Compendium of Plant Genomes
Series Editor: Chittaranjan Kole

Philipp Simon
Massimo Iorizzo
Dariusz Grzebelus
Rafal Baranski *Editors*

The Carrot Genome

Springer

Compendium of Plant Genomes

Series Editor

Chittaranjan Kole, ICAR-National Research Center on Plant Biotechnology, Pusa, Raja Ramanna Fellow, Government of India, New Delhi, India

Philipp Simon · Massimo Iorizzo ·
Dariusz Grzebelus · Rafal Baranski
Editors

The Carrot Genome

Springer

Editors
Philipp Simon
Vegetable Crops Research Unit
USDA-ARS
Madison, WI, USA

Massimo Iorizzo
Plants for Human Health Institute
North Carolina State University
Kannapolis, NC, USA

Dariusz Grzebelus
University of Agriculture in Krakow
Kraków, Poland

Rafal Baranski
Faculty of Biotechnology and
Horticulture
University of Agriculture in Krakow
Kraków, Poland

ISSN 2199-4781 ISSN 2199-479X (electronic)
Compendium of Plant Genomes
ISBN 978-3-030-03388-0 ISBN 978-3-030-03389-7 (eBook)
https://doi.org/10.1007/978-3-030-03389-7

Library of Congress Control Number: 2019934354

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Carrot Organelle Genomes: Organization, Diversity, and Inheritance

12

David M. Spooner, Philipp W. Simon, Douglas Senalik and Massimo Iorizzo

Abstract

Cultivated carrot (*Daucus carota* subsp. *sativus*) is one of about 25–40 related wild species in the genus *Daucus* depending on the classification. It is part of a widely distributed and taxonomically complex family Apiaceae (Umbelliferae) containing 466 genera and 3820 species that is one of the largest families of seed plants. Members of the Apiaceae, particularly the genus *Daucus*, have been the subject of intensive recent molecular studies on the structure and genetics of plastids and mitochondria. This chapter summarizes organellar (plastids and mitochondria) structure, function, mutational rates, and inter-organelle DNA transfer in the Apiaceae and inheritance in the genus *Daucus*, with a wider focus on the Apiaceae and the sister family Araliaceae, and places these data in the context of other studies in the angiosperms.

D. M. Spooner (✉) · P. W. Simon · D. Senalik
USDA-Agricultural Research Service, Vegetable Crops Research Unit, Department of Horticulture, University of Wisconsin-Madison, 1575 Linden Dr., Madison, WI 53706, USA
e-mail: David.Spooner@ars.usda.gov

M. Iorizzo
Department of Horticultural Sciences, North Carolina State University, 600 Laureate Way, Kannapolis, NC 28081, USA

12.1 Plastid Structure, Mutational Rates, and Inheritance in Angiosperms

Palmer (1985) provided an early review of plastid structure and gene content, documenting, in angiosperms, (1) its relatively small size (generally 120–160 knt); (2) high copy number (as many as 1000 per cell); (3) quadripartite circular structure comprising two inverted repeats (IR), flanking a large single-copy (LSC) region and a small single-copy (SSC) region; (4) labile structure of the IR region variously shrinking and expanding in different lineages with the junction between the inverted repeat and the large single-copy region located in a generally fixed position within the 276-nt *rps 19* gene; (5) repertoire of a complete set of rRNA, tRNA, and protein-encoding genes (Fig. 12.1); (6) only rare modifications of this basic structure in parasitic plants with reduced gene content, deletion of the IR region in the Fabaceae, or extensive gene rearrangements in the Geraniaceae. In summary, most of the over 200 angiosperm chloroplast genomes examined at that time were overwhelmingly similar in size, conformation, repeat structure, gene content, and gene order and arrangement, with the predominant mode of structural evolution consisting of small deletions and insertions occurring in intergenic spacers, 5′ and 3′ untranslated regions, and in the few introns found in their genes.

