A Higher-Level Nuclear Phylogenomic Study of the Carrot Family (Apiaceae)

James F. Smith
Boise State University

Publication Information
A higher-level nuclear phylogenomic study of the carrot family (Apiaceae)

James J. Clarkson1,8, Alexandre R. Zuntini1, Olivier Maurin1, Stephen R. Downie2, Gregory M. Plunkett3, Antoine N. Nicolas4, James F. Smith5, Mary Ann E. Feist6, Karime Gutierrez7, Panagiota Malakasi1, Paul Bailey1, Grace E. Brewer1, Niroshini Epitawalage1, Sue Zmarzty1, Félix Forest1, and William J. Baker1

PREMISSE: The carrot family (Apiaceae) comprises 466 genera, which include many well-known crops (e.g., aniseed, caraway, carrots, celery, coriander, cumin, dill, fennel, parsley, and parsnips). Higher-level phylogenetic relationships among subfamilies, tribes, and other major clades of Apiaceae are not fully resolved. This study aims to address this important knowledge gap.

METHODS: Target sequence capture with the universal Angiosperms353 probe set was used to examine phylogenetic relationships in 234 genera of Apiaceae, representing all four currently recognized subfamilies (Apioideae, Azorelloideae, Mackinlayoideae, and Saniculoideae). Recovered nuclear genes were analyzed using both multispecies coalescent and concatenation approaches.

RESULTS: We recovered hundreds of nuclear genes even from old and poor-quality herbarium specimens. Of particular note, we placed with strong support three incertae sedis genera (Platysace, Klotzchia, and Hermas); all three occupy isolated positions, with Platysace resolved as sister to all remaining Apiaceae. We placed nine genera (Apodicarpum, Bonannia, Grafia, Haplosciadium, Microsciadium, Physotrichia, Ptychotis, Tricholaser, Xatardia) that have never previously been included in any molecular phylogenetic study.

CONCLUSIONS: We provide support for the maintenance of the four existing subfamilies of Apiaceae, while recognizing that Hermas, Klotzchia, and the Platysace clade may each need to be accommodated in additional subfamilies (pending improved sampling). The placement of the currently apioid genus Phlyctidocarpa can be accommodated by the expansion of subfamily Saniculoideae, although adequate morphological synapomorphies for this grouping are yet to be defined. This is the first phylogenetic study of the Apiaceae using high-throughput sequencing methods and represents an unprecedented evolutionary framework for the group.

KEY WORDS: Angiosperms353; Apiales; molecular phylogenetics; target sequence capture; tree of life; Umbelliferae.
The Apiaceae (syn. Umbelliferae) is a large and cosmopolitan angiosperm family with 466 accepted genera and approximately 3820 species (Plunkett et al., 2018). Members of the family range from small herbs to trees and can be found in a wide range of habitats from temperate forests in the northern hemisphere to subantarctic coastlines (see Fig. 1). It is one of the most important edible plant groups, containing numerous vegetables, herbs, and spices. Carrots (Daucus carota) and parsnips (Pastinaca sativa) are among the world’s most extensively cultivated root crops, and celery (Apium graveolens) and fennel (Foeniculum vulgare) are widely cultivated “aboveground” vegetables. However, the family is perhaps best known for its herbs and spices, including aniseed (Pimpinella anisum), caraway (Carum carvi), coriander (Coriandrum sativum), cumin (Cuminum cyminum), dill (Anethum graveolens), and parsley (Petroselinum crispum). In contrast, it also includes notorious toxic species such as hemlock (Conium maculatum) and giant hogweed (Heracleum mantegazzianum). Despite the familiarity and importance of so many members of Apiaceae, uncertainties remain in the family’s classification due to gaps in our phylogenetic understanding.

The most influential and widely adopted classification of Apiaceae was that of Drude (1898), which was updated and slightly modified by Pimenov and Leonov (1993). However, the tribes recognized in these classifications are based primarily on fruit characters that do not always reflect monophyletic groups (see Downie et al., 2010). Early molecular phylogenetic analyses of the Apiaceae (Downie and Katz-Downie, 1996; Plunkett et al., 1996a, 1996b) showed that fruit characters have evolved multiple times in parallel resulting in nonmonophyletic tribes. Most recent studies (e.g., Nicolas and Plunkett, 2014; Plunkett et al., 2018) subdivided Apiaceae into four subfamilies, Apioideae, Azorelloideae, Mackinlayoideae and Saniculoideae, as proposed by Plunkett et al. (2004). By far the largest subfamily is Apioideae, and within this subfamily, 21 tribes and 20 other informal clades have been recognized based on results of phylogenetic analyses of plastid DNA and/or the internal transcribed spacer (ITS) region of nuclear ribosomal DNA (summarized by Downie et al., 2001, 2010). Here, we follow the classifications previously presented for each of the four subfamilies: Apioideae (Downie et al., 2010) (with additions of new tribes provided by Magee et al. [2010] and Zhou et al. [2009]), Saniculoideae (Calviño and Downie, 2007), and Azorelloideae and Mackinlayoideae (Nicolas and Plunkett, 2009). The currently recognized subfamilies, tribes, and other major clades of Apiaceae based on molecular evidence are presented in Table 1.

Over the last 25 years, data from Sanger sequencing has proved highly effective in elucidating phylogenetic relationships in Apiaceae. The earliest studies in the family used plastid DNA and found that the traditionally recognized subfamilies were largely monophyletic (with the significant exception of the now synonymized Hydrocotyleoideae, which was polyphyletic) but the tribes of Apioideae were polyphyletic (Plunkett et al., 1996a, 1996b). The internal transcribed spacer of nuclear ribosomal DNA (ITS) has been useful in understanding the problematic tribal relationships in Apioideae (summarized by Downie et al., 2010). The majority of molecular phylogenetic studies of Apiaceae relied on plastid loci (e.g., Downie et al., 1996, 2000; Plunkett et al., 1996a; Calviño and Downie, 2007; Nicolas and Plunkett, 2009, 2012, 2014; Magee et al., 2010), ITS (e.g., Downie and Katz-Downie, 1996; Spalik et al., 2004; Valiejo-Roman et al., 2006; Spalik and Downie, 2007; Banasiak et al., 2013), or a combination of the two (e.g., Spalik et al., 2009, 2010, 2019; Feist et al., 2012; Banasiak et al., 2016; Calviño et al., 2016; Smith et al., 2018; Ottenlips et al., 2020; Mousavi et al., 2021). More recently, high-throughput DNA sequencing has been used in Apiaceae to sequence the carrot genome (Iorizzo et al., 2016), whole chloroplasts (Downie and Jansen, 2015; Tan, 2020), and transcriptions (Wen et al., 2020) and to explore population genetics (Ottenlips et al., 2021), but these were deep rather than broad studies that gathered large quantities of sequence data but within only a few taxa. Apiaceae is thus a worthy focus for a high-throughput sequencing study spanning the entire family utilizing the unexplored potential of the nuclear genome for phylogenetic inference.

