

A *trnI*_CAU Triplication Event in the Complete Chloroplast Genome of *Paris verticillata* M.Bieb. (Melanthiaceae, Liliales)

Hoang Dang Khoa Do, Jung Sung Kim, and Joo-Hwan Kim*

Department of Life Science, Gachon University, Seongnam, Gyeonggi-do, Korea

*Corresponding author: E-mail: kimjh2009@gachon.ac.kr.

Accepted: June 16, 2014

Data deposition: The chloroplast genome sequence has been deposited at GenBank under the accession KJ433485.

Abstract

The chloroplast is an essential plant organelle responsible for photosynthesis. Gene duplication, relocation, and loss in the chloroplast genome (cpDNA) are useful for exploring the evolution and phylogeny of plant species. In this study, the complete chloroplast genome of *Paris verticillata* was sequenced using the 454 sequencing system and Sanger sequencing method to trace the evolutionary pattern in the tribe Parideae of the family Melanthiaceae (Liliales). The circular double-stranded cpDNA of *P. verticillata* (157,379 bp) consists of two inverted repeat regions each of 28,373 bp, a large single copy of 82,726 bp, and a small single copy of 17,907 bp. Gene content and order are generally similar to the previously reported cpDNA sequences within the order Liliales. However, we found that *trnI*_CAU was triplicated in *P. verticillata*. In addition, *cemA* is suspected to be a pseudogene due to the presence of internal stop codons created by poly(A) insertion and single small CA repeats. Such changes were not found in previously examined cpDNAs of the Melanthiaceae or other families of the Liliales, suggesting that such features are unique to the tribe Parideae of Melanthiaceae. The characteristics of *P. verticillata* cpDNA will provide useful information for uncovering the evolution within *Paris* and for further research of plastid genome evolution and phylogenetic studies in Liliales.

Key words: *trnI*_CAU triplication, *cemA* pseudogenization, chloroplast genome, *Paris verticillata*, Melanthiaceae, Liliales.

Introduction

The chloroplast of plants is believed to have evolved through an endosymbiotic event in which a eukaryotic heterotrophic organism became host to a cyanobacterium, with an interaction between them resulting in chloroplast formation (Douglas 1998; McFadden 1999). It contains a circular double-stranded DNA molecule ranging in length from approximately 100 kb to over 160 kb (Sugiura 1992). The chloroplast genome (cpDNA) has a quadripartite structure, which includes a large single copy (LSC), a small single copy (SSC), and two inverted repeat (IR) regions. The chloroplast genome contains genes that are responsible for photosynthesis (Sugiura 1992) and is inherited maternally, paternally, and even biparentally (Corriveau and Coleman 1988; Dong et al. 1992; Birky 1995; Yang et al. 2000; Zhang et al. 2003; Hansen et al. 2007; McCauley et al. 2007). Generally, the gene content and order are highly conserved in cpDNA across plant species. Therefore, cpDNA protein-coding sequences are believed to provide useful information for exploring phylogenetic relationships among plant species. For example, Jansen et al. (2007)

resolved the relationships among major clades of the angiosperms by analyzing 81 genes from the plastid genomes of 64 species. However, gene loss, relocation, and transformation to pseudogenes have also occurred. In parasitic plants such as those belonging to the genus *Cuscuta*, genes that encode photosynthesis proteins (e.g., *ndh* genes) are lost and have become pseudogenes due to the lifestyle of dependency on a host plant (McNeal et al. 2007). Lee et al. (2007) reported gene relocation in cpDNA of *Jasminum* and *Menodora* (Oleaceae), which was the result of multiple and overlapping inversions. In addition to these mutations, gene duplication events are another characteristic type of phylogenetically informative change in cpDNA. For example, *trnF* sequences were investigated and found to be duplicated in cruciferous plants, although only part of the gene was duplicated (Schmickl et al. 2007). For these reasons, the number of chloroplast genomes of green plants uploaded to NCBI has risen to 471, of which 372 are from angiosperms (<http://www.ncbi.nlm.nih.gov/genomes/GenomesGroup.cgi?taxid=2759&opt=plastid#pageTop>, last accessed May 13, 2014).

© The Author(s) 2014. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com



FIG. 1.—Photographs of *Paris verticillata*. (A) Young plant, (B) plant with flower, and (C) close up of flower.

Such increasing resources will allow exploration of evolution among plants. *Paris verticillata* M.Bieb (fig. 1) is a member of the tribe Parideae of the family Melanthyaceae (Angiosperm Phylogeny Group 2009). The species is widespread in East Asia and has been used as folk medicine for asthma and chronic bronchitis (Ahn 1998). Recently, a new phenolic amide and pyrrolizidine alkaloids extracted from the roots of *P. verticillata* were found to have potential cytotoxicity against human tumor cell lines (Lee et al. 2008; Kim et al. 2010). Although the medicinal features of *P. verticillata* have been studied, no genomic work on the plant has yet been conducted. We therefore analyzed the complete chloroplast genome sequence of *P. verticillata* and compared the cpDNA features with previously reported cpDNA in Liliales (Liu et al. 2012; Bodin et al. 2013; Do et al. 2013; Kim JS and Kim J-H 2013). Furthermore, we examined whether the specific changes in cpDNA of *P. verticillata* are found in other species in Liliales to investigate the evolutionary pattern and to provide useful molecular information for further research on this potential medicinal plant.

