Another first for the Apiaceae: evidence for mitochondrial DNA transfer into the plastid genome

Stephen R. Downie¹,*, Rhiannon M. Peery¹, and Robert K. Jansen²,³

¹ Department of Plant Biology, University of Illinois at Urbana-Champaign, Illinois 61801, U.S.A.
² Department of Integrative Biology, The University of Texas at Austin, Texas 78712, U.S.A.
³ Department of Biological Science, Faculty of Science, King Abdulaziz University, Jeddah 21589, Saudi Arabia

Abstract: The complete plastid genome (plastome) sequences of Anthriscus cerefolium, Crithmum maritimum, Hydrocotyle verticillata, Petroselinum crispum, and Tiedemannia filiformis subsp. greenmanii have recently been determined. With the exceptions of Crithmum and Petroselinum, which each demonstrate major shifts of their LSC-IRₐ (JLB) and LSC-IRₐ (JLA) junctions, all plastomes are typical of most other non-monocot angiosperm plastid DNAs in their structure, organization, and gene content. Crithmum and Petroselinum also incorporate novel DNA in the LSC region adjacent to JLA. These insertions show no sequence similarity to any other region of their plastid genomes and BLAST searches of the Petroselinum insert resulted in multiple significant hits to angiosperm mitochondrial genome sequences. We highlight results of this recent work on the comparative analysis of whole plastid genomes from the Apiales, present the circular plastome gene map of Petroselinum crispum (parsley), review the literature indicating other instances of mitochondrial DNA transfer into Apiaceae plastomes as evidenced by complete mitochondrial DNA sequencing of Daucus carota, and describe our ongoing research on the elucidation of mechanisms creating the many large IR junction shifts characteristic of the family.

Key words: Intracellular DNA transfer, plastome, mitochondrial DNA, inverted repeat

Introduction

The fascinating family Apiaceae has claims to many “firsts.” Because of their numerous distinctive attributes, they were “the first family of
flowering plants to achieve general recognition,” and with the publication of Robert Morison’s monograph *Plantarum Umbelliferarum Distributio Nova* in 1672, the group was subjected to “the earliest systematic study of any group of plants” (Morison, 1672; Constance, 1971). *Coriandrum* (coriander) was “one of the earliest words recognized” in the deciphering of Linear B, a syllabic script used by the Mycenaean Greeks during the fifteen to seventeenth centuries B.C.E. (Constance, 1971; Reduron, 1989), a testament to family’s characteristic fruits, odors, and flavorings. One of “the first historically well-documented cases of successful poisoning” was the drinking of an infusion of *Conium maculatum* (poison hemlock) by Socrates (Cutler, 1992).

The family Apiaceae can also be credited with “the first international symposium dedicated to systematic research on a plant family,” thereby providing a model for similar symposia on other plant families in the years that followed (Heywood, 1971; Watson et al., 2001). Additional “firsts” for the family include several aspects of plastid genome (plastome) structure, including what was at one time considered to be unprecedented variation in the frequency and size of its inverted repeat (IR) junction shifts (Plunkett and Downie, 2000). Most recently, the family provided “the first evidence for mitochondrial DNA transfer into the plastid genome” (Goremykin et al., 2009; Peery et al., 2011; Iorizzo et al., 2012a,b), a truly extraordinary discovery in angiosperm plastome evolution. In this paper, we highlight results of recent work on the comparative analysis of whole plastid genomes from the Apiales (Downie and Jansen, 2015), present the circular plastome gene map of the medicinally important species *Petroselinum crispum* (parsley), review the literature describing other instances of mitochondrial DNA transfer into the Apiaceae plastome, and describe our ongoing research on the group. Apiaceae are a family of tremendous economic, ecologic, and medicinal importance and provide a model system in which to better understand angiosperm plastome evolution.