P. Simon et al. (eds.), *The Carrot Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-030-03389-7_12

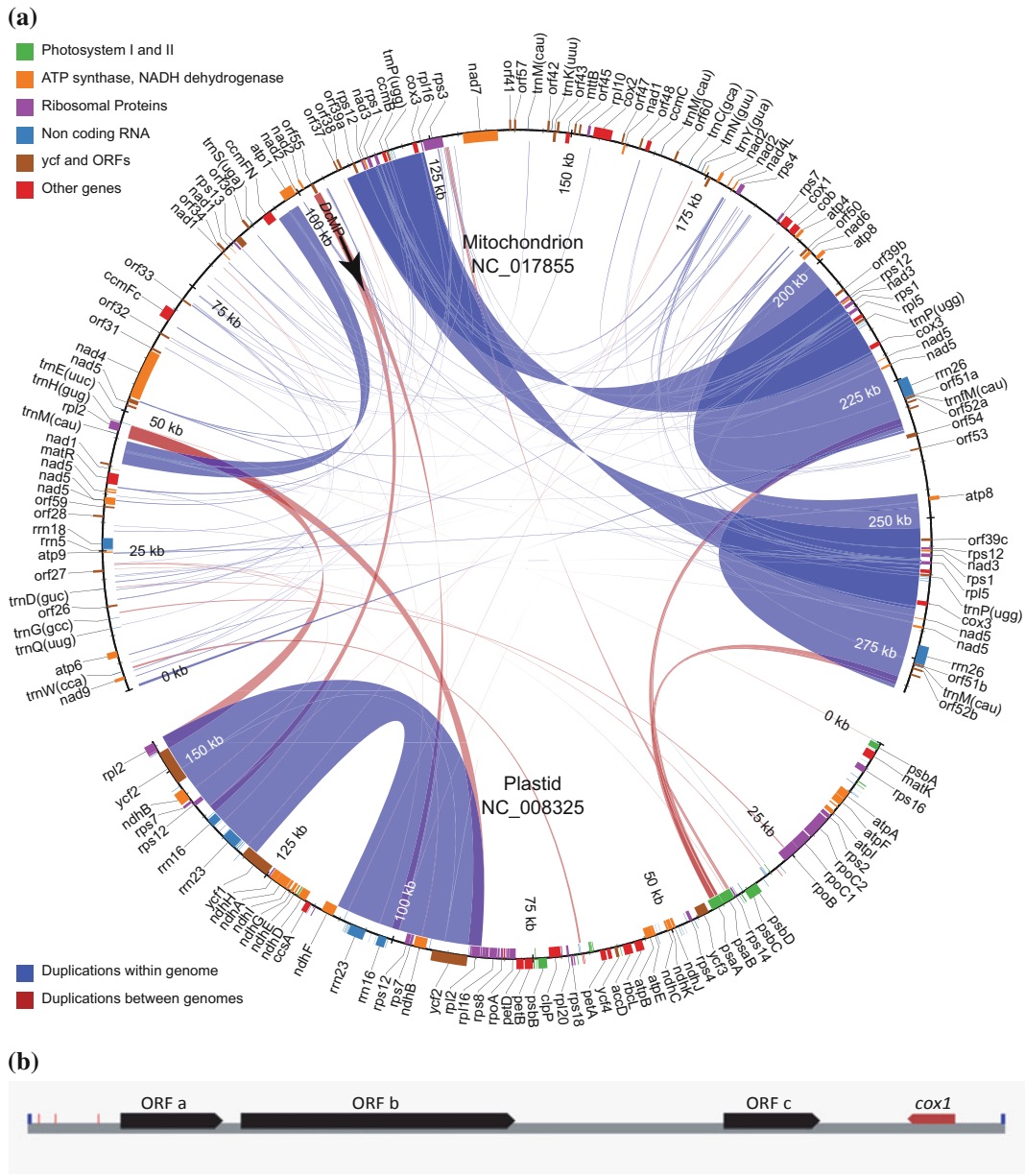

Fig. 12.1 Structure of the carrot mitochondrial and plastid genomes and inter-organelle DNA transfer; genome coordinates every 25 kb are listed inside the figure. **a** Mitochondrial (top) and plastid (bottom) genomes (visualized using Circos version 0.69-6; Krzywinski et al. 2009) and gene annotations of *Daucus carota*; these circularized genomes are drawn open to show gene transfers between them. For the plastid, only genes over 300 nt are annotated for space limitations, but these are collinear with those fully annotated in Ruhlman et al. (2006). Duplications within (blue) and between (red) genomes are shown by connected lines or ribbons. The direction of all duplications between genomes is presumed to be from plastid to mitochondrion except DcMP from mitochondrion to plastid (Iorizzo et al. 2012a, b) as labeled by the arrow. Organellar sequences and gene annotations were obtained from NCBI accessions NC_017855 (mitochondrion) and NC_008325 (plastid). Duplicated regions were detected using BLAST+ version 2.6.0 megablast program (Camacho et al. 2009) with minimum alignment length of 50, minimum percentage similarity of 80, and no dust filtering. **b** Structure of the plastid *D. carota* DcMP sequence. Open reading frames (ORFs) were detected using Open Reading Frame Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). The sequence was oriented according to 5′–3′ (indicated by arrows); ORF orientation is in opposite direction as related to other figures. Thick vertical blue lines indicate target site duplication (TSD). Thin red vertical lines indicate relative position of P1, P2, and P3 tnrV promoters. The red box indicates the region comprising partial sequence of cox1 gene. The scheme is drawn to scale

Palmer (1985) mentioned the maternal inheritance of plastid DNA, documented for most species by Tilney-Bassett (1978). Corriveau and Coleman (1988) developed a rapid cytological screen based on epifluorescence microscopy for maternal inheritance and examined 235 plant species from 80 angiosperm families. They detected putative plastid DNA in the generative and/or sperm cells of pollen from 43 species in 26 genera of 15 families, but not in the generative or sperm cells of pollen from the remaining 192 species (82%), strongly suggesting that they have only maternal inheritance. Their results corroborated most reports of maternal plastid inheritance, and suggested that biparental inheritance of plastids is rare, occurring in about 14% of flowering plant genera, scattered among 19% of the families examined. The carrot plastid genome follows a pattern of maternal inheritance (Vivek et al. 1999). Jansen and Ruhlman (2012) reviewed data on maternal inheritance of plastids in angiosperms and provided a similar figure (80%) for angiosperm species with maternal inheritance, the remaining 20% with biparental inheritance.

Wolfe et al. (1987) compared mutational rates among plant mitochondrial (mtDNA), plastid (cpDNA), and nuclear DNA (nDNA) sequences; and among plant and animal mitochondrial DNA sequences. He documented that (1) in contrast to mammals, where mtDNA evolves at least five times faster than nDNA, angiosperm mtDNA evolves at least five times slower than nDNA, (2) plant mtDNA undergoes much more frequent rearrangements and is larger and variable in size than mammalian mtDNA, (3) cpDNA evolves much slower than plant nDNA, and (4) DNA from the cpDNA IR region evolves much more slowly that the plant LSC or SSC regions. The relative structural conservatism and slower evolution rate of cpDNA in plants made it an ideal molecule for plant phylogenetic studies.