Results and Discussion

Genome Assembly, Features, and Comparisons with Other Liliales cpDNAs

The 96,169 reads (range: 40–680 bp) of *P. verticillata* cpDNA generated by the 454 system were assembled against the reference sequence of *Chionographis japonica* (Bodin et al. 2013). Out of them, 327 reads (0.34%) were assembled to *C. japonica*, with an average length of 365 bp and covering 35% of the reference sequences with 97% pairwise identity. Because of the low covering percentage of reads to reference sequence (35%) and low assembly coverage (1.1 \times), PCR and Sanger sequencing methods were conducted to complete the cpDNA sequence and to confirm the regions assembled from

454 system reads (supplementary fig. S1, Supplementary Material online).

The cpDNA sequence of *P. verticillata* (accession number: KF433485) was complete and of length 157,379 bp, in which 82,726 bp encompass the LSC region, 17,907 bp the SSC region, and 28,373 bp the length of each IR region (fig. 2). The AT and GC contents are 62.4% and 37.6%, respectively (table 1). The cpDNA consists of 115 unique genes, which are composed of 81 protein-coding genes, 30 tRNAs, and 4 rRNAs. Among the protein-coding genes, nine genes contain one intron and three genes possess two introns (*ycf3*, *clpP*, and *rps12*) of which *rps12* was *trans*-spliced. In addition, 25 coding regions are duplicated in the IR region. However, only part of *ycf1* was duplicated in the junction between the IRB and SSC regions. Similarly, *rps3* in the junction between the IRA and LSC regions is not functional because of incomplete duplication (6 bp). The *ycf15* and *ycf68* were pseudogenes according to the presence of several internal stop codons within the coding regions.

The basic features within Liliales were compared among reported cpDNA sequences (table 1). The results indicated that the length of *P. verticillata* (157,379 bp) is similar to that of *Smilax china* (157,878 bp) and longer than those of other species (range: 152,793–155,510 bp). The AT and GC contents of Melanthyaceae species were almost the same (62.3% and 37.7%, respectively), whereas these features are different in other families of Liliales (table 1). The gene content and order are similar among the species examined, but *rps16* was deleted completely in *C. japonica* and partially in *Veratrum patulum*. In addition, *infA* was lost in *S. china*. The IR borders were expanded varyingly not only in Melanthyaceae but also in other taxa (table 1). The IR/SSC boundary was identified by the incomplete duplication of *ycf1* in all species examined, whereas the IR/LSC junction was expanded to full *trnH_GUG* (*V. patulum*), part of *rps19* (*Alstroemeria aurea* and