Despite the sizable body of phylogenetic research on Apiaceae, fundamental uncertainties remain that must be addressed to consolidate systematic understanding across the family. We highlight three key areas of concern: (1) poorly understood relationships among some early divergences; (2) subfamily delimitation, especially the circumscription and placement of Saniculoideae with respect to Apioideae; and (3) relationships among the currently recognized tribes and informal clades of Apioideae. Resolving relationships among the early divergences in Apiaceae and the broader order Apiales has proved difficult in previous studies due to a lack of resolution and/or support. The major groups concerned are Myodocarpaceae (the sister group to Apiaceae), which is primarily New Caledonian, the isolated Australian genus Platysace, the primarily Australian subfamily Mackinlayoideae, the Brazilian genus Klotzschia, the primarily Andean and Sub-Antarctic-Australasian subfamily Azorelloideae, the South African genus Hermas, and the large cosmopolitan Saniculoideae + Apioideae clade. Delimiting Saniculoideae and Apioideae is particularly challenging. Three mostly African genera (Choritaeinca, Lichtensteiniæa, Marlothiella) were shown to form a grade at the base of the Apioideae in previous studies (e.g., Magee et al., 2010; Nicolas and Plunkett, 2014). However, the placement of another African genus Phlyctidocarpa, which has traditionally been treated in Apioideae, is particularly problematic. Phlyctidocarpa was shown to be sister to the Steganotaenieae clade (Saniculoideae sensu lato), calling into question the integrity of both subfamilies Saniculoideae and Apioideae (Magee et al., 2010). The authors suggested addressing this problem by sinking Saniculoideae into Apioideae. However, a subsequent study placed Phlyctidocarpa as a sister to the broadly defined Saniculoideae sensu lato clade (i.e., including both Steganotaenieae and Saniculoideae; Nicolas and Plunkett, 2014). Therefore, the recent treatment of Plunkett et al. (2018) used a more cautious approach by maintaining the four-subfamily system until a data set could be presented which more clearly resolves relationships at the root of the Apioideae with satisfactory levels of support. Apioideae is the largest subfamily in Apiaceae, with approximately 90% of all recognized genera. It is the most taxonomically complex group at both the generic and tribal levels because many morphological characters vary continuously across groups, making circumscription difficult (Downie et al., 1998; Katz-Downie et al., 1999; Nicolas and Plunkett, 2009; Valiejo-Roman, 2012; Jimenez-Mejias and Vargas, 2015; Banasiak et al., 2016). This is further confounded by the frequent convergent evolution of morphological characters (Spalik and Downie, 2001; Plunkett et al., 2018) often resulting in polyphyletic genera (Downie et al., 2010). Summaries of the taxa included in the currently recognized tribes and other major clades in Apioideae were presented by Downie et al. (2010). The resolution of relationships among these tribes, as well as the formal recognition of several major clades, must await supporting data.
FIGURE 1. A selection of the morphological diversity found in Apiaceae. (A) *Actinotus helianthi* (Mackinlayoideae), (B) *Xanthosia tomentosa* (Mackinlayoideae), (C) *Azorella* sp. (Azorelloideae), (D) *Thaspium chapmanii* (Apioideae), (E) *Klotzschia rhizophylla* (incertae sedis), (F) *Anethum graveolens* (Apioideae), (G) *Sanicula europaea* (Saniculoideae), (H) *Steganotaenia araliae* (Saniculoideae). Photo credits: (A), Ori Fragman-Sapir (B, G), Tiziana Ulian (C), Mary Ann Feist (D), Leandro Freitas/William Milliken (E), Igor Sheremetyev (F), and Hans Hillewaert (H).
### TABLE 1. A summary of the current classification of Apiaceae. Groupings for each subfamily follow: Apioideae (Downie et al., 2010) (with additions of new tribes provided by Magee et al. [2010]), Saniculoideae (Calviño and Downie, 2007) and Azorelloideae/Mackinlayoideae (Nicolas and Plunkett, 2009). An asterisk indicates the four groups not sampled in this study. Note: the final column remains blank for tribes that do not contain subtribes.

<table>
<thead>
<tr>
<th>No.</th>
<th>Subfamily</th>
<th>Tribe/Clade</th>
<th>Subtribal rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Apioideae</td>
<td>Aciphylleae M.F.Watson &amp; S.R.Downie</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Apioideae</td>
<td>Acronema clade</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Apioideae</td>
<td>Annesorhiza clade (Annesorhizeae Magee, C.I.Calviño, M.Liu, S.R.Downie, P.M.Tilney &amp; B.E. van Wyk; Magee et al., 2010)*</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Apioideae</td>
<td>Apieae Takht. ex V.M.Vinogr.</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Apioideae</td>
<td>Arcuatopterus clade*</td>
<td>Heteromorpha clade</td>
</tr>
<tr>
<td>6</td>
<td>Apioideae</td>
<td>Bupleureae Spreng.</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Apioideae</td>
<td>Cachrys clade</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Apioideae</td>
<td>Careae Baill.</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Apioideae</td>
<td>Chamassium clade (Chamaesieae J.Zhou &amp; F.D.Pu; Zhou et al., 2009)</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Apioideae</td>
<td>Choritaenieae Magee</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Apioideae</td>
<td>Conioselnum chinense Clade (North American Ligusticum clade; Zhou et al., 2020)</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Apioideae</td>
<td>Conium clade</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Apioideae</td>
<td>Coriandreae W.D.J.Koch</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Apioideae</td>
<td>Diplolophium clade</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Apioideae</td>
<td>Echinophoreae Benth. &amp; Hook. f.</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Apioideae</td>
<td>Eriogonieae Rydb.</td>
<td>Malagasy clade</td>
</tr>
<tr>
<td>17</td>
<td>Apioideae</td>
<td>Heteromorphae</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Apioideae</td>
<td>Heteromorphae</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Apioideae</td>
<td>Lichtensteinia clade (Lichtensteinieae Magee, C.I.Calviño, M.Liu, S.R.Downie, P.M.Tilney &amp; B.E. van Wyk; Magee et al., 2010)</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Apioideae</td>
<td>Marlothielleae Magee</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Apioideae</td>
<td>Oenantheae Dumort.</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Apioideae</td>
<td>Oppanaceae</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Apioideae</td>
<td>Physospermopsis clade or East Asia Clade</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Apioideae</td>
<td>Pimpinelleae Spreng.</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Apioideae</td>
<td>Pleurosympermeae M.F.Watson &amp; S.R.Downie</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Apioideae</td>
<td>Pleurospermopsis clade</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Apioideae</td>
<td>Pyramidoptereae Boiss.</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>Apioideae</td>
<td>Scandiceae Spreng.</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Apioideae</td>
<td>Scandiceae Spreng.</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>Apioideae</td>
<td>Scandiceae Spreng.</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>Apioideae</td>
<td>Scandiceae Spreng.*</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>Apioideae</td>
<td>Scandiceae Spreng.</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>Apioideae</td>
<td>Scandiceae Spreng.</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>Apioideae</td>
<td>Selinaceae Spreng.</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>Apioideae</td>
<td>Selinaceae Spreng.</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>Apioideae</td>
<td>Selinaceae Spreng.</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>Apioideae</td>
<td>Sinodelsia clade</td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>Apioideae</td>
<td>Smyrneae Spreng.</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>Apioideae</td>
<td>Tordyliieae W.D.J.Koch</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>Apioideae</td>
<td>Tordyliieae W.D.J.Koch</td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>Apioideae</td>
<td>Tordyliieae W.D.J.Koch</td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>Azorelloideae</td>
<td>Azorella clade</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>Azorelloideae</td>
<td>Asteriscium clade</td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>Azorelloideae</td>
<td>Bowlesia clade</td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>Azorelloideae</td>
<td>Diposis clade</td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>Mackinlayoideae</td>
<td>Centella clade</td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>Mackinlayoideae</td>
<td>Xanthosia clade</td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Saniculoideae</td>
<td>Saniculeae W.D.J.Koch</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>Saniculoideae</td>
<td>Steganaetaeae C.I.Calviño and S.R.Downie</td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>Incertae sedis</td>
<td>Hermas</td>
<td></td>
</tr>
<tr>
<td>52</td>
<td>Incertae sedis</td>
<td>Klotzschia</td>
<td></td>
</tr>
<tr>
<td>53</td>
<td>Incertae sedis</td>
<td>Platysace clade</td>
<td></td>
</tr>
</tbody>
</table>
Here, we present the first higher-level nuclear phylogenomic study of Apiaceae. We utilized the universal Angiosperms353 target capture probe set (Johnson et al., 2019) to sequence over half of the genera in Apiaceae, primarily with DNA sourced from herbarium specimens (Brewer et al., 2019). We used this data set to answer the following questions: (1) What are the relationships among the sub-families of Apiaceae, especially in relation to Mackinlayoideae and Azorilloideae, and the three isolated lineages, the Platyaceae clade, Klotzchia and Hermesia? (2) Is subfamily Saniculoideae nested within Apioidae? (3) What are the relationships among the currently recognized tribes and other major clades of Apioidae? In addressing these questions, we attempt not only to establish a new phylogenetic baseline for Apiaceae, but also to demonstrate the potential of the Angiosperms353 probe set in angiosperm phylogenomic studies.