Early plastid phylogenetic studies were based partly on DNA restriction site procedures, but were largely replaced by massive data from next-generation DNA sequencing, stimulating the rapid accumulation of whole plastid DNA sequences. For example, Jansen and Ruhlman (2012) reported the public availability of 200 plastid genomes that as of June 2018 has grown to over 3000 (https://www.ncbi.nlm.nih.gov/genomes/GenomesGroup.cgi?taxid=2759&opt=plastid), allowing for finer comparisons of plastid DNA sequences. Raubeson and Jansen (2005) documented varying rates of change in different regions of the plastid genome, favoring phylogenetic studies at different taxonomic levels. Plastid DNA analyses (first DNA restriction site studies, and then DNA sequences from portions of the genome) dominated much of the molecular phylogenetic literature in the 1980s and 1990s. Jansen and Ruhlman (2012) documented additional lineages of both gymnosperms and angiosperms (the Campanulaceae) deviating from stability of plastid architecture, gene and intron content, and gene order across seed plants. They documented highly rearranged plastomes to exhibit three general phenomena: (1) highly accelerated rates of nucleotide substitutions, (2) an increase in the number of dispersed repeats, many of which are associated with rearranged endpoints, and (3) biparental plastid inheritance. They reviewed studies (e.g., Lilly et al. 2001) documenting deviations from the typical circular arrangement of the plastid molecule, to include multimeric circles or linear and branched structures.

The phylogenetic analysis of 81 plastid genes in 64 sequenced genomes by Jansen et al. (2007) allowed lineage-specific correlations between rates of nucleotide substitutions. They documented gene and intron content in plastids to be highly conserved among the early diverging angiosperms and basal eudicots, but found 62 independent gene and intron losses limited to the more derived monocot and eudicot clades. They showed that most angiosperm plastid genomes contain 113 different genes, 16 of which are duplicated in the inverted repeat, for a total of 129 genes. Intron content was shown to be highly conserved across angiosperms with most genomes containing 18 genes with introns. Like gene losses, intron losses were shown to be restricted to the more derived monocot and eudicot clades. Their fully resolved and strongly supported phylogenetic tree supported the genus *Amborella* as the earliest diverging lineage of flowering plants (now estimated to contain over 257,400 species classified into 52 orders and

about 450 families; Judd et al. 2016), followed by the angiosperm orders Nymphaeales and Austrobaileyales, and provided strong support for a sister relationship between eudicots and monocots.

12.2 Plastid Structure in the Apiales (Apiaceae and the Sister Family Araliaceae)

Our literature survey of the in the Apiales (Table 12.1; data as of May 1, 2018) recovered 79 reports of published genomes in the Apiaceae and 33 reports (112 in total) in the Araliaceae. Like the Jansen et al. (2007) wider survey of the angiosperms, our survey of all 112 Apiales plastid genomes from these two families documents a single circular double-stranded DNA molecule, displaying the typical quadripartite structure of angiosperm plastid genomes, containing 111–114 nonduplicated genes. All plastid genomes are collinear, consistent with the rarity of recombination in plant plastomes (Palmer 1985). Total genome lengths varied from 146,512 in *Angelica nitida* to 171,083 in *Caucalis platycarpos*; with a large single-copy region from 83,553 in *Daucus crinitus* to 94,684 in *Pimpinella rhomboidea*; a small single-copy region ranging from 17,139 in *Crithmum maritimum* to 19,117 in *Schefflera delavayi*; and a pair of inverted repeats from 17,217 nt in *P. rhomboidea* to 27,993 in *C. maritimum*. Average CG contents range from 36.8% in *Eleutherococcus gracilistylus* to 38.1% in *Aralia undulata* and *Panax notoginseng*. The number of nonduplicated genes ranged from 111 in *Bupleurum falcatum* to 114 in many other species.

12.3 Plastid Structure in *Daucus* Sensu Lato

All reports of *Daucus* in its expanded sensu (sensu lato, Banasiak et al. 2016, see Chap. 2) likewise documented a typical chloroplast quadripartite circular genome consisting of a total length in nt varying from 155,441 in *Daucus involucratus* to 157,336 in *Daucus setulosus*; a large single-copy region from 83,553 in *D. crinitus* to 84,444 in *Rouya polygama*; a small single-copy region 17,314 in *R. polygama* to 17,887 in *Daucus tenuisectus*; and a pair of inverted repeats 26,924 nt in *Daucus bicolor* to 27,741 in *Daucus aureus*. Spooner et al. (2017) did not report average GC contents but they documented an inverse relationship between read coverage and GC content, most notably in the second half of the inverted repeat region, as seen in the coverage plots (Fig. 12.2). This observation is likely a reflection of the Illumina platform that introduces coverage bias in regions with high GC content (Ross et al. 2013). All reports documented 113 unique genes consisting of 80 protein-coding genes, 29 tRNA genes, and 4 RNA genes.

