. lycopersicum* with *S. seafortonianum* and *S. americanum* was observed in individual and concatenated phylogenetic analysis, similar to previous studies (Marshall et al., 2001; Melotto-Passarin et al., 2008; Weese and Bohs, 2007). In all of these studies the closest species to *S. lycopersicum* was *S. tuberosum* (potato), and as a result, *S. lycopersicum* has been considered part of the potato clade (Bohs, 2005). Moreover, *S. melongena* showed a closer relationship with *S. elaeagnifolium* when evaluated with the chloroplast DNA barcodes. This relationship has been observed in previous phylogenetics analyses that included chloroplast and nuclear DNA barcodes (Weese and Bohs, 2007; Zhang et al., 2013). In contrast, we observed a slightly closer relationship with *S. torvum* in the ITS and the concatenated phyllogenetic analyses.

Another agriculturally important crop genus included in this study is *Physalis*, represented by *P. pubescens* and *P. angulata*. Specimens of *Physalis* are difficult to identify to species morphologically. A recent study shows that the chloroplast barcode *psbA-trnH* was able to identify species of this genus (Feng et al., 2018). In our study, the DNA barcode that showed higher variability within *Physalis* was *psbA-trnH*, and when concatenated barcodes were used better discrimination was obtained (Fig. 4, Supplemental Fig. 2). Additionally, specimens of unknown species were added to our study from different resources to test if the neighbor joining analysis of the different barcodes was capable of identifying them to species. These species were *Solanum* sp. 49, *Physalis* sp. 52, *Physalis* sp. 17, and *Brunfelsia* sp. For the sample *Solanum* sp. 49 [Table 2 (no. 33)] *matK* was able to group it with *S. mammosum*. The unknown samples *Physalis* sp. 52 and *Physalis* sp. 17 [Table 2 (no. 27 and 28)] were identified as *Physalis ignota* and *Physalis pubescens*, respectively (Fig. 4, Supplemental Fig. 2). Traditional taxonomic identification also confirmed that *Solanum* sp. 49 was *Solanum mammosum*. Although the *Solanum* and *Physalis* species were identified with DNA barcoding, the unknown *Brunfelsia* species [Table 2 (no. 5)] could not be matched to any of the Puerto Rico *Solanaceae*, nor ones from GenBank. This plant was collected from a private citizen’s garden, who received it as a gift. Our results suggest that this plant does not correspond to any of the known *Brunfelsia* species from Puerto Rico and is not represented in the GenBank database.

Conclusions

Our study suggests that DNA barcoding can be used to identify the *Solanaceae* species of Puerto Rico. This molecular technique can be useful for future identification of the endemic and nonendemic plants of Puerto Rico for future conservation efforts or other studies. The majority of species were successfully identified with the barcodes individually. *MatK* was the most conserved region, and *ITS* the most variable region. As shown in previous studies, the *psbA-trnH* barcode by itself showed good discrimination in species identification. The concatenated analysis of *matK, psbA-trnH*, and *ITS* gave more resolution in comparison with the other analyses.

Literature Cited

Axelrod, F.S. 2011. A systematic vademecum to the vascular plants of Puerto Rico. Bot. Res. Inst. Texas Press, Fort Worth, TX.

Baum, D.A., R.L. Small, and J.F. Wendel. 1998. Biogeography and floral evolution of baobabs (*Adansonia, Bombacaceae*) as inferred from multiple data sets. Syst. Biol. 47:181–207.

Bohs, L. 2005. Major clades in *Solanum* based on ndhF sequence data, p. 27–49. In: T. Keating, R. Hollowell, and V. Croat (eds.). A festschrift for William G. D’Arcy: The legacy of a taxonomist. Missouri Bot. Garden Press, Chicago, IL.

Bolson, M., E. de Camargo Smidt, M.L. Brotto, and V. Silva-Pereira. 2015. ITS and trnH-psbA as efficient DNA barcodes to identify threatened commercial woody angiosperms from southern Brazilian Atlantic rainforests. PLoS One 10:1–18.