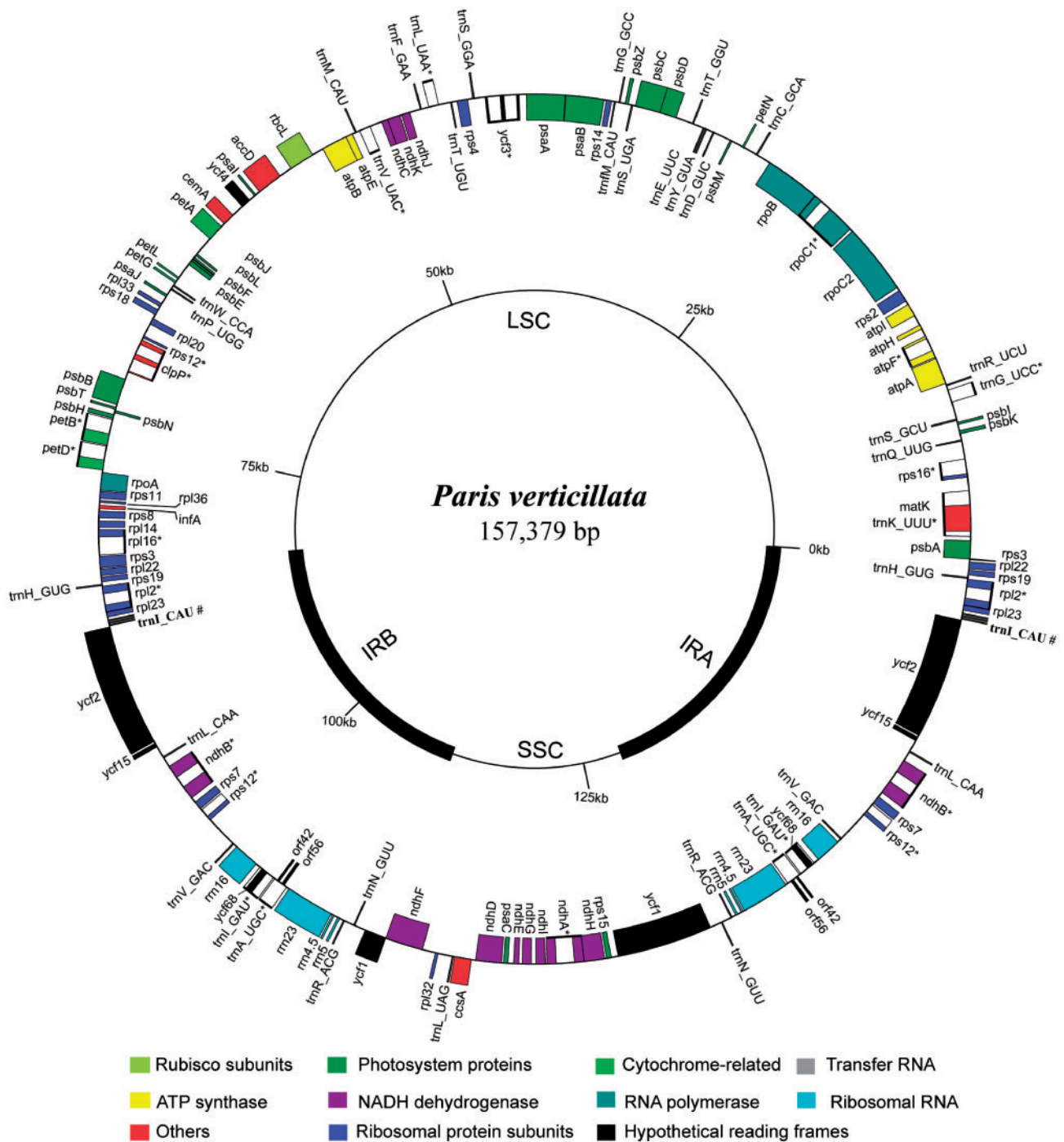


Fig. 2.—Map of the *Paris verticillata* chloroplast genome. Genes shown outside of the outer circle are transcribed counterclockwise, whereas those shown inside are transcribed clockwise. The thick lines in small circles indicate the IR regions. The asterisks indicate those genes with introns. # indicates triplicated genes.

Lilium longiflorum), part of *rpl22* (*S. china*), and part of *rps3* (*C. japonica* and *P. verticillata*). In general, the cpDNA structure of *P. verticillata* is similar to those of other Liliales species, such as *V. patulum* (Do et al. 2013), *C. japonica* (Bodin et al. 2013), *A. aurea*, *L. longiflorum* (Kim JS and Kim J-H 2013), and

S. china (Liu et al. 2012) (table 1). The gene content and order are also similar among the species examined. However, the loss of *infA* and *rps16* as well as partial deletion of *rps16*, which is uncommon in the Liliales, could be unique evolutionary events in such species.

Table 1

Characteristics of Chloroplast Genomes among Liliales

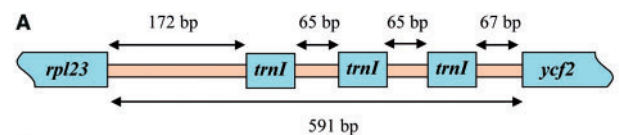
Species (Family)	<i>Paris verticillata</i> (Melanthiaceae)	<i>Chionographis japonica</i> (Melanthiaceae)	<i>Veratrum patulum</i> (Melanthiaceae)	<i>Lilium longiflorum</i> (Liliaceae)	<i>Smilax china</i> (Smilacaceae)	<i>Alstroemeria aurea</i> (Alstroemeriaceae)
Accession number	KJ433485	KF951065	KF437397	KC968977	HM536959	KC968976
Protein-coding genes	81	80	81	81	80	81
tRNAs	30	30	30	30	30	30
rRNAs	4	4	4	4	4	4
Length (bp)	157,379	154,646	153,699	152,793	157,878	155,510
LSC	82,726	81,653	83,372	82,230	84,608	84,241
SSC	17,907	18,195	17,607	17,523	18,536	17,867
IRs	28,373	27,399	26,360	26,520	27,367	26,701
AT content (%)	62.4	62.3	62.3	62.98	62.75	62.74
GC content (%)	37.6	37.7	37.7	37.02	37.25	37.26
IRB–SSC junction	<i>ycf1</i>	<i>ycf1</i>	<i>ycf1</i>	<i>ycf1</i>	<i>ycf1</i>	<i>ycf1</i>
	(pseudogene)	(pseudogene)	(pseudogene)	(pseudogene)	(pseudogene)	(pseudogene)
IRA–LSC junction	<i>rps3</i>	<i>rps3</i>	<i>trnH_GUG</i>	<i>rps19</i>	<i>rpl22</i>	<i>rps19</i>
	(pseudogene)	(pseudogene)		(pseudogene)	(pseudogene)	(pseudogene)
Length of IGS between <i>rpl23</i> and <i>ycf2</i> (bp)	591 bp	303 bp	305 bp	308 bp	308 bp	308 bp

The expansion and contraction of the IR region are highly variable among species, within not only Melanthiaceae but also the order Liliales, in that the junctions were differentiated from *trnH_GUG* to *rps3* (table 1). The IR junction of *P. verticillata* is in agreement with a previous report (Wang et al. 2008), which suggested that the junction of the order Liliales IR/LSC region included the *trnH-rps19* cluster. However, to date, the IR regions have only been examined in four of ten families of Liliales (Bodin et al. 2013; Do et al. 2013). Therefore, further studies covering the junctions of IR and LSC as well as the SSC region in all families of Liliales are required.