The inverted repeat junctions flanking the LSC were identical in all genotypes examined by Spooner et al. (2017), while those flanking the SSC were variable (Fig. 12.3). These variations form six distinct classes (A–F), with the out-group *Oenanthe virgata* (class F) having the largest fraction of the *ycf1* gene included in the inverted repeat, including a 9-nt insertion unique to this species. Relative to *Oenanthe*, class A consists of 15 accessions, which includes *D. carota*, and has a 326-nt contraction (reduction in the size of the inverted repeat); class B consisting of only *D. aureus* has the largest contraction, 422 nt; class C consisting of five accessions has a 318-nt contraction; class D consisting of 15 accessions has a 319-nt contraction; and class E consisting of only *C. platycarpos* has a 50-nt contraction. Relative to the plastid phylogeny of Spooner et al. (2017), there is a direct cladistic relationship of these inverted repeat junction classes with all accessions of *D. carota* and its immediate sister species *Pseudorlaya pumila* and *Rouya polygama* having class A; *D. aureus* class B; *D. muricatus*, *D. tenuisectus*, and *D. crinitus* class C; *D. conchitae*, *D. crinitus*, *D. glochidiatus*, *D. littoralis*, *D. pusillus*, *D. setulosus*, class D; out-group *Caucalis platycarpos* class E; and out-group *O. virgata* class F.

Table 12.1 Summary of genome statistics of fully sequenced plastids of members of the Apiaceae and sister family Araliaceae

Species	Reference	Total length in nucleotides	Large single copy	Small single copy	Percent average CG content	Inverted repeat	Number of unique (nonduplicated) genes
Apiaceae							
Anethum graveolens L.	NCBI: NC_029470	153,356					
Angelica acutiloba (Siebold & Zucc.) Kitag.	NCBI: NC_029391.1	147,074					
Angelica dahurica (Fisch.) Benth. & Hook.f.	NCBI: NC_029392	146,918					
Angelica decursiva (Miq.) Franch. & Sav.	Choi et al. (2016b)	146,719	93 256	17,497	37.56	17,983	113
Angelica gigas Nakai	Choi et al. (2016a)	146,916	93,118	17,582		18,108	113
Angelica gigas	NCBI: KX118044.1	152,185					
Angelica nitida H. Wolff	Deng et al. (2017)	146,512	93,298	18,068	37.48	17,573	113
Anthriscus cerefolium (L.) Hoffm	Downie and Jansen (2015)	154,719	84,774	17,551	37.4	26,197	
Arracacia xanthorrhiza Bancr.[a]	Alvarado et al. (2017)	143,989	49,169	17,439	37.48	31,370	106
Bupleurum boissieuanum H. Wolff	Wu et al. (2017)	156,108	86,007	17,495	37.7	26,303	112
Bupleurum falcatum L.	Shin et al. (2016)	155,989	85,912	17,517		26,280	111
Bupleurum latissimum Nakai	NCBI: NC_033346	155,621					
Carum carvi L.	NCBI: NC_029889.1	155,449					113
Caucalis platycarpos L.	Spooner et al. (2017)	171,083	85,042	17,553			113
Chuanminshen violaceum Sheh et Shan	Yuan et al. (2017)	154,529	84,171	17,800	37.8	26,279	112
Coriandrum sativum L.	NCBI: NC_029850	146,519					
Crithmum maritimum L.	Downie and Jansen (2015)	158,355	85,230	17,139	37.6	27,993	
Daucus aureus Desf.	Spooner et al. (2017)	156,984	83,655	17,846		27,741	113
Daucus bicolor Sm. in Sibth. and Sm	Spooner et al. (2017)	155,785; 155,833	84,282; 84,261	17,677; 17,677		26,924; 26,942	113; 113
Daucus capillifolius (Gilli) C. Arbizu	Spooner et al. (2017)	155,906	84,259	17,552		27,047	113
Daucus carota L. subsp. *carota*	Spooner et al. (2017)	155,676; 155,865; 155,870; 155,908; 155,909	84,102; 84,213; 84,250; 84,243; 84,243	17,503; 17,555; 17,527; 17,570; 17,571		27,035; 27,048; 27,046; 27,047; 27,047	113; 113; 113; 113; 113

(continued)

Table 12.1 (continued)

Species	Reference	Total length in nucleotides	Large single copy	Small single copy	Percent average CG content	Inverted repeat	Number of unique (nonduplicated) genes
Daucus carota subsp. *gummifer* (Syme) Hook.f.	Spooner et al. (2017)	155,857; 155,876; 155,883; 155,970	84,232; 84,257; 84,202; 84,323	17,528; 17,560; 17,594; 17,550		27,048; 27,028; 27,043; 27,048	113; 113; 113; 113
Daucus carota subsp. *maximus* (Desf.) Ball	Spooner et al. (2017)	155,870	84,250	17,527		27,046	113
Daucus carota subsp. *sativus*	Ruhlman et al. (2006)	155,911	84,243	17,571		27,048	113 (115)[b]
Daucus conchitae Greuter	Spooner et al. (2017)	155,835; 156,787; 156,821	84,227; 83,738; 83,735	17,676; 17,681; 17,682		26,966; 27,684; 27,702	113; 113; 113
Daucus crinitus Desf.	Spooner et al. (2017)	156,342; 156,388	83,553; 83,592	17,822; 17,829		27,483; 27,483	113; 113
Daucus glochidiatus (Labill.) Fisch., C. A. Mey. & Avé-Lall.	Spooner et al. (2017)	155,914	84,208	17,657		27,024	113
Daucus guttatus Sibth. and Sm.	Spooner et al. (2017)	157,194; 157,197	84,208; 84,231	17,657; 17,678		27,024; 27,644	113; 113
Daucus involucratus Sm.	Spooner et al. (2017)	155,441; 155,479	83,749; 83,738	17,717; 17,699		26,987; 27,021	113; 113
Daucus littoralis Sibth. and Sm.	Spooner et al. (2017)	156,923	83,940	17,698		27,642	113
Daucus muricatus L.	Spooner et al. (2017)	156,011; 156,052	83,905; 83,946	17,881; 17,881		27,112; 27,112	113; 113
Daucus pusillus Michx.	Spooner et al. (2017)	156,939; 157,032	84,191; 84,237	17,427; 17,451		27,667; 27,667	113; 113
Daucus setulosus Guss. ex DC.	Spooner et al. (2017)	157,292; 157,336	84,267; 84,311	17,681; 17,681		27,672 27,672	113; 113
Daucus syrticus Murb.	Spooner et al. (2017)	155,841; 155,898	84,208; 84,228	17,540; 17,585		27,046 27,042	113; 113
Daucus tenuisectus Coss. ex Batt.	Spooner et al. (2017)	156,931	83,615	17,887		27,714	113
Foeniculum vulgare Mill.	NCBI: NC_029469	153,628					
Glehnia littoralis	S-C. Lee et al. (2016b)	147,467	93 493	17,546		18,214	114
Glehnia littoralis	NCBI: KU866532	147,477					
Hansenia forbesii (H. Boissieu) Pimenov and Kljuykov	NCBI: NC_035054, NC_035056	159,287; 159,505					