Carlquist, S. 1988. Wood anatomy and relationships of Dukecodiesae and Goetzeeae. Intl. Assn. Wood Anatomists (IAWA) J. 9:3–12.

Carrizo Garcia, C., M. Barfuss, E.M. Sehr, G.E. Barboza, R. Samuel, E.A. Moscone, and F. Ehrendorfer. 2016. Phylogenetic relationships, diversification and expansion of chili peppers (*Capsicum, Solanaceae*). Ann. Bot. 118:35–51.

CBOL Plant Working Group. 2009. A DNA barcode for land plants. Proc. Natl. Acad. Sci. USA 106:12794–12797.

Cheng, T., C. Xu, L. Lei, C. Li, Y. Zhang, and S. Zhou. 2016. Barcoding the kingdom Plantae: New PCR primers for ITS regions of plants with improved universality and specificity. Mol. Ecol. Resour. 16:138–149.

Collins, R.A. and R.H. Crukshank. 2013. The seven deadly sins of DNA barcoding. Mol. Ecol. Resour. 13:969–975.

Crookston, A. 1981. An integrated system of classification of flowering plants. Columbia Univ. Press, New York, NY.

Cwynska, A., F.F. Hunter, and P.D. Hebert. 2006. Identifying Canadian mosquito species through DNA barcodes. Med. Vet. Entomol. 20:413–424.

Dahlgren, R.M.T. 1980. A revised system of classification of the angiosperms. Bot. J. Linn. Soc. 80:91–124.

Di Pinto, A., P. Marchetti, A. Mottola, G. Bozzo, E. Bonerba, E. Ceci, M. Botarro, and G. Tantillo. 2015. Species identification in fish fillet products using DNA barcoding. Fish. Res. 170:9–13.

Doyle, J.J. and J.L. Doyle. 1990. Isolation of plant DNA from fresh tissue. Focus 12:13–15.

Dunning, L.T. and V. Savolainen. 2010. Broad-scale amplification of matK for DNA barcoding plants, a technical note. Bot. J. Linn. Soc. 164:1–9.

Fay, M.F., R.G. Olmstead, J.E. Richardson, E. Santiago, G.T. Prance, and M.W. Chase. 1998. Molecular data support the inclusion of *Dukecodies cestroides* in *Solanaceae*. Kew Bull. 53:203–212.

Feliner, G.N. and J.A. Rosselló. 2007. Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. Mol. Phylogenet. Evol. 44:911–919.

Feng, S., K. Jiao, Y. Zhu, H. Wang, M. Jiang, and H. Wang. 2018. Molecular identification of species of *Physalis* (*Solanaceae*) using a candidate DNA barcode: The chloroplast *psbA-trnH* intergenic region. Genome 61:15–20.

Filipowicz, N., M. Nee, and S. Renner. 2012. Description and molecular diagnosis of a new species of *Brunfelsia* (*Solanaceae*) from the Bolivian and Argentinean Andes. PhytoKeys 10:83.

Filipowicz, N. and S.S. Renner. 2012. *Brunfelsia* (*Solanaceae*): A genus evenly divided between South America and radiations on Cuba and other Antillean islands. Mol. Phylogenet. Evol. 64:1–11.

González, M.A., C. Baraloto, J. Engel, S.A. Mori, P. Petronelli, B. Riéra, A. Roger, C. Thebaud, and J. Chavez. 2009. Identification of amazonian trees with DNA barcodes. PLoS One 4:e7483.

Hall, T.T. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids Symp. Ser 41:95–98.
Hebert, P.D.N., A. Cywinska, S.L. Ball, and J.R. deWaard. 2003. Biological identifications through DNA barcodes. Proc. Biol. Sci. 270:313–321.

Hebert, P.D.N., E. Penton, J.M. Burns, D.H. Janzen, and W. Hallwachs. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proc. Natl. Acad. Sci. USA 101:14812–14817.

Helmer, E.H., O. Ramos, T. López, M. Quinones, and W. Diaz. 2002. Mapping the forest type and land cover of Puerto Rico, A component of the Caribbean biodiversity hotspot. Caribb. J. Sci. 38:165–183.