trnI_CAU Triplication and *cemA* Pseudogenization

The length of the intergenic spacer (IGS) between *rpl23* and *ycf2*, which contains *trnI_CAU*, varies among *P. verticillata* and other Liliales species (303–591 bp; table 1). *Paris verticillata* possesses the longest IGS, containing three copies of *trnI_CAU* (fig. 3A). Such a pattern was not detected in other complete cpDNAs from Liliales (Liu et al. 2012; Bodin et al. 2013; Do et al. 2013; Kim JS and Kim J-H 2013). Further analysis through REPuter (Kurtz et al. 2001) recognized tandem repeat sequences of 139 bp, which included the 74-bp *trnI_CAU* sequences, in the relevant region of *P. verticillata* (fig. 3B), but not in other genomes (data not shown).

Gene duplication in the chloroplast genome occurs mainly within the IR regions due to the IR expansion (Goulding et al. 1996; Xiong et al. 2009). Most duplicated genes are associated with tRNAs, which were found in the IGS of *rbcl-psaI* of



B
AAATCCATTTCTCCCTATGAGTTCAGTATCGATAAGAATTCTAGTT
CTTACTGTTTCATATGTTATGGTATGAATATACCATATCAATTCGTTATG
TATGGATGCTTAACAGG**AATCATCGTAAATAAAAATAACCAAATTC**
AAGACTTATGAAACGTTCCATGGCGTGCATCCAGCAGGAATTGAACCT
ACGAATTGGCAATTATGAGTTGGGCGCTTAAACATTCAGCCATGGAT
GCTTAAACAGGAATCATCGTAAATAAAAATAACCAAATTC
CAATAGACTTATGAAACGTTCCGTTGGCGTGCATCCAGCAGGAATTGAACCTACGAATT
TGCCAATTATGAGTTGGGCGCTTAAACATTCAGCCATGGATGCTTAAAC
AGGAATCATCGTAAATAAAAATAACCAAATTC
CAATAGACTTATGAAACGTTCCGTTGGCGTGCATCCAGCAGGAATTGAACCTACGAATTG
CAATAGACTTATGAGTTGGGCGCTTAAACATTCAGCCATGGATGCTTAAACAGG
TATCATCGTAAATAAAAATAACCAAATTC
CAATAGACTTATGAAACGTTCCGTTGGCGTGCATCCAGCAGGAATTGAACCTACGAATTG
CAATAGACTTATGAGTTGGGCGCTTAAACATTCAGCCATGGATGCTTAAACAGG
ATCATAAATAACCAAATTCCAATTTAAATGAAATCTTTAGGAGGAAGC

Fig. 3.—Illustration of *trnI_CAU* composition in *Paris verticillata*. (A) Positions of the *trnI_CAU* copies. (B) The nucleotide sequence of the *rpl23-ycf2* IGS of *P. verticillata*, in which three tandem repeat units are highlighted in different colors. The bold italic characters indicate the sequences of *trnI_CAU*.

Jasminum and *Menodora* of the family Oleaceae (Lee et al. 2007) and in *trnC-rpoB* of *Ginkgo biloba* (Lin et al. 2012). In *G. biloba*, *trnC-GCA* was duplicated at least twice and evolved into a cluster of three tRNA genes: *trnY-AUA*, *trnC-ACA*, and *trnSeC-UCA*. Similarly, the duplication of *trnF-GAA* was also detected in Brassicaceae (Schmickl et al. 2007). Although gene duplication events were reported previously, the mechanisms underlying these events remain unclear (Lee et al. 2007; Erixon and Oxelman 2008; Schmickl et al. 2007; Lin et al. 2012). Guo et al. (2007) proposed that tandem repeats



Fig. 4.—The alignment of partial *cemA* sequences among *Paris verticillata* and related taxa. (A) Alignment of partial nucleotide sequences of *cemA*. The poly(A) sequences are underlined. The colored boxes show the SSR. (B) Alignment of partial amino acid sequences of *cemA*. The asterisks indicate stop codons. The letters shaded in color show differences in the amino acid compositions of the *cemA* genes among species. The underlined characters show amino acid sequences that are similar among *P. verticillata* and other species.

found in legume chloroplast genomes were generated by homology-facilitated illegitimate recombination. We suggest that the *trnI*_CAU duplication events may be attributable to homology-facilitated illegitimate recombination, which has occurred twice to generate three tandem repeats in *P. verticillata*.

The *cemA*-coding sequence of *P. verticillata* contained a poly(A) sequence (9 bp) and a small single repeat (SSR) CA unit (fig. 4A). The other species of Liliales have poly(A) sequences of variable lengths (8–12 bp) following the start codon, and *C. japonica* has the longest such sequence (12 bases). However, the CA SSR was not detected in these taxa. Although the lengths of poly(A) sequences are variable, the amino acid sequences are very similar (fig. 4B). In contrast, the poly(A) sequence and the SSR unit caused a frameshift mutation in *cemA* of *P. verticillata*. Consequently, the *cemA* amino acid sequence of *P. verticillata* differs from those of other species, except for the first four amino acids (fig. 4B).