**Materials and methods**

The complete plastome sequences of *Anthriscus cerefolium* (L.) Hoffm. (chervil), *Crithmum maritimum* L. (sea samphire), *Hydrocotyle verticillata* Thunb. (whorled marshpennywort), *Petroselinum crispum*...
(Mill.) Fuss (parsley), and *Tiedemannia filiformis* (Walter) Feist & S.R. Downie subsp. *greenmanii* (Mathias & Constance) Feist & S.R. Downie (giant water cowbane) have previously been determined (Downie and Jansen, 2015). Methodological details of plastid genome isolation, rolling circle amplification of the entire plastome, library preparation, Roche/454 sequencing, and genome assembly and annotation are presented in this earlier paper. The circular genome map of *Petroselinum crispum* was constructed using OGDraw (Lohse et al., 2007). These plastome data were compared to published data for *Daucus carota* subsp. *sativus* (Apiaceae; GenBank accession number DQ898156; Ruhlman et al., 2006) and *Panax schin-seng* (Araliaceae; AY582139; Kim and Lee, 2004), as well as to results reported previously by R. Peery and colleagues (Peery et al., 2006, 2007, 2011). BLAST searches were conducted against the GenBank nucleotide database to identify regions of similarity between the novel inserts and plant mitochondrial DNA sequences.

**Results**

A summary of the major structural features of each of the five Apiales plastomes is presented in Table 1. These plastomes ranged in size from 152,890 bp (*Petroselinum*; Fig. 1) to 158,355 bp (*Crithmum*), with size extremes caused by major shifts of their LSC-IR<sub>B</sub> (J<sub>LB</sub>) and LSC-IR<sub>A</sub> (J<sub>LA</sub>) junctions relative to the ancestral angiosperm plastid genome structural organization, as typically represented by *Nicotiana tabacum* L. (tobacco). All plastomes contain the same 113 uniquely occurring tRNA, rRNA, and protein-coding genes, with different numbers of genes duplicated by the IR. In *Petroselinum*, the IR has contracted 1,550 bp such that all of *rps19* and most of *rpl2* are now single-copy; in *Crithmum*, the IR has expanded 1,500 bp so that all of *rps19* and previously single-copy genes *rpl22* and *rps3* are now contained within the IR (Fig. 2). Additionally, in *Petroselinum* and *Crithmum*, novel noncoding DNA of 345 and 1,463 bp has been incorporated into the LSC region adjacent to the J<sub>LA</sub> boundary (Fig. 3). In the other Apiales plastomes sequenced, as well as in *Daucus* and *Panax* and the many other non-monocot angiosperm plastomes sequenced to date whose genomes are unrearranged relative to the ancestral plastome gene order, there is commonly between 0 and 30 bp of noncoding sequence in this region (Downie and Jansen, 2015). The large insertions detected
Another first for the Apiaceae: evidence for mitochondrial DNA transfer into the plastid genome

in *Petroselinum* and *Crithmum* show no sequence similarity to any other regions of their respective plastomes, nor do they match any plastid DNA sequence currently available in GenBank. BLAST searches querying the 345-bp *Petroselinum* insert resulted in multiple hits of a 122-bp region to angiosperm mitochondrial DNA sequences from a variety of angiosperms, with the best alignment score showing a 91% sequence similarity to a portion of the cytochrome b (*cob*) - *atp4* (*ORF25*) intergenic spacer region in *Daucus carota* (Fig. 4). BLAST searches of the *Crithmum* insert resulted in no significant hits, except for a small (39-bp) region that also matched a portion of the *D. carota* mitochondrial genome fragment (Fig. 4). Junctions *J*, and *J*, are in the same relative gene positions in all Apiales plastomes, that is within *ycf1* (*J*) and between genes *ndhF* and *ycf1* pseudogene (*J*; Fig. 1).

**Table 1.** Plastome characteristics for five Apiales taxa.

<table>
<thead>
<tr>
<th>Plastome characteristic</th>
<th><em>Hydrocotyle verticillata</em></th>
<th><em>Anthriscus cerefolium</em></th>
<th><em>Tiedemannia filiformis subsp. greenmanii</em></th>
<th><em>Petroselinum crispum</em></th>
<th><em>Crithmum maritimum</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Genbank accession number</td>
<td>HM596070</td>
<td>GU456628</td>
<td>HM596071</td>
<td>HM596073</td>
<td>HM596072</td>
</tr>
<tr>
<td>Plastome size (bp)</td>
<td>153,207</td>
<td>154,719</td>
<td>154,737</td>
<td>152,890</td>
<td>158,355</td>
</tr>
<tr>
<td>LSC size (bp)</td>
<td>84,352</td>
<td>84,774</td>
<td>84,585</td>
<td>86,116</td>
<td>85,230</td>
</tr>
<tr>
<td>SSC size (bp)</td>
<td>18,739</td>
<td>17,551</td>
<td>17,140</td>
<td>17,508</td>
<td>17,139</td>
</tr>
<tr>
<td>IR size (bp)</td>
<td>25,058</td>
<td>26,197</td>
<td>26,506</td>
<td>24,633</td>
<td>27,993</td>
</tr>
<tr>
<td>Number of different genes</td>
<td>113</td>
<td>113</td>
<td>113</td>
<td>113</td>
<td>113</td>
</tr>
<tr>
<td>Number of different genes duplicated by IR</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>16</td>
<td>20</td>
</tr>
<tr>
<td>Length of noncoding region between <em>J</em> and 3’<em>trnH</em>-GUG (bp)</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>345</td>
<td>1,463</td>
</tr>
</tbody>
</table>