(continued)

Table 12.1 (continued)

Species	Reference	Total length in nucleotides	Large single copy	Small single copy	Percent average CG content	Inverted repeat	Number of unique (nonduplicated) genes
Hansenia oviformis (R. H. Shan) Pimenov and Kljuykov	NCBI: NC_035055	157,292					
Hansenia weberbaueriana (Fedde ex H. Wolff) Pimenov and Kljuykov	NCBI: NC_035053	158,625					
Ledebouriella seseloides (Hoffm.) H. Wolff	H. O. Lee et al. (2016a)	147,880	93,222	17,324	37.5	18,667	113
Ligusticum tenuissimum (Nakai) Kitag	NCBI: NC_029394	158,500					
Notopterygium forrestii H. Wolff	Yang et al. (2017)	159,607	88,870	18,212	37.70	26,262	113
Notopterygium franchetii H. de Boissieu	Yang et al. (2017)	159,389	88,749	18,260	37.70	26,175	113
Notopterygium incisum C. C. Ting ex H. T. Chang	Yang et al. (2017)	158,684	88,260	18,232	37.70	26,096	113
Notopterygium oviforme R. H. Shan	Yang et al. (2017)	157,462	87,303	17,996	37.90	26,081	113
Oenanthe virgata Poir	Spooner et al. (2017)	154,218	84,411	17,163		26,445	
Ostericum koreanum Kitagawa	Choi et al. (2016c)	147,282	93,185	17,663	37.54	18,217	113
Pastinaca pimpinellifolia M. Bieb.	NCBI: NC_027450.1	149,758					
Petroselinum crispum (Mill.) Fuss	Downie and Jansen (2015)	152,890	86,116	17,508	37.8	24,633	
Peucedanum insolens Kitag.	NCBI: NC_033344	156,912					
Peucedanum japonicum Thunb.	NCBI: NC_034644	164,653					
Pimpinella rhomboidea var. *tenuiloba* Shan and Pu	Tan and Yu (2018)	146,655	94,684	17,537		17,217	113
Pleurospermum camtschaticum Hoffm.	NCBI: NC_033343.1	155,415					
Prangos trifida (Mill.) Herrnst. et Heyn	Samigullin et al. (2017)	153,510	86,481	17,445		24,792	113
Pseudorlaya pumila Grande	Spooner et al. (2017)	155,672	84,042	17,570		27,030	113
Pterygopleurum neurophyllum (Maxim.) Kitag.	NCBI: NC_033345.1	154,369					
Rouya polygama Coincy	Spooner et al. (2017)	155,864	84,444	17,314		27,053	113
Seseli montanum L.	Samigullin et al. (2016)	147,823	92,620	17,481	37.57	18,861	114
Tiedemannia filiformis subsp. *greenmannii* (Mathias and Constance) M. A. Feist and S. R. Downie	Downie and Jansen (2015)	154,737	84,535	17,140	37.3	26,506	

(continued)

Table 12.1 (continued)