In addition to the pseudogenization of hypothetical open-reading frames (*ycf15*, *ycf68*) in *C. japonica*, *V. patulum*, *A. aurea*, *L. longiflorum*, and *S. china*, dysfunctional protein-coding genes have also been reported in many land plants. Lin et al. (2012) reported the pseudogenization of *rpl23* in *G. biloba* caused by the truncated 5'-region. The loss of functional genes has also been found in parasitic plants. For example, many genes such as *atpB*, *rbcl*, *ndhF*, and *rpoC2* were found to have been lost or to be nonfunctional in *Cistanche deserticola* due to its parasitic lifestyle with its host *Haloxylon ammodendron* (Li et al. 2013). The *cemA* gene, assumed to encode a b-type heme protein of unknown function (Willey and Gray 1990), was reported to have been lost in *C. deserticola*, *Epifagus virginiana*, *Rhizanthella gardneri*, and *Neottia nidus-avis* (Wolfe et al. 1992; Delannoy et al. 2011;

Logacheva et al. 2011; Li et al. 2013). Different lengths of poly(A) sequence have also been observed in other species (Yang et al. 2010). Although transcriptome data have been analyzed, Yang et al. (2010) concluded that whether *cemA* can be translated into protein remains unclear. The *cemA* sequenced in this study is suspected to be pseudogene because of the presence of several stop codons caused by the poly(A) sequence and SSR of CA at the beginning of the coding region in *P. verticillata* (tribe Melanthieae of the Melanthiaceae; fig. 4B). However, it is thought to be functional in other species such as *V. patulum* (tribe Melanthieae of Melanthiaceae), *C. japonica* (tribe Chionographideae of Melanthiaceae), *A. aurea* (Alstroemeraceae), *L. longiflorum* (Liliaceae), and *S. china* (Smilacaceae) because of the absence of internal stop codons. Consequently, this mutation may occur only in the tribe Parideae of Melanthiaceae and could be useful for further research in not only Melanthiaceae but also Liliales species. The loss of *cemA* in parasitic plants could be explained by the dependence on the host plants. However, *P. verticillata* is autotrophic. Therefore, further research is required to clarify the impact of *cemA* pseudogenization in this species.

Phylogenetic Analysis

Phylogenetic relationships between species in Liliales and other monocots, and dicots, were explored (fig. 5). The results showed that Liliales was a monophyletic group; the bootstrap value was high (BP 100). Liliales is a sister group of other species of Aparagales (*Oncidium* Gower Ramsey, *Erycina pusilla*, *Cymbidium aloifolium*, *Phalaenopsis aphrodite* subsp. *formosana*, and *Yucca schidigera*), Poales (*Anomochloa marantoidea*, *Bambusa emeiensis*, *Dendrocalamus latiflorus*,