*Pseudogenes excluded and one copy of IR removed.*
Figure 1. Circular *Petroselinum crispum* (parsley) plastome gene map, at 152,890 bp in size. The inner circle shows the four major regions of the plastome: the two copies of the inverted repeat (designated as IR$_A$ and IR$_B$), and the large single copy (LSC) and small single copy (SSC) regions. The outer circle shows the gene map with the transcribed regions shown as boxes proportional to their size. Features of the map and gene products are provided in the legend, with the color of the gene boxes indicating the functional group to which the gene belongs. Genes inside the circle are transcribed in a clockwise direction; genes outside the circle are transcribed counterclockwise. Intron-containing genes are indicated by asterisks.
Another first for the Apiaceae: evidence for mitochondrial DNA transfer into the plastid genome

Figure 2. Variation in position of the LSC-IR$_B$ junction (J$_{LB}$) in seven species of Apiales. A. Expansion (Crithmum) and contraction (Petroselinum) of the IR at J$_{LB}$ (thin line = LSC region; thick line with arrow = IR$_B$ region). The vertical line represents the position of J$_{LB}$ in the five other species of Apiales examined, as well as in Nicotiana tabacum (tobacco) and many other non-monocot angiosperm plastomes sequenced to date. Gene map is
based on *Daucus carota* (Ruhlman et al. 2006); coding regions are shaded, introns are not. B. The location of $J_{LB}$ with regard to the gene map. In *Crithmum*, $J_{LB}$ occurs between genes *rpl16* and *rps3*, whereas in *Petroselinum*, $J_{LB}$ occurs in the *rpl2* 5′exon. In all other Apiales examined, $J_{LB}$ occurs at different positions within *rps19*.

**Figure 3.** Gene maps of seven species of Apiales in the vicinity of $J_{LA}$, showing the presence of large regions of novel DNA, of 345 and 1,463 bp, in *Petroselinum* and *Crithmum* plastomes, respectively. Asterisks denote pseudogenes in IR$_A$.

**Figure 4.** Results of BLAST searches querying the (A) 345-bp *Petroselinum* and (B) 1,463 *Crithmum* plastid DNA (ptDNA) inserts between $J_{LA}$ and 3′*trnH*-GUG. A. The best alignment score, with 35% query coverage (*Petroselinum* positions 224-345), showed a 91% sequence similarity to a *Daucus carota* intergenic spacer region between mitochondrial genes cytochrome b (*cob*) and *atp4* (AY007821, Bach et al., 2002; JQ248574, Iorizzo et al., 2012b). B. Only a portion of the *Crithmum* insert (positions 45-83) showed significant (88%) similarity to the same *D. carota* mitochondrial DNA (mtDNA) spacer region.
Discussion

The typical angiosperm plastome is highly conserved in structure, organization, and gene content. Most angiosperm plastomes range between 120 and 170 kilobasepairs (kb) in size, with an IR between 20 and 30 kb and usually about 25 kb (Wicke et al., 2011; Ruhlman and Jansen, 2014). Variation in size of the plastome is due most typically to the expansion and contraction of the IR and changes in sequence complexity due to gene and intron loss and additional duplications outside of the IR (Jansen and Ruhlman, 2012). With the exceptions of *Crithmum* and *Petroselinum*, all examined Apiales plastomes are typical of most other non-monocot angiosperm plastomes.