Species	Reference	Total length in nucleotides	Large single copy	Small single copy	Percent average CG content	Inverted repeat	Number of unique (nonduplicated) genes
Araliaceae							
Aralia elata (Miq.) Seem.	Kim et al. (2017)	156,220					
Aralia undulata Hand.-Mazz.	Li et al. (2013)	156,333	86,028	18,089	38.1	26,108	114
Brassaiopsis hainla (Buch.-Ham.) Seem.	Li et al. (2013)	156,459	86,566	18,021	38.0	25,936	114
Dendropanax dentiger (Harms) Merr.	Wang et al. (2016)	156,687	86,680	18,247	38.0	25,880	114
Dendropanax morbifera H. Lev.	Kim et al. (2017)	156,366					
Eleutherococcus brachypus (Harms) Nakai	Zhang et al. (2018)	156,981	86,921	18,184		25,938	114
Eleutherococcus gracilistylus (W. W. Sm.) S. Y. Hu	Kim et al. (2016a)	156,770	86,729	18,175	36.8	25,938	113
Eleutherococcus senticosus (Rupr. & Maxim.) Maxim.	Yi et al. (2012)	156,768	86,755	18,153		25,930	
Eleutherococcus sessiliflorus (Rupr.& Maxim.) S.Y.Hu	Kim et al. (2017)	156,730					
Fatsia japonica (Thunb.) Decne. & Planch.	Chen et al. (2016)	155,613	86 487	17,866	37.91	25,929	114
Hydrocotyle sibthorpioides Lam.	Ge et al. (2017)	152,880	84,064	18,690		25,063	113
Hydrocotyle verticillata Thunb., non Turcz.	Downie and Jansen (2015)	153,207	84,352	18,739	37.6	25,058	
Kalopanax septemlobus (Thunb.) Koidz.	Li et al. (2013)	156,413	86,466	18,119	38.0	25,914	114
Metapanax delavayi (Franch.) J. Wen and Frodin	Li et al. (2013)	156,343	86,360	18,131	38.0	25,926	114
Panax bipinnatifidus Seem.	Manzanilla et al. (2018)	156,248					
Panax ginseng C. A. Mey.	Zhao et al. (2015)	156,354; 156,355	86,129; 86,130	18,007; 18,007		26,074; 26,074	114; 114
Panax ginseng	Kim et al. (2017)	156,248					
Panax japonicas C. A. Mey.	Kim et al. (2017)	156,188					
Panax notoginseng (Burk.) F. H. Chen	Dong et al. (2014)	156,387					
Panax notoginseng	Zhang et al. (2016)	156,324	86,082	18,032	38.1	26,105	114
Panax notoginseng	Kim et al. (2017)	156,466					
Panax quinquefolius L.	Han et al. (2016)	156,359	86,184	18,081	38.08	26,076	114
Panax quinquefolius	Kim et al. (2016b)	156,088					
Panax schin-seng T.Nees	Kim and Lee (2004)	156,318	86,106	18,070		26,071	114

(continued)

Table 12.1 (continued)

Species	Reference	Total length in nucleotides	Large single copy	Small single copy	Percent average CG content	Inverted repeat	Number of unique (nonduplicated) genes
Panax stipuleanatus H. T. Tsai and K. M. Feng	Manzanilla et al. (2018)	156,090					
Panax stipuleanatus	NCBI: NC_030598.1	156,064					
Panax vietnamensis Ha and Grushv.	Kim et al. (2017)	155,993					
Panax vietnamensis	Manzanilla et al. (2018)	156,022; 156,099					
Schefflera delavayi (Franch.) Harms	Li et al. (2013)	156,341	86,112	19,117	37.8	25,551	114
Schefflera octophylla (Lour.) Harms	Zong et al. (2016)	156,685	86,609	18,146	37.93	25,965	

[a]We report the numbers for *Arracacia xanthorrhiza* from Alvarado et al. (2017) but do not use them in our summaries in the text because of the atypical calculations in this paper
[b]Ruhlman et al. (2006) report 115 unique plastid genes, but Jansen et al. (2007) correct this to 113

The plastids of members of *D. carota* sensu lato have variable numbers of repeats (scanned for minimum length 30 nt) between 13 and 18, with a minimum size of 70 nt for *R. polygama* and a maximum size of 127 nt in *D. crinitus*. Twenty-five accessions share a maximum repeat size of 88, three accessions 106 nt, and two accessions 109 nt. Species in closely related clades share a larger number of repetitive sequences (Spooner et al. 2017).

12.4 Mitochondrial Structure and Function in Angiosperms

Mitochondrial DNA has the same basic role in plants as it does in other eukaryotes, encoding a small number of essential genes of the mitochondrial electron transfer chain. For the expression of these few genes, the mitochondrion has its own translation system that is also partially encoded by the mtDNA, including rRNAs, tRNAs, and a variable number of ribosomal proteins that vary across different species (Kubo and Newton 2008). A few proteins involved in the assembly of functional respiratory complexes are encoded by the plant mtDNA. However, all factors required for the maintenance of the mtDNA and the expression of its genes are encoded in the nucleus and imported from the cytosol, thus placing mtDNA replication, structural organization, and gene expression under nuclear control.

Although the number of mitochondrial genes varies little between species, the size of the mtDNA varies over more than a 100-fold, with land plant mitochondrial genomes by far the largest. Angiosperm mitogenomes are usually in the range of 200–700 kb, but can be as large as 11 Mb in *Silene conica* (Sloan et al. 2012). Although a few additional genes exist in plant mitogenomes, and several genes contain introns, these features do not contribute significantly to the large size or the size variation of plant mtDNA. Rather, most of the genome consists of noncoding sequences that are not conserved across species. Horizontal transfer seems to be responsible for the acquisition of exogenous

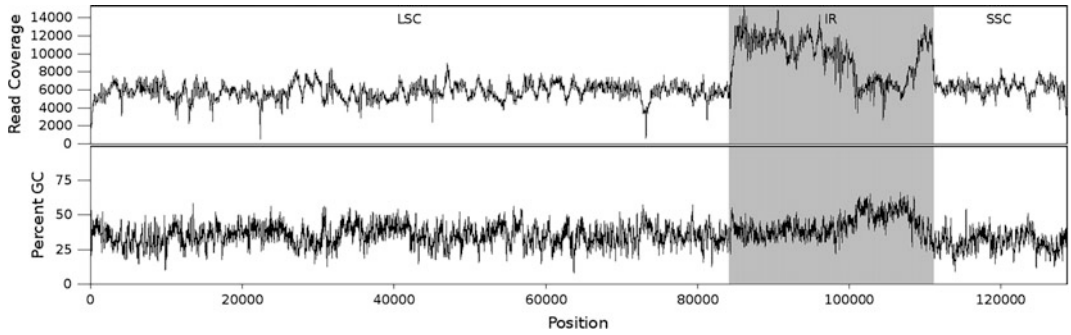

Fig. 12.2 Read coverage and percent GC plots spanning the plastid genome of *Daucus carota* subsp. *carota* PI 274297; inverted repeat regions highlighted in gray