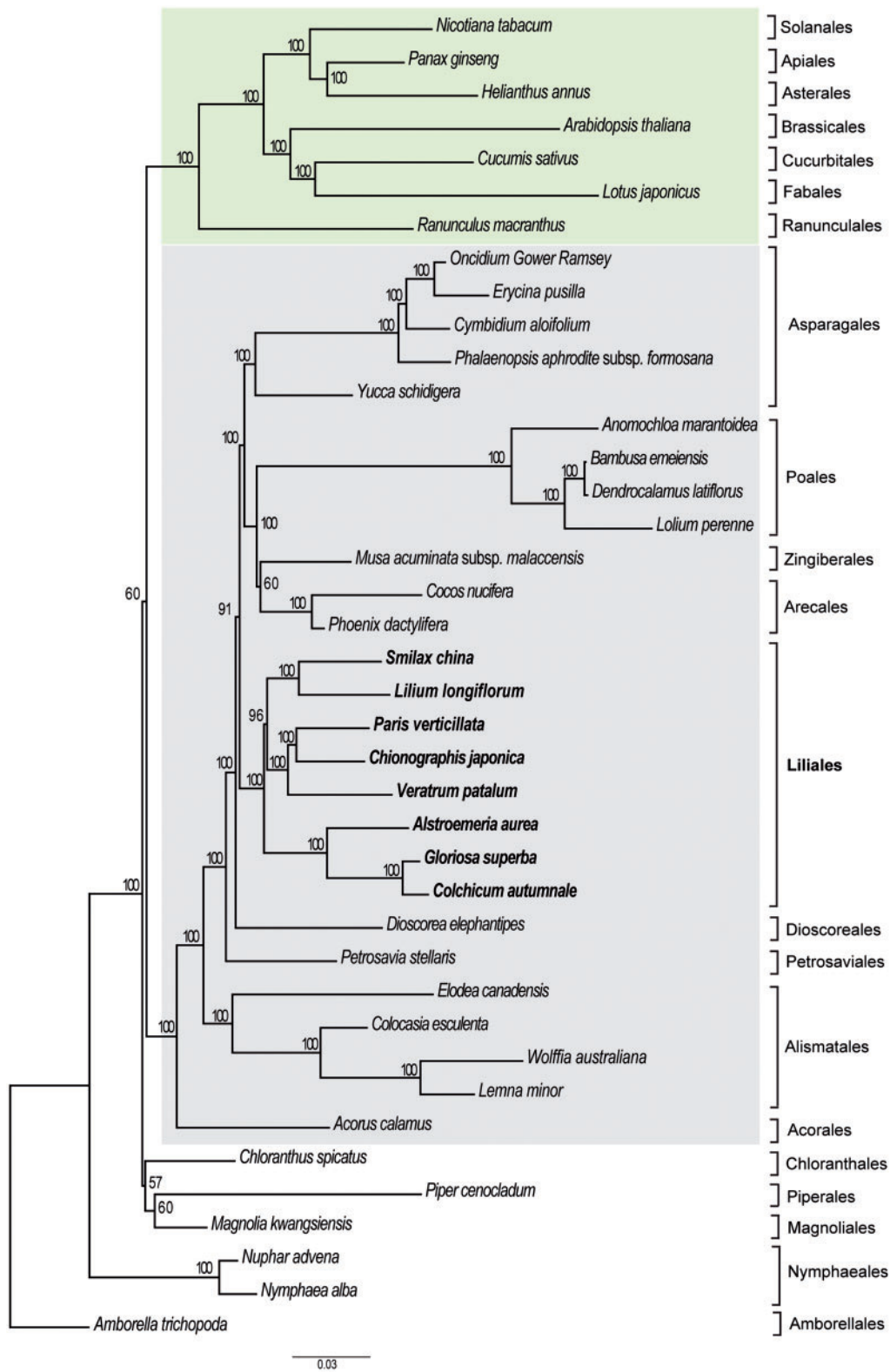


Fig. 5.—Phylogenetic tree inferred by RAxML using nucleotide sequences of 76 protein-encoding regions from 40 species. Bootstrap values (>50) are shown above the branches. The light green color box shows the eudicots group whereas the light gray color box indicates the monocots species. The names in the right side of phylogenetic tree represent the classification of species at order level.

and *Lolium perenne*), Zingiberales (*Musa acuminata* subsp. *malaccensis*), and Arecales (*Cocos nucifera* and *Phoenix dactylifera*) (BP 91). Within Liliales, Melanthiaceae (including *P. verticillata*, *V. patulum*, and *C. japonica*) is sister to the Smilacaceae (*S. china*) and Liliaceae (*L. longiflorum*). Also, the Alstroemeriaceae (*A. aurea*) is sister to the Colchicaceae (*Colchicum autumnale* and *Gloriosa superba*). The familial relationships defined in this study were identical to those delineated in a previous work (Kim et al. 2013), in which relationships among all families in the Liliales were investigated.

Conclusions

Here, we report the first data of *trnL*_CAU triplication and *cemA* pseudogenization in Melanthiaceae inferred from the complete cpDNA sequence of *P. verticillata*. Notably, these features were not noted in the previous studies on cpDNA of either the Melanthiaceae or the Liliales. Therefore, these patterns will be useful for understanding the phylogeny and evolution of these species. However, the detailed mechanisms and evolutionary impacts of these findings remain unclear, and further investigations are therefore required.

Materials and Methods

Taxon Sampling, cpDNA Extraction, Sequencing, and Assembly

Paris verticillata was collected in South Korea, and a voucher specimen was deposited in the herbarium of Gachon University (voucher number: GCU02222). The plant materials used in this study were obtained from the Korean National Research Resource Center (Medicinal Plants Resources Bank NRF-2010-0005790) supported by the Korea Research Foundation with resources provided by the Ministry of Education, Science, and Technology in 2013. Fresh leaves (50 g) of *P. verticillata* were used for chloroplast isolation employing the Percoll gradient buffer method (Kim JS and Kim J-H 2013). A DNeasy Plant Mini Kit (Qiagen, Seoul, South Korea) was used to extract cpDNA from purified chloroplasts. The 454 sequencing system (Roche Applied Science, Penzberg, Germany) was employed to sequence cpDNA of *P. verticillata*. The raw data were uploaded to Geneious Version 6.1, created by Biomatters Ltd. (Auckland, New Zealand), to assemble sequencing data. The *C. japonica* cpDNA sequence (Bodin et al. 2013) was chosen as a reference to identify gaps in the sequence of *P. verticillata*. Prior to assembly, raw sequences that included ambiguous bases ("N") were excluded. The remaining sequences were mapped using the default settings of Geneious. Only reads that were in excess of 90% identical to reference sequence were selected to identify gaps in *P. verticillata* cpDNA. After assembly, identification of gaps and calculation of assembly coverage, specific primers, designed with the aid of Primer3

(Untergrasser et al. 2012), and candidate Liliales primers (Bodin et al. 2013) were used to fill all gaps, to confirm ambiguous sequences including low assembly coverage regions and to identify the borders of the LSC, SSC, and IR regions through PCR and Sanger sequencing method.