MATERIALS AND METHODS

Sampling
We sampled 248 taxa belonging to 234 genera of Apiaceae (just over half of the 466 recognized genera) along with eight species representing other families in Apiales, which served as outgroup taxa. This sampling represents all four subfamilies (Apioidae, Azorilloideae, Mackinlayoideae and Saniculoideae), and 37 of the 41 tribes and other major apioid clades recognized by Downie et al. (2010). The specimens were selected to represent the breadth of taxonomic diversity, including all major clades according to Nicolas and Plunkett (2009). Included in this sampling were the following nine genera, which have never been included in any previous molecular phylogenetic study: Apodicarpum, Bonannia, Grafia, Haplosciadium, Microsciadum, Physotrichia, Psychotis, Tricholaser, and Xatartia. Our list of accepted generic names follows Plunkett et al. (2018). Material was sourced from herbarium specimens, silica-dried samples, the Kew DNA & Tissue Bank (http://dnabank.science.kew.org/homepage.html), and the living collections at the Royal Botanic Gardens, Kew (Appendix S1). To explore the root of Apiaceae, we used several other families from Apiales (sensu APG, Santa Clara, CA, USA) as outgroup taxa. Our trees were rooted with Grieselinae on the basis of earlier studies, which indicate that this family is sister to the clade comprising Apiaceae, Araliaceae, Myodocarpaceae, and Pittosporaceae (e.g., Nicolas and Plunkett, 2009, 2014).

DNA extraction and library preparation
DNA was extracted from 40 mg of herbarium material, 20 mg of silica gel-dried material (Chase and Hills, 1991), or 100 mg of fresh material using a modified CTAB extraction method, with chloroform–isoamyl alcohol ("Sevag") separation and isopropanol precipitation at –20°C (Doyle and Doyle, 1987). Plant tissue was pulverized using a Mixer Mill MM400 (Retsch GmbH, Haan, Germany). The extraction was followed by magnetic bead clean-up using AMPure XP beads (Beckman Coulter, Indianapolis, IN, USA), according to the manufacturer’s protocols. Many existing genomic DNA samples were available from the DNA & Tissue Bank at the Royal Botanic Gardens, Kew (http://dnabank.science.kew.org/homepage.html). DNA was extracted from these samples using a modified CTAB method (Doyle and Doyle, 1987), followed by cesium chloride/ethidium bromide density gradient centrifugation and dialysis. Extracted DNA was quantified using Quanta (Promega, Madison, WI, USA) or Qubit (Thermo Fisher Scientific, Inchinnan, UK) fluorometers and then separated in a 1% agarose gel to assess the average fragment size. Samples with very low concentration (not visible in a 1% agarose gel) were assessed on a 4200 TapeStation System using Genomic DNA ScreenTapes (Agilent Technologies, Santa Clara, CA, USA). DNA extracts with average fragment size above 350 bp were sonicated using an M220 Focused-ultrasonicator with microTUBES AFA Fiber Pre-Slit Snap-Cap (Covaris, Woburn, MA, USA) following the manufacturer’s protocols and with varied shearing times depending on the DNA fragment size profile, to obtain an average fragment size of 350 bp. Dual-indexed libraries for Illumina sequencing were prepared using the DNA NEBNext Ultra II Library Prep Kit and the NEBNext Multiplex Oligos for Illumina (Dual Index Primer Sets 1 and 2) according to the manufacturer protocols (New England BioLabs, Ipswich, MA, USA) and at either the recommended volumes or half these volumes. Six to 12 cycles of PCR were required for libraries depending on the initial sample concentration. Quality of libraries was evaluated on a 4200 TapeStation System using D1000 ScreenTapes and the libraries were quantified using a Quanta fluorometer. The final average library size including the adapters was ca. 500 bp, or lower when input DNA fragments were smaller than 350 bp on average.

Target enrichment and sequencing
The libraries were pooled (20–24 libraries per pool) and enriched using the Angiosperms353 probe kit (Arbor Biosoiences myBaits Target Sequence Capture Kit, Angiosperms-353 v1, Catalog #308196; Johnson et al. [2019]) following the manufacturer’s protocols (v4.0; http://www.arborbiosci.com/mybaits-manual). Hybridizations were performed at 65°C for 24 h in a Hybex Microsample Incubator (SciGene, Sunnyvale, CA, USA) with a layer of red Chill-out Liquid Wax (Bio-Rad, Hercules, CA, USA) on the surface to prevent evaporation. Enriched products were amplified with a KAPA HiFi 2X HotStart ReadyMix PCR Kit (Roche, Basel, Switzerland) for 8 cycles. PCR products were then cleaned using Agencourt AMPure XP Beads and quantified with a Quanta fluorometer (in some cases, weaker products required a second amplification for 3 to 6 cycles). Final products were run on a 4200 TapeStation System using High Sensitivity D1000 ScreenTapes to assess quality and average fragment size. The enriched pools were normalized to 6 nM and multiplexed for sequencing, to include between 24 and 384 samples, depending on the sequencing platform and service provider requirements. Library pools were multiplexed and sequenced on an Illumina MiSeq with v2 (300-cycles of 2 x 150 bppaired-end reads) or v3 (600-cycles of 2 x 300-bp paired-end reads) chemistry (Illumina, San Diego, CA, USA) at the Royal Botanic Gardens, Kew or on an Illumina HiSeq 4000 producing 2 x 150-bp paired-end reads at Genewiz (Takeley, UK) or Macrogen, (Seoul, South Korea).

Recovery of target loci, multiple sequence alignments, and gene trees
The reads of the sequencing output (.fastq files) were trimmed using Trimmomatic (Bolger et al., 2014) to remove reads with a quality score below 30 and reads that had any 4-bp window below 30, retaining reads with at least 36 bp (LEADING:30 TRAILING:30 SLIDINGWINDOW:4:30 MINLEN:36). The HybPiper pipeline version 1.3.1 (Johnson et al., 2016) was used to recover target sequences from paired reads and combined unpaired reads, using the default settings (with the exception of minimum coverage, which was set to 4x).
Reads were mapped to de-gapped medoid sequences using BLASTx (Camacho et al., 2009), since it has been found to result in longer sequences (Murphy et al., 2019). Subsequently, each gene was assembled de novo using SPAdes version 3.13.1 (Bankevich et al., 2012), and coding sequences were extracted using exonerate version 2.2 (Slater and Birney, 2005). The 353 target genes were recovered from the transcriptomes of 19 taxa that were sequenced in a previous study (Wen et al., 2020), nine from OneKP (OTPI, 2019) and five from individual studies, all available via the NCBI database (https://www.ncbi.nlm.nih.gov/). *Daucus* sequences were mined from the latest available genome assembly (GCA_001625215.1). These publicly available sequences were added to our data set and analyzed alongside the novel sequence data generated specifically for this study.