Hollingsworth, P.M. 2011. Refining the DNA barcode for land plants. Proc. Natl. Acad. Sci. USA 108:19451–19452.

Hollingsworth, P.M., S.W. Graham, and D.P. Little. 2011. Choosing and using a plant DNA barcode. PLoS One 6:e19254.

Howard, R.A. 1966. Notes on some plants of Puerto Rico. J. Arnold Arbor. 47:138–139.

Knapp, S. 2009. Synopsis and lectotypification of *Solanum* (Solanaceae) species endemic in the West Indies. An. Jardin Botanico Madr. 1979:665–84.

Kress, W.J. 2017. Plant DNA barcodes: Applications today and in the future. J. Syst. Evol. 55:291–307.

Kress, W.J. and D.L. Erickson. 2007. A two-locus global DNA barcode for land plants: The coding rbcL gene complements the non-coding trnH-psbA spacer region. PLoS One 2:e508.

Kress, W.J., D.L. Erickson, F. Jones, N.G. Swenson, R. Perez, E. Bermingham, and O. Sanjur. 2009. Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. Proc. Natl. Acad. Sci. USA 106:18621–18626.

Kress, W.J., K.J. Wurdack, E.A. Zimmer, L.A. Weigt, and D.H. Janzen. 2005. Use of DNA barcodes to identify flowering plants. Proc. Natl. Acad. Sci. USA 102:8369–8374.

Kumar, S., G. Stecher, and K. Tamura. 2016. MEGA7: Molecular evolutionary genetics analysis Version 7.0 for bigger datasets. Mol. Biol. Evol. 33:1870–1874.

Li, D.Z., L. Gao, H. Li, H. Wang, X. Ge, J. Liu, and G. Duan. 2011. Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. Proc. Natl. Acad. Sci. USA 108:19641–19646.

Li, X., Y. Yang, R. Henry, M. Rossetto, Y. Wang, and S. Chen. 2014. Plant DNA barcoding: From gene to genome. Biol. Rev. Camb. Philos. Soc. 90:157–166.

Ligier, H.A. 1995. Descriptive flora of Puerto Rico and adjacent islands. Univ. Puerto Rico Press, Rio Piedras, PR.

Marshall, J.A., S. Knapp, M.R. Davey, J.B. Power, E.C. Cocking, M.D. Bennett, and A. Cox. 2001. Molecular systematics of *Solanum* section *Lycopersicum* (Lycopersicon) using the nuclear ITS rDNA region. Theor. Appl. Genet. 103:1216–1222.

Melotto-Passarin, D.M., I.J. Berger, K. Dressano, V.F. De Martin, G. Conde, G.C. Oliveira, and H. Carrer. 2008. Phylogenetic relationships in Solanaceae and related species based on cpDNA sequence from plastid trnL-trnF region. Crop Breed. Appl. Biotechnol. 8:85–95.

Moscone, E.A., M. Scialfaferro, M. Grabbage, N.M. Cecchini, Y.S. Garcia, R. Jarret, and F. Ehrendorfer. 2007. The evolution of chili peppers (*Capsicum*-Solanaceae): A cytogenetic perspective. Acta Hort. 745:137–170.

Musarella, R., M. Uriarte, D. Erickson, N.G. Swenson, J.K. Zimmerman, and W.J. Kress. 2014. A well-resolved phylogeny of the trees of Puerto Rico based on DNA barcode sequence data. PLoS One 9:e12843.

Myers, N., R.A. Mittermeier, C.G. Mittermeier, G.A.B. da Fonsecaavo, and J. Kent. 2000. Biodiversity hotspots for conservation priorities. Nature 403:853–858.

Olmstead, R.G., V.F. Garcia, H.A. Migid, S.M. Collier, L. Bohs, and E. Santiago-Valentin. 2008. A molecular phylogeny of the Solanaceae. Taxon 57:1159–1181.