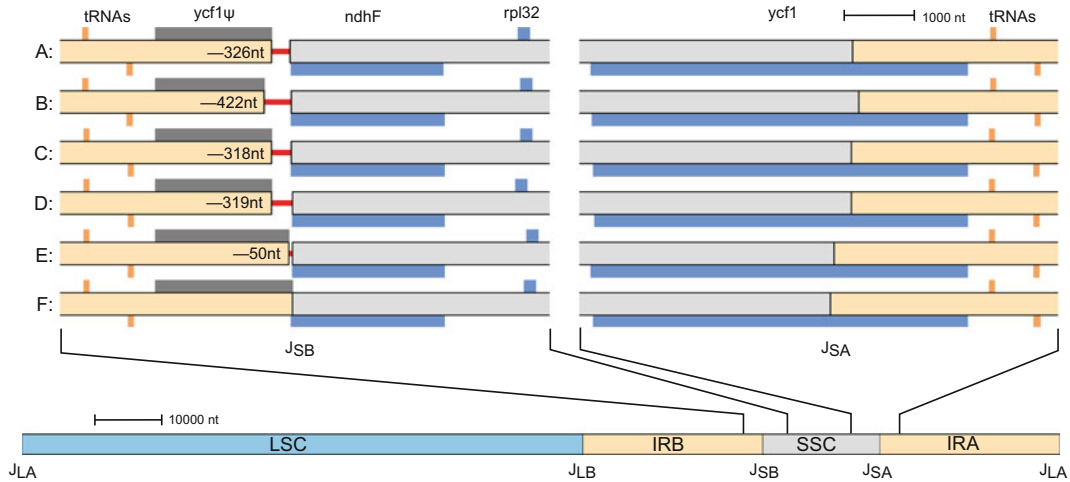

Fig. 12.3 Junctions of the inverted repeats and small single-copy plastid regions. Functional genes are represented in blue, tRNA in tan, and pseudogenes in gray. Numbers in the figure represent the number of nucleotides no longer present in inverted repeat B relative to *Oenanthe virgata*. A (*Daucus carota* NC_008325.1) is representative of 14 additional genotypes; B (*D. aureus* 319,403) is unique to this genotype; C (*D. crinitus* 652,413) is representative of four additional genotypes; D (*D. guttatus* 286,611) is representative of 14 additional genotypes; E (*Caucalis platycarpos* 649,446) and F (*O. virgata* Ames 30,293) are unique to these genotypes

sequences (Bergthorsson et al. 2003), and a fraction of plant mitogenomes can be recognized as derived from plastid, nuclear, or viral DNA. However, most noncoding sequences are of unknown origin.

The structure of angiosperm mitochondrial genomes is frequently characterized by repeat sequences (Gualberto et al. 2014). The number and the size of these repeats are important, as they influence the size of the genome, and they are the sites of intragenomic recombination, underlining evolutionary changes in mitochondrial genome organization and structural dynamism in vivo (Guo et al. 2017; Gupta et al. 2013). The repeats have often been classified as large repeats (>500 nucleotides), which can be involved in frequent homologous recombination; intermediate-size repeats (50–500 nucleotides), which are involved in infrequent ectopic homologous recombination; and small repeats (<50 nucleotides), which can promote illegitimate microhomology-mediated recombination (Arrieta-Montiel et al. 2009; Davila et al. 2011; Gualberto et al. 2014). Based on the very active recombination behavior of large

repetitive sequences, early studies postulated that the entire genetic content of mtDNA could be assembled into a circular molecule, the so-called master circle, from which multiple subgenomic circular molecules are generated by intramolecular recombination across direct repeats. Although the repetitive sequences across species are not conserved, their organization and structure, which drive the recombination process, are conserved. Recent studies based on gel-based approaches or electron microscopy and quantitative sequence data from next-generation sequencing have indicated that circular and linear forms of mtDNA co-exist in vegetative tissue. Sequencing data also revealed the evolution of multichromosomal genomes associated with genome size expansion.

An economically important trait that can result from intraspecific variation promoted by recombination within mitogenomes is cytoplasmic male sterility (CMS)—the maternally transmitted inability of a plant to produce viable pollen. CMS is widespread in natural plant populations and is important for the evolution of gynodioecious species, in which females and hermaphrodites co-occur in populations (Dufay et al. 2007). In crop breeding, including in carrot it is an economically valuable trait used extensively for the production of hybrid seeds (see Chap. 3). It usually results from the expression of a chimeric gene created de novo by recombination processes, particularly microhomology-mediated recombination events, each of which involves just a few nucleotides of sequence identity. Multiple CMS phenotypes in carrot have been described and are used in breeding programs. A maternal mode of inheritance of the mitochondrial (mt)DNA has been observed in carrot CMS plants by several authors, and different genes/ORFs have been proposed to control this important trait (see Chap. 3).

Given the larger genome size relative to plastid, the diversity of repetitive sequences, and its dynamic organization, assembling mitochondrial genomes is challenging, and for this reason the number of mitochondrial genomes available is far lower than the plastomes.