Gene matrices were aligned separately using MAFFT (Katoh and Standley, 2013) version 7 (mafft-7.419-gcc_f6.x86), with Accuracy-oriented methods (--localpair 2 --maxiterate 1000). Matrices were subsequently trimmed using phylutility (Smith and Dunn, 2008), which is available at GitHub (https://github.com/blackrim/phylutility), to delete sites that are missing 30% or more data (--clean 0.3). Gene trees from trimmed matrices were generated using IQ-TREE v2.10 (Minh et al., 2020), using ultrafast bootstrapping with data set partitioning (Chernomor et al., 2016). Due to matrix size, model selection was not feasible, and therefore we implemented the model GTR+I+G for all partitions (i.e., the same model for all partitions, but the parameters may differ between partitions). In a first iteration, we generated gene trees that were subsequently evaluated using TreeShrink (Mai and Mirarab, 2018) to identify and exclude branches that increased the diameter of each gene tree by more than 20% with centroid re-rooting (--b 20 -c). Each locus was then realigned, trimmed, and analyzed using IQ-TREE with bipartition support (i.e., support values for all splits). One thousand ultrafast bootstrap replicates (UFBootStrap) were run (--B 1000; Hoang et al., 2018) in IQ-TREE and branches with support values below 10% were collapsed (Mirarab, 2019 [Preprint]) using Newick Utilities v1.6 (Jenner and Zdobnov, 2010). Samples which contained less than 10 genes were excluded.

**Phylogeny reconstruction**

A species tree was generated by inputting IQ-TREE gene tree files based on each individual exon alignment. Extensive branch annotations were generated using ASTRAL-III version 5.5.11 (Zhang et al., 2018) using alternative quartet topologies (-t 2) indicating the local posterior probabilities of the percentage of quartets in genes. Additionally, we generated a phylogenetic tree from the concatenation of exon re-alignments (after exclusion of outliers indicated by TreeShrink). The resulting data set was analyzed as a partitioned alignment in IQ-TREE v2.10 (Minh et al., 2020) with 1000 ultrafast bootstrap replicates. The GTR+I+G model was implemented for all partitions (i.e., the same model for all partitions, but the parameters may differ between partitions). All trees were plotted in R (R Core Team, 2020) using packages ape (Paradis and Schliep, 2019), ggimage (Yu, 2019), ggtree (Yu et al., 2017), and their dependencies.

**RESULTS**

**Overview**

Overall, we obtained a total of 384 billion reads, 8.2% of which were on target (31,890,522 reads) for the 256 samples examined. Over 45,596 targets (corresponding to 353 genes) with sequences that were >50% of the target length were assembled. There is significant variation in the percentage of the target length of genes obtained for different samples (see Fig. 2), probably due to the inherent variation in DNA degradation found in the source material (old and new herbarium sheets). Only slight variation exists between the recovery statistics obtained for the four different subfamilies of Apiaceae (no significant taxonomic bias was observed). However, the most notable anomaly in the recovery statistics was found for the outgroups, which have a significantly higher number of genes above 50% of the target length (Fig. 3), because most outgroups were obtained from transcriptomic data (rather than target sequence capture data) and therefore they have a higher frequency of complete genes. The full recovery statistics table is presented in Appendix S2. Some genes (presumably the more conserved ones) are more efficiently captured than others in the Apiaceae taxa examined using the universal probes (see dark versus light columns in the heatmap in Appendix S3). The pseudo-coalescent tree (hereafter referred to as the coalescent tree for convenience) and the concatenation tree are broadly congruent; thus, the support levels for clades are mostly quoted from both data sets for comparison. Therefore, local posterior probability (LPP) values are followed by bootstrap percentage (BP) values when clades are discussed. However, where relationships differ between our analyses, these are highlighted and discussed for the relevant clades. We use the following four terms to discuss support values in the phylogenetic trees: (1) nodes with LPP = 1 and BP = 100 are described as fully supported; (2) nodes with LPP > 0.85 and BP > 100 are described as highly supported; (3) nodes with LPP 0.85 to 0.6 and BP > 95 are described as moderately supported; and (4) nodes with LPP < 0.6 are described as weakly supported.

**Subfamily relationships in Apiaceae**

Myodocarpaceae and Apiaceae are fully supported (LPP = 1, BP = 100) as sister families (see coalescent tree in Fig. 4 and concatenation tree in Appendix S4). Starting at the root of Apiaceae, *Platysace* (*Platysace clade*) is sister to all remaining Apiaceae, followed by the core Mackinlayioidae clade, the genus *Klotzschia*, subfamily Azorellioideae, the genus *Hermas*, and a clade that unites Saniculoideae sensu stricto with tribe S Stefanotaenieae (Saniculoideae sensu lato; Calviño and Downie, 2007) and *Phlyctidocarpa* of Apiaceae. Support for this set of relationships is high for all nodes except Stefanotaenieae + Saniculoideae (LPP = 0.82, BP = 90, with quartet scores indicating high gene tree incongruence). However, the placement of *Phlyctidocarpa* renders Apiideae paraphyletic, although this relationship is only moderately supported (LPP = 0.66, BP = 99, with equivocal quartet scores). The relationships between the major groups in Apiaceae can therefore be summarized as follows: (*Platysace* (Mackinlayioidae (*Klotzschia* (Azorellioideae (*Hermas* (Phlyctidocarpa + Saniculoideae sensu lato) (Apiaceae)))))), with all nodes receiving high support except the two nodes stated above.

**Monophyly of tribes and other major clades in Apioidae**

Apart from *Phlyctidocarpa* (*Phlyctidocarpaceae*), the earliest divergence in Apiaceae is a clade uniting three monotypic tribes, Lichtensteinieae, Choritaenieae and Marlothielleae (LPP = 1, BP = 100), and this clade is in turn sister to the rest of the subfamily. The following tribes and other previously defined clades form monophyletic groups within Apioidae, all with maximum support (LPP = 1,
BP = 100) unless otherwise stated: Heteromorphae, Annesorhizeae, Pleurospereae, Oenantheae, Smyrineae, Aciphylla (LPP = 0.90, BP = 100), Acronema clade, Careae, Pyramidoptereae, Pimpinelleae, Sinodielsia clade, Opopanax clade, Apieae, Cymbocarpum clade, Echinophorea, and Cachrys clade. However, our analyses indicate that three groups are paraphyletic (see Fig. 4 and Appendix S4). The Physospermopsis clade is well supported (LPP = 1, BP = 100), Physospermopsis and previously named clades with high support (LPP = 1, BP = 100): Heteromorphae, Annesorhizeae, Conioselinum chinense clade (now called the North American Ligusticum clade; see Zhou et al., 2020), Diplolophium clade, Conium clade, and Coriandreae. Within Heteromorphae, the Heteromorpha and Malagasy clades are each reconstructed as monophyletic. Within Scandiceae, subtribes Daucinae, Ferulinae, Scandicinae, and Torilidinae and the Glaucoaciadium clade are each monophyletic; Scandiceae subtribe Artediinae is monotypic. Within Tordylieae, subtribe Tordyliae and the Cymbocarpum and Lefebvrea clades are each monophyletic. The Arracacia clade of Selineae is also monophyletic.

**Placing the newly sequenced genera**

The nine newly sequenced genera are placed in the following tribes and previously delimited major clades are monotypic or only represented by a single sample in our data set, and therefore, it was not possible to evaluate their monophyly: Phylcidiocarpeae, Lichtensteinieae, Choritaenieae, Marlothielleae, Chamaesieae, Bupleureae, Komarovieae, Conioselinum chinense clade (now called the North American Ligusticum clade; see Zhou et al., 2020), Diplolophium clade, Conium clade, and Coriandreae. Within Heteromorphae, the Heteromorpha and Malagasy clades are each reconstructed as monophyletic. Within Scandiceae, subtribes Dauniae, Ferulinae, Scandicinae, and Torilidinae and the Glaucoaciadium clade are each monophyletic; Scandiceae subtribe Artediinae is monotypic. Within Tordylieae, subtribe Tordyliae and the Cymbocarpum and Lefebvrea clades are each monophyletic. The Arracacia clade of Selineae is also monophyletic.