Olmedo, R.G., J.A. Sweere, R.E. Spangler, L. Bohs, and J.D. Palmer. 1999. Phylogeny and provisional classification of the Solanaceae based on chloroplast DNA, p. 111–137. In: J. Nee, M. Symon, D. Lester, and R. Jessop (eds.). Solanaceae IV: Advances in biology and utilization. Royal Botanic Gardens, London, UK.

Ratnasingham, S. and P.D. Herbert. 2007. BOLD: The barcode of life data system. Mol. Ecol. Notes 7:355–364.

Rolo, E.A., A.R. Oliveira, C.G. Dourado, A. Farinha, M.T. Rebelo, and D. Dias. 2013. Identification of sarcosaprophagous diptera species through DNA barcoding in wildlife forensics. Forensic Sci. Int. 228:160–164.

Sang, T., D.J. Crawford, and T.F. Stuessy. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). Amer. J. Bot. 84:1120–1136.

Selvaraj, D., R.K. Sarma, and R. Sathishkumar. 2008. Phylogenetic analysis of chloroplast matK gene from Zingiberaceae for plant DNA barcoding. Bioinformation 3:24–27.

Spooner, D.M. 2009. DNA barcoding will frequently fail in complicated groups: An example in wild potatoes. Amer. J. Bot. 96:1177–1189.

Storchová, H. and M.S. Olson. 2007. The architecture of the chloroplast psbA-trnH non-coding region in angiosperms. Plant Syst. Evol. 268:235–256.

Strickland-Constable, R., H. Schneider, S.W. Ansell, S.J. Russell, and S. Knapp. 2010. Species identity in the *Solanum bahamense* species group (Solanaceae) *Solanum subgenus Leptoselunon*. Taxon 59:209–226.

Swofford, D.L. 2002. PAUP*: Phylogenetic analysis using parsimony (*and other methods), version 4.0b3a. Sinauer Associates, Sunderland, MA.

Tate, J.A. and B.B. Simpson. 2003. Paraphyly of *Tarasa* (Malvaceae) and diverse origins of the polyploid species. Syst. Bot. 28:723–737.

Urbatsch, L.E., B.G. Baldwin, and M.J. Donoghue. 2000. Phylogeny of the coneflowers and relatives (Heliantheae: Asteraceae) based on nuclear rDNA internal transcribed spacer (ITS) sequences and chloroplast DNA restriction site data. Syst. Bot. 25:539–565.

U.S. Fish and Wildlife Service. 1985. Four plants given endangered species act protection. Endangered Species Tech. Bul. MB-815.

Vaidya, G., D.J. Lohman, and R. Meier. 2011. SequenceMatrix: Cladistics multi-gene datasets with character set and codon information. Cladistics 27:171–180.

Vargas, M. 2013. Matabuey (*Goetzia elegans*) 5-year review: Summary and evaluation. 19 Jan. 2019. <https://ecos.fws.gov/docs/five_year_review/doc4257.pdf>.

Vargas, M. 2015. Erubia (*Solanum dymophilum*) 5-year review: Summary and evaluation. 19 Jan. 2019. <https://ecos.fws.gov/docs/five_year_review/doc4652.pdf>.

Walsh, B.M. and S.B. Hoot. 2001. Phylogenetic relationships of *Capsicum* (Solanaceae) using DNA sequences from two noncoding regions: The chloroplas atpB-rbcL spacer region and nuclear waxy introns. Intl. J. Plant Sci. 162:1409–1418.

Weese, T.L. and L. Bohs. 2007. A three-gene phylogeny of the genus *Solanum* (Solanaceae). Syst. Bot. 32:445–463.

White, T.J., T. Bruns, S. Lee, and J. Taylor. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR protocols: A Guide to Methods and Applications*. Academic Press, New York, NY.

Yu, J., J.H. Xue, and S.L. Zhou. 2011. New universal matK primers for DNA barcoding angiosperms. J. Syst. Evol. 49:176–181.

Zhang, W., X. Fan, S. Zhu, H. Zhao, and L. Fu. 2013. Species-specific identification from incomplete sampling: Applying DNA barcodes to monitor invasive solanum plants. PLoS One 8:e55927.