A plastid DNA (ptDNA) restriction site mapping study revealed that the frequency and large size of \( J_{LB} \) junction shifts in Apiaceae was, at the time this work was done, unprecedented among angiosperms (Plunkett and Downie, 2000). In that study, one expansion and seven different major contractions of the IR relative to the tobacco IR at \( J_{LB} \) were detected, each ranging in size from \( \sim 1 \) to 16 kb. The expansion of the IR in *Crithmum* (the “*Aegopodium* group” of umbellifers or tribe Pyramidoptereae; Downie et al., 2000), resulting in a shift of \( J_{LB} \) of \( \sim 1.1 \) kb, and the contraction of the IR in *Petroselinum* (the “*Apium* group” or tribe Apieae; Downie et al., 2001), resulting in a shift of \( J_{LB} \) of \( \sim 1.6 \) kb, are consistent with the results reported herein. *Coriandrum sativum* L. (coriander; tribe Coriandreae) possesses a \( J_{LB} \) some 16.1 kb away from that of tobacco and most Apiaceae (Plunkett and Downie, 2000), confirming a previous observation by Palmer (1985) that its IR has shrunk to no more than half the normal size. *Coriandrum* also has a 5.7 kb insertion of unknown composition near the 16S rRNA gene at \( J_{LB} \) (Plunkett and Downie, 2000), a region subsequently identified through whole plastome sequencing as a duplication of genes \( trnH \)-GUG and \( psbA \) (Peery et al., 2006 and unpubl.). Apparently, in coriander, the IR contracted initially to include only the rRNA gene cluster and then subsequently expanded to include formerly single-copy genes \( trnH \) and \( psbA \) (Peery et al., 2006 and unpubl.). The mechanisms creating these large-scale expansions and contractions of the IR were only speculated upon and included intramolecular recombination between small, dispersed repeats and tRNA genes (Plunkett and Downie, 2000).
In addition to the major IR junction shifts characteristic of Apiaceae, whole plastome sequencing has revealed unique, noncoding DNA in the region bounded by \( J_{LA} \) and \( trnH \)-GUG of the LSC region that does not show significant similarity to any other known cpDNA sequence. Such novel insertions in this region (of 214-392 bp) have been reported in other Apiaceae, including all examined members of tribe Apieae (Peery et al., 2007, 2011 and unpubl.). BLAST searches of these sequences resulted in multiple significant hits to angiosperm mitochondrial sequences, with the highest values registered for \( Daucus carota \) (Bach et al., 2002; Iorizzo et al., 2012b). We have suggested that these novel insertions likely represent a unique transfer of mitochondrial DNA into the plastome (Peery et al., 2011; Downie and Jansen, 2015).

Goremykin et al. (2009), while analyzing the \( Vitis vinifera \) L. (grape) mitochondrial genome, provided the first evidence of mitochondrial DNA transfer into an angiosperm plastome. Sequences having 95% similarity to the mitochondrial gene cytochrome c oxidase subunit 1 (\( cox1 \)) were found integrated into the plastid genome of cultivated \( Daucus carota \) (Ruhlman et al., 2006) in a unique stretch of noncoding DNA between IR genes 3’\( rps12 \) and \( trnV \)-GAC. This region, identified as being 1,439 bp in size by Goremykin et al. (2009), does not occur in any other Apiales plastome we have examined including \( Anthriscus \) of the same tribe Scandiceae, nor was it present in any other plastome available to Goremykin et al. (2009) at the time of their study. Subsequently, Iorizzo et al. (2012a,b), as a result of characterizing the full carrot mitochondrial genome, designated this site as the \( Daucus carota \) mitochondrial-plastid (\( Dc \)MP) region and reported that it is 1,452 bp in size. They discovered that this region is present as three non-contiguous, rearranged sequences in the mitochondrial genome of \( D. carota \) (Iorizzo et al., 2012b). In the plastome, however, the \( Dc \)MP sequence, or a large portion of it, is present only in \( Daucus \) (seven species) and its close relative \( Cuminum \) L. (cumin), both of Scandiceae subtribe Daucinae. Iorizzo et al. (2012a) concluded that their results provided strong evidence of a mitochondrial to plastid transfer of DNA in the ancestor of subtribe Daucinae.