**FIGURE 2.** Number of genes recovered. (A) Number of genes recovered per sample, categorized by percentage of target length. (B) Overall number of genes per category. The percentage of the target length is color-coded (green = >75%, yellow = 50% to 75%, red = 25% to 50% and blue = <25%).

[Diagram showing the number of genes recovered per sample, with categories by percentage of target length.]

[Text continued with details on the number of genes recovered per sample, categorized by percentage of target length.]
DISCUSSION

These data, generated using the universal Angiosperms353 probe set, constitute the largest phylogenomic data set to be analyzed to date for the Apiaceae. As predicted, the percentage of reads on target is relatively low (8.2%), due to the universal nature of Angiosperms353 probe set, but this shortcoming is offset by the time and money saved in not having to develop a custom probe set. In the future, a custom 353 probe set could be constructed, using the sequence data generated in this study, to increase probe efficiency for Apiaceae. However, despite the fairly low percentage of reads on target, we were able to recover 31,890,522 reads on target (corresponding to 353 genes). The approach resulted in trees that resolve the backbone, major clades of Apiaceae, and outgroup taxa with high levels of support almost universally. However, we note that several “key nodes” that have been problematic in previous studies here receive only moderate support, suggesting that underlying biological processes are behind these patterns.

The target sequence capture data set generated for this study yields few paralog warnings in HybPiper (mostly 0–2, see final column of Appendix S2) although high numbers of paralogs were identified in the data mined from transcriptomes/genomes (64 paralogs were recovered from Torilis scabra). While paralogy demands further investigation, especially in relation to its impact on phylogenetic inference (Maddison, 1997), it has recently been demonstrated that species tree inference in the presence of paralogs is accurate using coalescent-based approaches (Yan et al., 2020 [Preprint]). In this study, the main cause of discordance between the coalescent

FIGURE 3. Recovery plots for Apiaceae data set. Recovery statistics are presented individually for each of the four subfamilies to illustrate differences in representation of the genes.

FIGURE 4. Coalescent tree showing relationships in Apiaceae, using nuclear exons and inferred in ASTRAL III. Local posterior probability values are presented below branches, and all unlabeled nodes received maximum support. The subfamilies are shown in gray bars, the tribes and major clades are displayed as indented groups, and the tip labels show the genus and species for each taxon. Table 1 summarizes the classification that was followed. “U” refers to “unplaced”, i.e., taxa that are absent from the published classification; “N” refers to “new”, i.e., genera that have not appeared in any previous molecular phylogenetic study. Pie charts show the quartet support values for each node (blue = species tree topology quartet support; red = first alternative topology quartet support; gray = second alternative topology quartet support).
tree and the concatenation tree seems likely to be low gene recovery, rather than paralogy. Low gene recovery also reduces support values for corresponding clades. Haploid chromosome number varies widely across Apiaceae, from \( n = 3 \) in *Sium suave* to \( n = 77 \) in *Lomatium columbianum* due to dysploidy, aneuploidy, and polyplody (Pimenov et al., 2003; Plunkett et al., 2018). However, paralog number (see Appendix S2) does not appear to correlate with the ploidy level of taxa but seems to be directly linked to the data collection method (more paralogs from transcriptomic data compared to target sequence capture data).

**Delimitation and relationships**

The sister group relationship between Apiaceae and Myodocarpaceae (LPP = 1) confirms earlier studies (e.g., Nicolas and Plunkett, 2009, 2014). Myodocarpaceae is an Australasian family (primarily New Caledonian) that was formerly included in Araliaceae (e.g., by Cronquist, 1988). Its phylogenetic placement is reflected in its intermediate morphology, which is vegetatively similar to Araliaceae, but shares reproductive features with both Apiaceae and Araliaceae (Rodriguez, 1957; Plunkett and Lowry, 2001; Plunkett et al., 2004; Lowry and Plunkett, 2018).

Over the past 25 years, many molecular studies have investigated the major lineages in Apiaceae (Plunkett et al., 1996a, 1996b, 1997; Downie and Katz-Downie, 1999; Downie et al., 1998; Plunkett and Lowry, 2001), but understanding the relationships among them has proved difficult due to a lack of informative characters at key nodes (Chandler and Plunkett, 2004; Nicolas and Plunkett, 2009, 2014). Plunkett et al. (2018) has provided a detailed description of these issues. We revisit these below with respect to the four widely recognized subfamilies and the isolated genera *Platysace* from Australia (representing the *Platysace* clade, which also includes *Homalosciadium*, not sampled here), *Klotzschia* from Brazil, and *Hermas* from South Africa.

**Phylogenetic relationships**—Our results yield the following relationships among the major lineages with almost all of these major nodes receiving high support: (*Platysace* (Mackinlayoideae (*Klotzschia* (Azorelloideae (*Hermas* (Phlyctidocarpia + Saniculoideae) (Apiioideae)))))) (see Fig. 4 and Appendix S4). These relationships differ from previous findings in important ways. Chandler and Plunkett (2004) resolved ((Mackinlayoideae + *Platysace*) (Azorelloideae (Saniculoideae + Apiioideae))) using two plastid genes (*matK* and *rbcL*) and nuclear ribosomal 26S, but did not sample *Klotzschia* and *Hermas*. More recently, Nicolas and Plunkett (2009, 2014) used two noncoding plastid regions (*rpl16* intron and the *trnD-trnT* spacer) and recovered the following topology: (*Mackinlayoideae (Platysace clade (Azorelloideae (Klotzschia (Hermas (Saniculoideae + Apiioideae)))))). Other relevant studies placed *Platysace* in the Mackinlayoideae clade, based on 17 loci but with relatively sparse taxon sampling (Soltis et al., 2011), and as sister to Apiioideae with low support (Andersson et al., 2006). The differences between our findings and earlier studies may in part be attributed to plastid/nuclear gene incongruence. However, this possibility needs to be confirmed with a more extensive exploration of the plastid genome, which could be achieved by recovering plastid sequences from the off-target reads in our data sets. That said, gene sampling is low in the majority of the earlier studies compared to the new data set; therefore, a lack of informative characters is likely the primary explanation for the differences.

**Mackinlayoideae and Platysace**—Our results support the monophyly of Mackinlayoideae, consistent with the delimitation of the subfamily in more recent plastid studies (Nicolas and Plunkett, 2009, 2014). Our results agree on the exclusion of the *Platysace* clade, but disagree on the relative placement of the two groups. The node separating *Platysace* from Mackinlayoideae in our study is slightly less well supported than many (LPP = 0.93, BP = 100), with somewhat equivocal quartet support. Further investigations of gene tree conflict with additional taxon sampling (especially more species of *Platysace* and its relative *Homalosciadium*) are required to gain a greater understanding of the credibility of this relationship.

Within Mackinlayoideae, the genus *Actinotus* is sister to the rest of the subfamily, as shown in previous studies (Nicolas and Plunkett, 2009; Liu et al., 2016). *Actinotus* has an unusual floral structure (particularly its distinctive gynoecium). Melikian and Konstantinova (2006) even suggested that a new monotypic family should be erected to accommodate this "oddball" taxon, but our results do not support such a taxonomic change. The broader Actinotus clade also contains the former araliaceous genus *Aioipetatum* (Nicolas and Plunkett, 2009), but this genus was not sampled in the present study.