Zona, S. 1989. Leaf anatomy of the Goetzieaeae. Aliso J. Systematic Evol. Bot. 12:303–312.
Supplemental Fig. 1. Neighbor joining analysis of the concatenated sequences of ITS + matK. Bootstraps scores are shown (100 replicates, ≥50%) for each branch.
Supplemental Fig. 2. Neighbor joining analysis of the concatenated sequences of ITS + *psbA-trnH*. Bootstraps scores are shown (100 replicates, ≥50%) for each branch.
Supplemental Fig. 3. Neighbor joining analysis of the concatenated sequences of matK+psbA-trnH. Bootstraps scores are shown (100 replicates, ≥50%) for each branch.
Supplemental Table 1. Interspecific and intraspecific pairwise distance for the *MatK* DNA barcode.

| No. | Species                    | Intrasp. difference | Intersp. difference |
|-----|----------------------------|---------------------|---------------------|
| 1   | *Acnistus arborescens*     | 0.004               | —                   |
| 2   | *Browallia americana*      | 0.004               | —                   |
| 3   | *Brugmansia × candida*     | 0.000               | 0.004               |
| 4   | *Brugmansia suaveolens*    | 0.001               | 0.004               |
| 5   | *Brunfelsia americana*     | 0.004               | 0.001               |
| 6   | *Brunfelsia densifolia*    | 0.000               | 0.001               |
| 7   | *Brunfelsia lactea*        | 0.004               | 0.001               |
| 8   | *Brunfelsia nitida*        | —                   | 0.001               |
| 9   | *Brunfelsia portoricensis* | 0.000               | 0.001               |
| 10  | *Capsicum chinense*        | —                   | 0.000               |
| 11  | *Capsicum frutescens*      | 0.000               | 0.000               |
| 12  | *Capsicum annuum*          | 0.000               | 0.000               |
| 13  | *Cestrum citrifolium*      | 0.000               | 0.002               |
| 14  | *Cestrum diurnum*          | 0.000               | 0.002               |
| 15  | *Cestrum macrophyllum*     | 0.002               | 0.002               |
| 16  | *Cestrum nocturnum*        | 0.001               | 0.002               |
| 17  | *Datura inoxia*            | 0.002               | 0.005               |
| 18  | *Datura metel*             | 0.000               | 0.005               |
| 19  | *Datura stramonium*        | 0.002               | 0.005               |
| 20  | *Goetzea elegans*          | 0.000               | —                   |
| 21  | *Lycium americanum*        | —                   | —                   |
| 22  | *Lycianthes virgata*       | —                   | —                   |
| 23  | *Physalis angulata*        | 0.000               | 0.000               |
| 24  | *Physalis pubescens*       | 0.000               | 0.000               |
| 25  | *Physalis ignota*          | 0.000               | 0.000               |
| 26  | *Solanandra grandiflora*   | 0.017               | —                   |
| 27  | *Solanum americanum*       | 0.000               | 0.017               |
| 28  | *Solanum bahamense*        | 0.000               | 0.017               |
| 29  | *Solanum capsicoides*      | 0.001               | 0.017               |
| 30  | *Solanum elaeagnifolium*   | 0.003               | 0.017               |
| 31  | *Solanum ensifolium*       | 0.000               | 0.017               |
| 32  | *Solanum erianthum*        | 0.000               | 0.017               |
| 33  | *Solanum jamaicense*       | 0.001               | 0.017               |
| 34  | *Solanum mammosum*         | —                   | 0.017               |
| 35  | *Solanum melongena*        | 0.000               | 0.017               |
| 36  | *Solanum nudum*            | 0.009               | 0.017               |
| 37  | *Solanum polygamum*        | 0.018               | 0.017               |
| 38  | *Solanum rugosum*          | 0.002               | 0.017               |
| 39  | *Solanum seaforthianum*    | —                   | 0.017               |
| 40  | *Solanum torvum*           | 0.000               | 0.017               |
| 41  | *Solanum viarum*           | 0.001               | 0.017               |
| 42  | *Solanum woodburyi*        | 0.000               | 0.017               |
| 43  | *Solanum lycopersicum*     | 0.011               | 0.017               |

| Outgroups | Minimum difference | Maximum difference |
|-----------|--------------------|--------------------|
| *Ipomoea obscura* | 0.200 | 0.237 |
| *Buxus microphylla* | 0.166 | 0.296 |
Supplemental Table 2. Interspecific and intraspecific pairwise distance for the ITS DNA barcode.