Within flowering plants, the intracellular transfers of DNA from the mitochondrion and plastid into the nucleus and from the plastid and nucleus into the mitochondrion are well documented (Kleine et al., 2009; Iorizzo...
et al., 2012a; Park et al., 2014). Until recently, there was no evidence of transfer from the nucleus or mitochondrion into the angiosperm plastome, giving credence to the phrase “plastids can dish it out but can’t take it” (Smith, 2011). To date, only members of families Apiaceae (Goremykin et al., 2009; Iorizzo et al., 2012a,b) and Apocynaceae (Ku et al., 2013; Straub et al., 2013) have plastids containing foreign DNA determined to be of mitochondrial provenance. A portion of \textit{DcMP} is highly conserved across mitochondrial genomes of several disparate lineages of angiosperms, suggesting that the direction of transfer was from mitochondrion to plastid (Goremykin et al., 2009; Iorizzo et al., 2012b). The high sequence similarity of novel DNA adjacent to \textit{J_{LA}} in \textit{Petroselinum} and other Apiaceae to a variety of angiosperm mitochondrial DNA sequences suggests strongly that these transfers came from the mitochondrion too. While the possibility also exists that this novel DNA in Apiaceae plastomes is of a nuclear origin, we believe this is unlikely. There has been substantial transfer of DNA from the mitochondrion to the nucleus during angiosperm evolution (Kleine et al., 2009), but it is much less parsimonious to suggest that the mitochondrial DNA was transferred to the nucleus and then was subsequently transferred to the plastid. The Apiaceae are the first family of flowering plants to provide convincing evidence of intracellular DNA transfer from the mitochondrion to the angiosperm plastid, a direction of transfer heretofore considered highly unlikely possibly because of the lack of an efficient DNA uptake system within plastids (Richardson and Palmer, 2007; Bock, 2010; Bock and Timmis, 2008; Kleine et al., 2009; Smith, 2011; Wicke et al., 2011; Iorizzo et al., 2012a). It appears that the mitochondrion to plastid transfer of DNA can indeed occur and that angiosperm plastomes are not as impenetrable to foreign DNA as previously thought (Smith, 2014).

To date, fragments of the mitochondrial \textit{cox1} gene and \textit{cob-atp4} intergenic spacer region have been integrated into the plastomes of some Apiaceae. Such rare genomic changes are additionally significant for they can provide complimentary markers to DNA sequencing for resolving relationships and designating major clades (Downie and Palmer, 1992; Rokas and Holland, 2000). The \textit{DcMP} insertion circumscribes Scandiceae subtribe Daucinae and would likely have been transferred from the mitochondrion to the plastid in the common ancestor of the group (Iorizzo
et al., 2012a). The putative mitochondrial insertion at $J_{\text{LA}}$ characterizes *Petroselinum* and all other members of tribe Apieae (Peery et al., 2007), and specific IR junction shifts have established sister group relationships, such as that between tribes Careae and Pyramidoptereae (Plunkett and Downie, 2000; Downie and Jansen, 2015). The distribution of rare genomic changes has much potential to demarcate major clades in the family, particularly those within the apioid superclade where higher-level relationships remain elusive.

We have also sequenced whole plastomes from *Anethum graveolens* L. (dill), *Carum carvi* L. (caraway), *Coriandrum sativum*, and *Foeniculum vulgare* Mill. (fennel) and examined LSC-IR junctions in a phylogenetically diverse selection of other Apiaceae (R. Peery et al., unpubl). Characterization of these plastomes and junction regions against a robust phylogeny for the group is revealing inversions, new IR junction shifts, additional transfers of mitochondrial DNA into the plastome, and when these rare genomic changes may have occurred. The mechanisms leading to IR junction shifts in land plant plastomes are numerous and include such phenomena as gene conversion, double-strand break repair, dispersed sequence repeats, and double reciprocal recombination between IR segments (reviewed in Goulding et al., 1996). Unique to Apiaceae, the presence of mitochondrial DNA sequences in the LSC region adjacent to $J_{\text{LA}}$ may either be the cause or consequence of these IR junction shifts. The mechanisms responsible for these dynamic IR junction shifts in Apiaceae are currently being assessed.

Apiaceae contain many plants used as food and medicine. As examples, carrot is one of the ten most economically important vegetable crops in the world (Simon et al., 2008), parsley is believed to be among the world’s most potent disease-fighting spices (Karimi et al., 2012), and *Crithmum maritimum* and *Foeniculum vulgare* demonstrate strong antioxidant and antibacterial properties (Ruberto et al., 2000). The continued sequencing of Apiaceae plastomes not only adds to our knowledge of angiosperm plastome evolution but the information obtained also is important for studies of plastid genetic engineering, as the advantages of plastid over nuclear transformations are numerous (Ruhlman et al., 2006; Ruhlman and Jansen, 2014).
Another first for the Apiaceae: evidence for mitochondrial DNA transfer into the plastid genome

Acknowledgements

The authors thank the Apiales Istanbul 2014 Organizing Committee for arranging this 8th international multidisciplinary symposium on the Apiales and these resultant proceedings, and Deborah Katz-Downie for comments on the manuscript.

References