Data Analysis

The gene content and order were identified with the aid of Geneious and adjusted manually. The tRNAscan-SE (Schattner et al. 2005) was used to confirm tRNAs. Ambiguous bases in coding regions were checked using data in NCBI (<http://blast.ncbi.nlm.nih.gov/>, last accessed May 13, 2014) and confirmed by Sanger sequencing. In particular, specific primer pairs were used to verify triplication of *trnL*_CAU and pseudogenization of *cemA*, *ycf15*, and *ycf68*. The cpDNA map of *P. verticillata* was constructed using GenomeVx (Connant and Wolfe 2008). The sequences of *cemA* were aligned using MUSCLE (Edgar 2004), which is included in the Geneious program, and manual adjustments were made when necessary. Sequencher version 5.0 (Gene Codes Co., Ann Arbor, MI) was used to assemble complete sequences of *trnL*_CAU and *cemA*. The REPuter program (Kurtz et al. 2001) was used to detect repeat units in the *rpl23-ycf2* regions.

Phylogenetic Analysis

Phylogenetic analysis was performed using data on 76 protein-encoding genes from cpDNA sequences of 40 species (supplementary table S1, Supplementary Material online). The nucleotide sequences were aligned using CLUSTALW (Hall 1999), which is included in the Geneious program. Phylogenetic trees were reconstructed using RAxML (Stamatakis et al. 2008), which is available online (<http://embnet.vital-it.ch/raxml-bb/index.php>, last accessed May 13, 2014). Substitution of GTR+G was modeled for the entire data matrix. Maximum-likelihood bootstrap analysis was calculated using 100 replications employing the rapid bootstrapping approach implemented in RAxML. The phylogenetic tree was drawn using FigTree v1.3 program.

Supplementary Material

Supplementary figure S1 and table S1 are available at *Genome Biology and Evolution* online (<http://www.gbe.oxfordjournals.org/>).

Acknowledgments

The authors thank Sang-Chul Kim of Gachon University, Korea, for collecting the plant materials and anonymous reviewers for suggestions and comments on the article. This work was supported by the National Research Foundation of Korea (NRF) grant fund (MEST 2010-0029131).