**Azorelloideae, Hermes, and Klotzschia**—With respect to the delimitation of subfamily Azorelloideae, our results are in conflict with other authors, who tentatively placed *Hermas* and *Klotzschia* within the subfamily after the dissolution of the polyphyletic Apiaceae subfamily Hydrocotylidoideae (Nicolas and Plunkett, 2009). The two genera and the main Azorelloideae clade form a paraphyletic group within which Saniculoideae and Apiioideae are nested. These results are consistent with the classification of Plunkett et al. (2018) who treated *Hermas* and *Klotzschia* as incertae sedis. As in the case of *Platysace*, the nodes separating *Hermas* and *Klotzschia* from the main Azorelloideae clade are not maximally supported with what equivocal quartet scores. Again, more in-depth scrutiny of gene tree conflict accompanied by added taxon sampling will be important to understand these placements better.

The 11 remaining taxa sampled from Azorelloideae form a well-supported monophyletic group. Azorelloideae is composed of three distinct clades, each receiving maximum support, corresponding to the *Asteriscium* clade, the *Azorella* clade and the Bowlesia clade established in previous studies (e.g., Andersson et al., 2006; Nicolas and Plunkett, 2009). Within Azorelloideae, several genera have been shown to be nonmonophyletic and in need of further study (Nicolas and Plunkett, 2009, 2012; Fernández et al., 2017). Plunkett and Nicolas (2017) resolved these issues in the *Azorella* clade, but additional species-level studies are required, especially in the *Asteriscium* clade.

**Saniculoideae and Apiioideae**—Since the advent of molecular systematics, there have been several alternative delimitations of subfamily Saniculoideae, based on transfers between subfamilies (e.g., Plunkett et al., 1996, 2004; Valiejo-Roman et al., 2002; Plunkett and Lowry, 2001; Chandler and Plunkett, 2004). The placement of several African genera (*Phlyctidocarpia*, *Polemoniannis*, and *Steganotaenia*) previously in Apiioideae was particularly problematic. Calviño and Downie (2007) found that *Steganotaenia* and *Polemoniannis* (tribe *Steganotaeniaceae*) were sister to tribe Saniculoideae and suggested that Saniculoideae be broadened to include both tribes. They also suggested that the Namibian genus *Phlyctidocarpia*, traditionally treated in subfamily Apiioideae, might also eventually be transferred to Saniculoideae on the basis of preliminary molecular evidence. In the
study of Magee et al. (2010), Phlyctidocarpa was resolved as sister to tribe Steganaotaeae (with moderate to weak branch support, depending upon the analysis), while Nicolas and Plunkett (2014) showed Phlyctidocarpa as sister to a clade of Steganaotaeae and Saniculeae. Magee et al. (2010) highlighted that expanding Saniculoideae by adding Steganaotaeae (sensu Calviño and Downie [2007]) makes it more difficult to define either Saniculoideae or Apioidaea on the basis of morphological synapomorphies. The genera of Saniculoideae share some defining features, but no characters found in Saniculeae and Steganaotaeae (i.e., Saniculoideae sensu lato) are not also found in Apioidaea. Their solution was to propose the inclusion of these tribes in an expanded Apioidaea. This complex situation had not been fully resolved by the time Plunkett et al. (2018) published their revised treatment of the family.

Our study yields a moderately supported group comprising the sister tribes Steganaotaeae and Saniculeae (LPP = 0.82, BP = 90, with quartet scores indicating high gene tree incongruence). Phlyctidocarpa is resolved as sister to Saniculeaeoideae, albeit rather weakly (LPP = 0.66, BP = 99; Fig. 4 and Appendix S4), although we note that this result is consistent with evidence from the plastid, a distinct genomic compartment (Calviño and Downie, 2007). Our evidence is thus consistent with both the proposed classifications of Calviño and Downie (2007; including the possible subsuming of Phlyctidocarpa in Saniculoideae) and of Magee et al. (2010). Although our placement of Phlyctidocarpa differs significantly from Magee et al. (2010) their proposed classification which sinks Saniculoideae into Apioidaea is consistent with our results. In light of this equivocal outcome, other considerations come into play, most importantly perhaps the practical usability of any classification. Nevertheless, further attention to factors causing lower support would be advisable before action is taken.

With the exclusion of Phlyctidocarpa, the remaining Apioidaea sampled in this study form a strongly supported monophyletic group, which is consistent with accepted subfamily circumscription (summarized in Downie et al., 2010).

### Prospects for a new subfamily classification in Apiaceae

Currently, a few genera of Apiaceae are incompletely classified to subfamly, which is an important impediment to communication about the family’s diversity. Our study increases the amount of data applied to the family’s phylogenetic relationships by at least two orders of magnitude and thus affords an opportunity to reflect on the potential structure of a revised, subfamily classification. Based on the evidence presented herein, the four currently recognized subfamilies, Mackinlayoideae, Azorelloideae, Saniculoideae, and Apioidaea can be maintained (a system first proposed by Plunkett et al. [2004]), thereby promoting taxonomic stability, provided appropriate adjustments are made to address the position of Phlyctidocarpa in Saniculoideae. However, the lack of obvious morphological synapomorphies for the delimitation of Saniculoideae and Apioidaea remains problematic. This could be addressed by reducing the two to a single, broad Apioidaea (sensu Magee et al., 2010), or alternatively initiating a search for as yet undiscovered characters, which may come in more subtle guises, for example in chemistry or genome biology.

Three lineages remain unclassified in our tree, Hermas, Klotzschia, and the Platsyace clade (Platsyace and the unsampled relative Homalosciadium), each of which could be accommodated in a new subfamily. Such an approach was hinted at by Nicolas and Plunkett (2009), at least with respect to the Platsyace clade, and would be a practical and informative solution that appropriately augments the four accepted subfamilies. Alternatively, these three groups might be treated as incertae sedis, but in reality their position is far from uncertain. We can be confident, for example, that they do not fall within any of the accepted subfamilies, as supported by our data. Moreover, by recognizing them as three subfamilies, their distinctness from each other would be represented clearly in future classifications. Complete generic sampling is required to further explore the case for this putative seven-sub-family system, as well as additional species sampling in the three critical genera. Target capture, with its ready application to herbarium samples, is the ideal method to achieve this.

### Tribes and other major clades in Apioidaea

The integrity of the tribes and other major clades of Apioidaea inferred on the basis of cladistic analyses of molecular data, summarized initially by Downie et al. (2001) and later expanded (Downie et al., 2010), are largely upheld in the present study. These summaries were primarily based on results of analyses of ITS data, a locus that proved too small and too rapidly evolving to infer deep-level (i.e., intertribal) relationships within the subfamily (Downie et al., 1998; Katz-Downie et al., 1999). Subfamily Apioidaea is notorious for its abundance of nonmonophyletic genera, most of which are species rich but poorly delimited. For example, Downie et al. (2010) listed 18 genera having species that were assigned to two or more clades, including the genus Ligusticum whose species have been spread across to six tribes/major clades on the basis of molecular phylogenetic studies (Zhou et al., 2020).

With regard to specific groups, our trees show the genus Parasilaus (currently in Komarovieae) is here embedded within the Physospermopsis clade with high support (LPP = 1.0, BP = 100). In previous studies, Komarovieae and the Physospermopsis clade form monophyletic sister groups (Banasiak et al., 2013). Parasilaus is the sole representative of Komarovieae in our analysis, but in its entirety, the tribe comprises a total of seven genera (sensu Downie et al. 2010). Therefore, we highlight this as an area of focus for future studies and refrain from drawing any firm conclusions until further sampling from the two clades (Komarovieae and Physospermopsis) can be undertaken. The genus Conium, which includes the poison hemlock (C. maculatum) and six other species, is the sole member of the Conium clade (Cordes, 2009; Downie et al., 2010). Its isolated position in Apioidaea is supported by our data, where Conium is sister to the clade uniting Selineae and Coriandrae in the coalescent analysis and the Arracacia clade of Selineae in the concatenation analysis.