| No. | Species                      | Intraspecific difference | Interspecific difference |
|-----|------------------------------|--------------------------|--------------------------|
| 1   | Acnistus arborescens         | 0.008                    | —                        |
| 2   | Browallia americana          | —                        | —                        |
| 3   | Brugmansia × candida         | 0.000                    | 0.015                    |
| 4   | Brugmansia suaveolens        | 0.015                    | 0.015                    |
| 5   | Brunfelsia americana         | 0.001                    | 0.010                    |
| 6   | Brunfelsia densifolia        | 0.000                    | 0.010                    |
| 7   | Brunfelsia lactea            | 0.000                    | 0.010                    |
| 8   | Brunfelsia nitida            | 0.000                    | 0.010                    |
| 9   | Brunfelsia portoricicensis   | 0.000                    | 0.010                    |
| 10  | Capsicum chinense            | —                        | 0.050                    |
| 11  | Capsicum frutescens          | 0.008                    | 0.050                    |
| 12  | Capsicum annuum              | 0.006                    | 0.050                    |
| 13  | Cestrum citrifolium          | 0.000                    | 0.007                    |
| 14  | Cestrum diurnum              | 0.005                    | 0.007                    |
| 15  | Cestrum macrophyllum         | 0.000                    | 0.007                    |
| 16  | Cestrum nocturnum            | 0.008                    | 0.007                    |
| 17  | Datura inoxia                | 0.014                    | 0.007                    |
| 18  | Datura metel                 | 0.000                    | 0.007                    |
| 19  | Datura stramonium            | 0.000                    | 0.007                    |
| 20  | Goetzea elegans              | 0.000                    | —                        |
| 21  | Lycium americanum            | —                        | —                        |
| 22  | Lycianthes virgata           | —                        | —                        |
| 23  | Physalis angulata            | 0.016                    | 0.011                    |
| 24  | Physalis pubescens           | 0.000                    | 0.011                    |
| 25  | Physalis cordata             | —                        | 0.011                    |
| 26  | Physalis ignota              | 0.000                    | 0.011                    |
| 27  | Solandra grandiflora         | 0.001                    | —                        |
| 28  | Solanum americanum           | 0.005                    | 0.088                    |
| 29  | Solanum bahamense            | 0.011                    | 0.088                    |
| 30  | Solanum capsicoides          | 0.000                    | 0.088                    |
| 31  | Solanum elaeagnifolium       | 0.033                    | 0.088                    |
| 32  | Solanum ensifolium           | —                        | 0.088                    |
| 33  | Solanum erianthum            | 0.013                    | 0.088                    |
| 34  | Solanum jamaicensense        | 0.007                    | 0.088                    |
| 35  | Solanum mammosum             | 0.000                    | 0.088                    |
| 36  | Solanum melongena            | 0.002                    | 0.088                    |
| 37  | Solanum nudum                | —                        | 0.088                    |
| 38  | Solanum polygamum            | 0.162                    | 0.088                    |
| 39  | Solanum rugosum              | 0.012                    | 0.088                    |
| 40  | Solanum seaforthianum        | 0.284                    | 0.088                    |
| 41  | Solanum nigrescens           | —                        | 0.088                    |
| 42  | Solanum torvum               | 0.045                    | 0.088                    |
| 43  | Solanum viarum               | 0.011                    | 0.088                    |
| 44  | Solanum wendlandii           | —                        | 0.088                    |
| 45  | Solanum lycopersicum         | 0.000                    | 0.088                    |