Literature Cited

- Ahn DK. 1998. Illustrated book of Korean Medicinal Herbs. Seoul (Korea): Kyo-hak Publishing Co. p. 107.
- Angiosperm Phylogeny Group. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Bot J Linn Soc.* 161:105–121.
- Birky CW Jr. 1995. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proc Natl Acad Sci U S A.* 92: 11331–11338.
- Bodin SS, Kim JS, Kim J-H. 2013. Complete chloroplast genome of *Chionographis japonica* (Willd.) Maxim. (Melanthiaceae): comparative genomics and evolution of universal primers for Liliales. *Plant Mol Bio Rep.* 31(6):1407–1421.
- Conant GC, Wolfe KH. 2008. GenomeVx: simple web-based creation of editable circular chromosome maps. *Bioinformatics* 24(6):861–862.
- Corriveau JL, Coleman AW. 1988. Rapid screening method to detect potential biparental inheritance of plastid DNA and results for over 200 angiosperms. *Am J Bot.* 75:1443–1458.
- Delannoy E, Fujii S, Colas des Francs-Small C, Brundrett M, Small I. 2011. Gene loss in the underground orchid *Rhizanthella gardneri* highlights evolutionary constraints on plastid genomes. *Mol Biol Evol.* 28(7): 2077–2086.
- Do HDK, Kim JS, Kim J-H. 2013. Comparative genomics of four Liliales families inferred from the complete chloroplast genome sequence of *Veratrum patulum* O. Loes. (Melanthiaceae). *Gene* 530:229–235.
- Dong J, et al. 1992. Paternal chloroplast DNA inheritance in *Pinus contorta* and *Pinus banksiana*: independence of parental species or cross direction. *J Hered.* 83(6):419–422.
- Douglas SE. 1998. Plastid evolution: origins, diversity, trends. *Curr Opin Genet Dev.* 8:655–661.
- Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5:1–19.
- Erixon P, Oxelman B. 2008. Whole-genome positive selection, elevated synonymous substitution rates, duplication, and indel evolution of the chloroplast *clpP1* gene. *PLoS One* 3(1):e1386.
- Fay MF, et al. 2006. Phylogenetic of Liliales: summarized evidence from combined analyses of five plastid and one mitochondrial loci. *Aliso* 22: 559–565.
- Goulding SE, Olmstead RG, Morden CW, Wolfe KH. 1996. Ebb and flow of the chloroplast inverted repeat. *Mol Gen Genet.* 252:195–206.
- Guo X, et al. 2007. Rapid evolutionary change of common bean (*Phaseolus vulgaris* L.) plastome, and the genomic diversification of legume chloroplasts. *BMC Genomics.* 8:228.
- Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser.* 41:95–98.
- Hansen AK, Escobar LK, Gilbert LE, Jansen RK. 2007. Paternal, maternal, and biparental inheritance of the chloroplast genome in *Passiflora* (Passifloraceae): implications for phylogenetic studies. *Am J Bot.* 94: 42–46.
- Jansen RK, et al. 2007. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *Proc Natl Acad Sci U S A.* 104(49):19369–19374.
- Kim JS, Hong JK, Chase MW, Fay MF, Kim J-H. 2013. Familial relationships of the monocot order Liliales based on a molecular phylogenetic analysis using four plastid loci: *matK*, *rbcl*, *atpB* and *atpF-H*. *Bot J Linn Soc.* 172(1):5–21.
- Kim JS, Kim J-H. 2013. Comparative genome analysis and phylogenetic relationship of order Liliales insight from the complete plastid genome sequences of two Lilies (*Lilium longiflorum* and *Alstroemeria aurea*). *PLoS One* 8(6):e68180.
- Kim KH, Lee KH, Choi SU, Kim KR, Lee KR. 2010. Pyrrolizidine alkaloids from the roots of *Paris verticillata*. *Heterocycles* 81(6):1493–1502.
- Kurtz S, et al. 2001. REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* 29(22):4633–4642.
- Lee HL, Jansen RK, Chumley TW, Kim KJ. 2007. Gene relocations within chloroplast genomes of *Jasminum* and *Menodora* (Oleaceae) are due to multiple, overlapping inversions. *Mol Biol Evol.* 24(5):1161–1180.
- Lee KH, et al. 2008. A new phenolic Amide from the roots of *Paris verticillata*. *Molecules* 13:41–45.
- Li X, et al. 2013. Complete chloroplast genome sequence of holoparasite *Cistanche deserticola* (Orobanchaceae) reveals gene loss and horizontal gene transfer from its host *Haloxylon ammodendron* (Chenopodiaceae). *PLoS One* 8(3):e58747.
- Lin CP, Wu CS, Huang YY, Chaw SM. 2012. The complete chloroplast genome of *Ginkgo biloba* reveals the mechanism of inverted repeat contraction. *Genome Biol Evol.* 4(3):374–381.
- Liu J, Qi ZC, Zhao YP, Fu CX, Xiang QY. 2012. Complete cpDNA genome sequence of *Smilax china* and phylogenetic placement of Liliales— influences of gene partitions and taxon sampling. *Mol Phylogenet Evol.* 64(3):545–562.
- Logacheva MD, Schelkunov M, Penin AA. 2011. Sequencing and analysis of plastid genome in mycoheterotrophic orchid *Neottia nidus-avis*. *Genome Biol Evol.* 3:1296–1303.
- McCaughey DE, Sundby AK, Bailey MF, Welch ME. 2007. The inheritance of chloroplast DNA is not strictly maternal in *Silene vulgaris* (Caryophyllaceae): evidence from experimental crosses and natural populations. *Am J Bot.* 94:1333–1337.
- McFadden GI. 1999. Endosymbiosis and evolution of the plant cell. *Curr. Opin. Plant Biol.* 2(6):513–519.
- McNeal JR, Heuhl JV, Boore JL, dePamphilis CW. 2007. Complete plastid genome sequences suggest strong selection for retention of photosynthetic genes in the parasitic plant genus *Cuscuta*. *BMC Plant Biol.* 7: 57.
- Schattner P, Brooks AN, Lowe TM. 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res.* 33(Suppl 2):W686–W689.
- Schmickl R, Kiefer C, Dobes C, Koch MA. 2007. Evolution of *trnF* (GAA) pseudogenes in cruciferous plants. *Plant Syst Evol.* 282:229–240.
- Stamatakis A, Hoover P, Rougemont J. 2008. A rapid bootstrap algorithm for the RaxML web-servers. *Syst Biol.* 75(5):758–771.
- Sugiura M. 1992. The chloroplast genome. *Plant Mol Biol.* 19:149–168.
- Untergrasser A, et al. 2012. Primer3-new capabilities and interfaces. *Nucleic Acids Res.* 40(15):e155.
- Wang RJ, et al. 2008. Dynamics and evolution of the inverted repeat-large single copy junctions in the chloroplast genomes of monocots. *BMC Evol Biol.* 8:36.
- Willey DL, Gray JC. 1990. An open reading frame encoding a putative haem-binding polypeptide is contrascribed with the pea chloroplast gene for apocytochrome f. *Plant Mol Biol.* 15: 357–356.
- Wolfe KH, Morden CW, Palmer JD. 1992. Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. *Proc Natl Acad Sci U S A.* 89(22):10648–10652.
- Xiong AS, et al. 2009. Gene duplication, transfer, and evolution in the chloroplast genome. *Biotechnol Adv.* 27:340–347.
- Yang M, et al. 2010. The complete chloroplast genome sequence of date palm (*Phoenix dactylifera* L.). *PLoS One* 5(9):E12762.
- Yang TW, Yang YA, Xiong Z. 2000. Paternal inheritance of chloroplast DNA in interspecific hybrids in the genus *Larrea* (Zygophyllaceae). *Am J Bot.* 87:1452–1458.
- Zhang Q, Liu Y, Sodmergen 2003. Examination of the cytoplasmic DNA in male reproductive cells to determine the potential for cytoplasmic inheritance in 295 angiosperm species. *Plant Cell Physiol.* 44:941–951.

Associate editor: Bill Martin