The circumscription of tribe Scandiceae is also affected by the results presented herein. Downie et al. (2010) recognized six major lineages as comprising the tribe. The current data set places the Acrorema clade as the sister group to the clade comprising...
four Scandiceae subtribes, Arctediinae, Daucinae, Ferulinae, and Torilidinae. The remaining two members of this tribe (sensu Downie et al., 2010) include subtribe Scandiceae and the Glaukosciadium clade. A close association between the Acrornema clade and Scandiceae was evident in a recent 27-taxon transcriptome-based study by Wen et al. (2020), but their results are less detailed because they sampled only three taxa from Scandiceae. Our results provide strong evidence for the inclusion of the Acrornema clade in tribe Scandiceae, particularly the results yielded by the concatenation analysis (Appendix S4, BP = 100). Tordylieae is not monophyletic in the concatenation analysis due to the placement of the Lefebvreia clade (a monophyletic Tordylieae clade recognized at the subtribal rank), which is sister to the Opopanax clade (BP = 100) rather than the rest of Tordylieae. However, Tordylieae is monophyletic in the coalescent analysis (LPP = 0.90, with quartet scores indicating high gene tree incongruence) which is a similar result to previous studies (e.g., Banasiak et al., 2013). Therefore, due to the incongruence between our two analyses (see Fig. 5, tanglegram) these relationships should be further investigated before firm conclusions can be drawn. The Cachrys clade (represented by Alococarpum, Cachrys, Diplotaenina, Ferulago) is embedded within Selineae (LPP = 0.83, with equivocal quartet support) in the coalescent analysis. However, in the concatenation analysis the Cachrys clade is sister to the Selineae clade with high support (BP = 100), a result similar to that of Ajani et al. (2008) based on ITS sequence data, where the Cachrys clade and Selineae (plus Coriandraceae) comprise monophyletic sister clades. The Arracacia clade of Selineae (Ottoa, Myrrhidendron, Tauschia, Neosclerisa, Enantiophyllum) is resolved as monophyletic (LPP = 1, BP = 100) and placed in a large unresolved clade with many other major clades resulting in a paraphyletic Selineae in the concatenation analysis. However, this relationship is not yielded by the coalescent analysis as the Arracacia clade falls within Selineae (LPP = 0.99), similar to the placement in previous studies (e.g., Downie et al., 2010; Banasiak et al., 2013). Therefore, due to the incongruence between our two analyses (see Fig 5, tanglegram) these relationships should be further investigated before firm conclusions can be drawn. We note that Caucalis platycarpus, Diplolophium africamum, and Ormosciadium aucheri are placed differently than the positions yielded in previous studies (Logacheva et al., 2010; Downie and Katz-Downie, 1999), and therefore, we refrain from drawing any firm conclusions regarding these taxa until more dense sampling can be undertaken. Within Selineae, the Perennial Endemic North American (PENA) clade (Podistera, Harbouria, Orogenia, Oreonana, Thaspium, Zizia) is not monophyletic, although branch support in this portion of the tree is poor due to low gene recovery in the target sequence capture (see Appendix S3). The PENA clade forms a well-supported monophyletic group in previous phylogenetic analyses (Downie et al., 2002; Sun et al., 2004) and is widely regarded as a natural group supported by a number of morphological synapomorphies (Downie et al., 2002; Sun and Downie, 2010). Therefore, this unexpected result requires further investigation with additional taxon sampling for the clade. Placing new genera in the tree of life

We specifically targeted nine genera that do not appear in any published molecular phylogenetic studies. These have now been placed in the tree of life for the first time. Many are small or monotypic genera and some are from inaccessible areas (e.g., Afghanistan, Iran, Iraq). It was therefore critical to use natural history collections (mostly the Herbarium at the Royal Botanic Gardens, Kew) as a source of samples for the target sequence capture. Most are poorly understood genera with very few traces of their existence in the published literature. All nine genera were described long before the advent of DNA sequencing, and the oldest was described almost 200 years ago (Psychotis; Koch, 1824).

**Apodicarpum**—The monotypic genus *Apodicarpum* Makino (1891) is endemic to Japan, and its sole species is *Apodicarpum ikenosii* Makino. Hardway et al. (2004), citing morphological similarities of fruit and vegetative characters, suggested that *Apodicarpum* belongs to tribe Oenantheae. Spalik et al. (2009) concurred and provided several morphological characteristics supporting this placement, including roots that are thickened similarly to those of the East Asian species in the genus *Sium*, suggesting that *Apodicarpum* likely has affinities to this taxon. Our data confirm these earlier hypotheses and place *Apodicarpum* in tribe Oenantheae (strongly supported as monophyletic; LPP = 1, BP = 100) where it is sister to *Sium* (LPP = 0.84, BP = 100). Future studies, with more dense sampling of *Sium*, may prove that *Apodicarpum* should be included in *Sium*.

**Bonannia**—*Bonannia* Gussone (1843) is monotypic and native to southern Europe (Greece, Italy and Sicily). *Bonannia* has previously been included in *Cicuta* (Oenantheae), *Foeniculum* (Apiaceae), and *Sium* (Oenantheae), and therefore, its affinities are unclear (POWO 2020). Our data place *Bonannia* resinsiera as sister to the *Opopanax* clade (Ajani et al., 2008; Downie et al., 2010) with LPP = 1 /BP = 100, where it is sister to the clade of *Petroedmondia* (SW Asia), *Magydaris* (Europe and N. Africa), and *Opopanax* (Europe and Asia). On the basis of branch support, we include *Bonannia* within the *Opopanax* clade.

**Grafia**—The genus *Grafia* Reichenbach (1837) is monotypic and native to the former Yugoslavia, Albania, and Italy [containing only *Grafia golaka* (Hacq.) Rchb.]. At various times, *Grafia* was treated under *Aethamanta*, *Hladnikia*, *Malabalia*, and *Pleurospermum* (POWO, 2020), and therefore, its affinities have remained unclear. In the most recent treatment of the family, Plunkett et al. (2018) indicated that *Grafia* has sometimes been included in *Pleurospermum*. We were unable to sample *Pleurospermum*, but among the taxa sampled, our data place *Grafia* in tribe Pleurosermeae (sister to *Physospermum*, with high support LPP = 1, BP = 100). Spalik et al. (2004) placed *Grafia* in tribe Selineae, based on the inclusion of this genus in the *Angelica* clade of Downie et al. (2001) using molecular evidence. However, it seems that the “*Grafia*” sample used for this previous study was likely misidentified material of *Cnidium silafo- lium* (S. R. Downie, University of Illinois, personal communication), which is why Downie et al. (2010) did not include *Grafia* in their summary classification of Apioidae.

**Haplosciadium**—The genus *Haplosciadium* Hochst (1844) is monotypic, containing only the African species *H. abyssinicum* Hochst. *Haplosciadium* is used as an ethno-veterinary medicine in Ethiopia and was traditionally fed to livestock and painted onto livestock due to its reputed medicinal properties (Yineger et al., 2007). Previously, Pimenov and Leonov (1993) placed *Haplosciadium* in tribe Apiaceae based on morphology, but our data place it in the *Lefebvreia* clade of tribe Tordylieae (Winter et al., 2008; Downie et al., 2010) with
Pyramidoptereae. — The genus Pyramidoptereae Boiss. is monotypic, containing only Pyramidopterum minutum (L’Urb.) Briq., which is native to the eastern Aegean Islands (Greece) and Turkey. It is a rare taxon, known from only a small number of localities and is currently threatened with extinction, requiring urgent conservation action (Akalin et al., 2011). Celestani (1905) placed Pyramidopterum in tribe Scandiceae, whereas Plunkett et al. (2018) hypothesized its affinity to either Bunium (Pyramidoptereae) or Carum (Caraeae). Our results show that Pyramidopterum is sister to a clade that unites Astomaea (Turkey and Egypt) and Bunium (widespread, Europe, Asia, and NW Africa) (LPP = 1, BP = 100), placing it in tribe Pyramidoptereae.