**Outgroups**

| Species                  | Minimum difference | Maximum difference |
|--------------------------|--------------------|--------------------|
| Ipomoea obscura          | 0.172              | 0.392              |
| Buxus microphylla        | 0.307              | 0.467              |
Supplemental Table 3. Interspecific and intraspecific pairwise distance for the *psbA-trnH* DNA barcode.

| No. | Species                       | Intraspecific difference | Interspecific difference |
|-----|-------------------------------|--------------------------|--------------------------|
| 1   | *Acnistus arborescens*       | 0.000                    | —                        |
| 2   | *Browallia americana*         | 0.000                    | —                        |
| 3   | *Brugmansia × candida*        | 0.001                    | 0.011                    |
| 4   | *Brugmansia suaveolens*      | 0.016                    | 0.011                    |
| 5   | *Brunfelsia americana*        | 0.000                    | 0.010                    |
| 6   | *Brunfelsia densifolia*       | 0.000                    | 0.010                    |
| 7   | *Brunfelsia lactea*           | 0.000                    | 0.010                    |
| 8   | *Brunfelsia portoricensis*    | 0.000                    | 0.010                    |
| 9   | *Capsicum chinense*           | —                        | 0.000                    |
| 10  | *Capsicum frutescens*         | 0.000                    | 0.000                    |
| 11  | *Capsicum annuum*             | 0.000                    | 0.000                    |
| 12  | *Cestrum citrifolium*         | 0.002                    | 0.006                    |
| 13  | *Cestrum diurnum*             | 0.000                    | 0.006                    |
| 14  | *Cestrum macrophyllum*        | 0.005                    | 0.006                    |
| 15  | *Cestrum nocturnum*           | 0.004                    | 0.006                    |
| 16  | *Datura inoxia*               | 0.000                    | 0.008                    |
| 17  | *Datura metel*                | 0.000                    | 0.008                    |
| 18  | *Datura stramonium*           | —                        | 0.008                    |
| 19  | *Goetzea elegans*             | 0.000                    | —                        |
| 20  | *Lycium americanum*           | 0.000                    | —                        |
| 21  | *Lycianthes virgata*          | —                        | —                        |
| 22  | *Physalis angulata*           | 0.013                    | 0.029                    |
| 23  | *Physalis pubescens*          | 0.002                    | 0.007                    |
| 24  | *Solanandra grandiflora*      | 0.003                    | —                        |
| 25  | *Solanum americanum*          | 0.024                    | 0.054                    |
| 26  | *Solanum bahamense*           | 0.000                    | 0.054                    |
| 27  | *Solanum capsicoides*         | 0.000                    | 0.054                    |
| 28  | *Solanum elaeagnifolium*      | 0.000                    | 0.054                    |
| 29  | *Solanum ensifolium*          | 0.000                    | 0.054                    |
| 30  | *Solanum erianthum*           | 0.008                    | 0.054                    |
| 31  | *Solanum jamaicense*          | —                        | 0.054                    |
| 32  | *Solanum mammosum*            | —                        | 0.054                    |
| 33  | *Solanum melongena*           | 0.000                    | 0.054                    |
| 34  | *Solanum nudum*               | —                        | 0.054                    |
| 35  | *Solanum polygamum*           | 0.000                    | 0.054                    |
| 36  | *Solanum rugosum*             | —                        | 0.054                    |
| 37  | *Solanum seaforthianum*       | 0.000                    | 0.054                    |
| 38  | *Solanum turvum*              | 0.000                    | 0.054                    |
| 39  | *Solanum viarum*              | 0.000                    | 0.054                    |
| 40  | *Solanum woodburyi*           | 0.000                    | 0.054                    |
| 41  | *Solanum lycopersicum*        | 0.000                    | 0.054                    |

| Outgroups | Minimum difference | Maximum difference |
|-----------|--------------------|--------------------|
| *Ipomoea obscura* | 0.350               | 0.580               |
| *Buxus microphylla* | 0.608               | 0.670               |

6 J. AMER. SOC. HORT. SCI. 144(5):1–6. 2019.