Physotrichia — The genus Physotrichia Hiern (1873) contains six species (P. atropurpurea (C. Norman) Cannon, P. heracleoides H. Wolff, P. longiradiatum H. Wolff, P. muriculata (Welw. ex Hiern) S. Droop & C. C. Towns., P. verdickii C. Norman and P. welwitschii Hiern), which are all native to central Africa. Some of these species have previously been included in Daucus (Scandiceae subtribe Daucinae) and Peucedanum (Selinaceae), and therefore, their affinities are unclear (POWO, 2020). The present data set places Physotrichia (represented by P. muriculata) in tribe Echinophoreae with high support (LPP = 1, BP = 100) as sister to Pseudoselinum. Physotrichia is morphologically similar to another African genus, Diplolophium (Plunkett et al., 2018). Due to the unexpected placement of Physotrichia in tribe Echinophoreae, and because this genus is poorly known, we suggest that the tribal placement of Physotrichia should be confirmed in future studies.

Ptychotis — Ptychotis Koch (1824) is monotypic, containing only Ptychotis saxifraga (L.) Loret & Barrandon, which is native to Europe (Corsica, France, Italy, Sardinia, Spain, and Switzerland). Plunkett et al. (2018) reported that the genus remains understudied and that it is unclear whether one or two species should be recognized. Ptychotis was hypothesized to be related to Ammoides in tribe Pyramidoptereae (Koch, 1824; Plunkett et al., 2018), but this relationship is not supported by our data set, which places Ptychotis in tribe Careae with high support (LPP = 1, BP = 100).

Tricholaser — The genus Tricholaser Gilli (1959) comprises two species, T. cachemiricum (C.B. Clarke) Alava and T. ovatilobum (Dunn & R.S. Williams) Alava, which are native to the Western Himalayas of Afghanistan, Pakistan and northwestern India (Plunkett et al., 2018). Our data place Tricholaser (here represented by T. cachemiricum) in Tordylieae subtribe Tordylinae (Downie et al., 2010), with LPP = 1 / BP = 100, as sister to the Asian genus Zosima. This result confirms the taxonomy of Pimenov and Leonov (1993), which also placed Tricholaser in tribe Tordylieae.

Xatartia — The genus Xatartia Meisnser (1838) is monotypic, including only Xatartia scabra (Lapeyr.) Meisn., which survives at high elevations in the eastern Pyrenees (Spain and France). The present data set places Xatartia in tribe Selineae (with LPP = 1, BP = 100), confirming the classification of Pimenov and Leonov (1993), which also placed Xatartia in tribe Selineae (their tribe Angeliceae).

LPP = 1 and BP = 100, where it is sister to another African genus, Afriligusticum.

CONCLUSIONS

Our study has effectively shed light on the higher-level relationships in Apiaceae and highlights the benefits of novel phylogenomic methods, such as targeted sequence capture with the universal Angiosperms353 probe set. These data form a significant contribution to the genus-level phylogeny currently being constructed for all flowering plants utilizing Angiosperms353 (Baker et al., 2021). This method has led to the phylogenetic placement of nine apioid genera that have never previously been included in molecular phylogenetic trees, an important advance toward a comprehensive understanding of the family. It has also allowed us to test the prevailing classification across the family, providing support for the existing subfamilies (and lower groups), while also shedding new light on long-standing concerns regarding the delimitation of Saniculoideae and the placement of three isolated lineages, the Platysace clade, Hermas, and Klotzschia, which may require their own subfamilies.

We have provided further evidence for the successful application of targeted sequence capture to herbarium material (Brewer et al., 2019), paving the way for completing the sampling of all angiosperm genera and, in future, all species. Species-level phylogenies will be required to address other systematic challenges in Apiaceae, such as the suspected polyploidy of many genera and the preponderance of small (1–5 species) genera (Apiaceae contains twice the number of species as its close relative Araliaceae, but 10 times the number of genera). Species level phylogenies also create expanded opportunities for broader comparative analyses so that a deeper understanding of the diversification of this important family may be achieved.

ACKNOWLEDGMENTS

We gratefully acknowledge the following people for their support of this project: Robyn Cowan and Dion Devey for assisting with laboratory work; Nicola Biggs, Sara Edwards, Federico Fabriani, Sue Frisby and Anna Haigh for sampling assistance in the Kew Herbarium; Imalka Kahandawala and Rosy Woods for assisting with DNA bank sampling; Paul Kersey, Ilia Leitch, Kevin Leempeol, Shyamali Roy, Elaine Francoos, Wolf Eiserhardt, Lisa Pokorny, Steven Dodsworth, Joe Parker, and Jan Kim for their contributions to the establishment and running of the PAFTOL project. We thank the two anonymous reviewers for identifying significant improvements for the manuscript. J.C. thanks all those who encouraged and inspired this work through discussions at the Botany 2019 meeting in Arizona, particularly Michael Ottenlips, Jun Wen, Hayley Hale, and Matt Johnson.

AUTHOR CONTRIBUTIONS

J.C., O.M., F.E., and W.J.B. conceived the study. S.R.D., G.M.P., A.N., J.F.S., and M.A.F. all contributed specialist knowledge on the Apiaceae family. O.M. and S.Z. oversaw the specimen sampling at Kew. J.F.S. and G.M.P. provided samples. J.C., K.G., and P.M. conducted all laboratory work with support from G.E.B. and N.E.A.R.Z. carried out the analyses with support from P.B. J.C. wrote the first draft of the manuscript with contributions from A.R.Z. All authors provided feedback on the manuscript. F.E. and W.J.B. secured funding and supervised the PAFTOL project.
FUNDING
This work was supported by grants from the Calleva Foundation and the Sackler Trust to the Plant and Fungal Trees of Life Project (PAFTOL) at the Royal Botanic Gardens, Kew. The internship of Karime Gutierrez was supported by the Muriel Kohn Pokross Travel Internship Fund from Smith College (Massachusetts).

DATA AVAILABILITY
All sequences generated for this study are available in the European Nucleotide Archive (ENA) in a custom project area at https://www.ebi.ac.uk/ena/browser/view/PRJEB35285 (see Appendix S2 for accession numbers). The data set is also available via the Kew Tree of Life Explorer (https://treeoflife.kew.org/) using the FTP tab (http://sftp.kew.org/pub/paftol/current_release/). Taxa that had been sequenced previously were mined from the ENA and NCBI (National Center for Biotechnology Information) databases. Recovery statistics were generated using custom scripts available at https://github.com/sidonieB/.

SUPPORTING INFORMATION
Additional Supporting Information may be found online in the supporting information tab for this article.

APPENDIX S1. Voucher specimens associated with sampling. Collection details are listed for all samples studied and binomial authorities follow POWO (2020).

APPENDIX S2. Recovery statistics table.

APPENDIX S3. Heatmap of gene recovery efficiency plotted with coalescent tree.

APPENDIX S4. Concatenation tree showing relationships in Apiaceae using nuclear exons and inferred in IQ-TREE.

LITERATURE CITED
Cordes, J. M. 2009. A systematic study of poison hemlock (Conium, Apiaceae). M.S. thesis, University of Illinois at Urbana-Champaign, IL, USA.


Gussone, G. 1843. Florae siculae Synopsis 1. Ex Typis Tramater, Naples, Italy.


Hoehst. 1844. Haplosciadium.