12.5 Carrot Mitochondrial Genome, Structure, and Organization

In 2012, Iorizzo et al. (2012a) assembled and characterized the carrot mitochondrial genome, the first and still the only mitochondrial genome sequenced in the Apiaceae. With 281,132 nt, the carrot mitogenome is among the smallest mitochondrial genomes sequenced to date among the angiosperms and confirmed previous estimation (255,000 nt) made by Robison and Wolyn (2002) based on restriction digestion mapping. Although the genome could be assembled and represented as a master circle, Southern blot analysis confirmed the presence of two recombinant sub-circles. The overall GC content of carrot (45.4%) is comparable to other angiosperms (Alverson et al. 2011; Rodriguez-Moreno et al. 2011).

Annotation of the genome identified 44 protein-coding sequences and three ribosomal RNAs, which confirmed the previous report of Adams et al. (2002) based on Southern hybridization that surveyed mitochondrial gene presence or loss across 280 angiosperms. Truncated copies of *atp1* and *atp9* were detected, confirming observations previously reported by Bach et al. (2002). Considering a set of 51 mitochondrial conserved genes, the carrot mitogenome lack 7 genes (*sdh3*, *sdh4*, *rpl2*, *rps2*, *rps10*, *rps14,* and *rps19*), and three of them were identified in the carrot genome assembly. In addition to coding genes, the carrot mitogenome contains 18 tRNAs that recognize 15 amino acids and is missing tRNA genes for six amino acids, which are likely coded by the nuclear genome.

As expected, intergenic spacer regions represent the largest part of the genome, 224,526 nt (79.9%), with repetitive sequences occupying the majority of this space (49%). With 74 repeats ranging from 37 to 14,749 nt, the carrot mitochondrial genome has the lowest number of repeats among the sequenced plant mitochondrial genomes, which reflect its small genome size. All but one are dispersed repeats. Most of the repeats (about 90%) are between 20 and 202 nt in length accounting for just 2.0% of the total genome

coverage. Nine large repeats ranging from 4220 to 14,749 nt account for 44.0% of the genome. The insertion of the large repeat 1, between repeat 2 and 3, forms a 35 kb super-repeat. After wild cabbage (Chang et al. 2011), this is the largest repeat region described in eudicot mitochondrial genomes to date. Other sequences in the intergenic spacer regions include additional open reading frames not associated with any conserved mt genes, and DNA of nuclear or plastid origin, derived from intracellular gene transfer (IGT) or possibly horizontal gene transfer (HGT), a prevalent and ongoing process in plant evolution.

12.6 Intracellular DNA Transfer in Angiosperms

While nuclear and mitochondrial genomes integrate foreign DNA via IGT and HGT, plastid genomes (plastomes) have resisted foreign DNA incorporation and only recently has IGT been uncovered in the plastomes of a few land plants. The emergence of contemporary genomics has dispelled traditional hypotheses of the sole evolution by vertical descent with modification. Drawing on phenotypic data, early investigators could not have predicted the impact of HGT on both the universality of the genetic code and diversity of organisms found on earth (Vetsigian et al. 2006). Although first recognized among eubacteria (Tatum and Lederberg 1947), HGT occurs across all domains of life and has shifted our views on the phylogeny of organisms from one of bifurcation to a more reticulate, web-like mode of evolution (Soucy et al. 2015).

Just as the sharing of DNA sequences among unrelated organisms has shaped their evolutionary history, so has the transfer of sequences among the genome-bearing compartments of individual cells shaped the evolution of eukaryotic species. Intracellular gene transfer, along with HGT, has played a pivotal role in the evolution of multicellularity and the oxygenation of earth's atmosphere, facilitating the evolution of plant and animal life (Timmis et al. 2004). The free-living, single-celled organisms that ultimately became mitochondria, and later plastids, of eukaryotic cells through endosymbiosis contained the necessary complement of genetic material for survival in the extracellular environment. Once housed within the host cell, much of that genetic material was transferred to the host nuclear genome. This massive transfer of DNA sequence fully integrated the processes of the organelles with those of the host nucleus.

Since the establishment of the cellular organelles, both mitochondrial and plastid genomes (mitogenomes and plastomes) of plants have continued to divest themselves of both coding and noncoding DNA. While mitogenomes exhibit more variation in overall size and retained gene content (Adams et al. 2002), most plastomes harbor a conserved set of coding sequences within a relatively stable size and configuration, with a small set of genes that tend to be transferred to the nucleus across the plant phylogeny (Jansen and Ruhlman 2012). The transfer of DNA sequence from both organelles to the nucleus is an ongoing process that has contributed to the evolution of the nuclear genome, regardless of whether those sequences were eventually purged from their original location or activated for their ancestral function elsewhere in the cell following nuclear transcription (Timmis et al. 2004). Likewise, plant mitogenomes contain extensive insertions of both plastid and nuclear DNA (nDNA), although, for the most part, these remain nonfunctional (Mower et al. 2012). Plastomes, however, appear to be recalcitrant to the incorporation of foreign DNA either by HGT or IGT, possibly because of the lack of an efficient DNA uptake system within plastids (Bock 2015; Richardson and Palmer 2007; Smith 2011).

Among the >3000 complete angiosperm plastomes now available in GenBank (https://www.ncbi.nlm.nih.gov/genbank/), just a few lineages have been recognized to contain DNA of nonplastome origin. Although a few studies explored putative plastome sequences with high identity to mtDNA, for the most part, the identity was due to the presence of sequences of plastid or nuclear origin in mitogenomes (Chumley et al. 2006; Ohtani et al. 2002).

The notion that land plant plastomes could incorporate foreign DNA sequences without biotechnological intervention was unheard of prior to 2009 (Goremykin et al. 2009). To date, legitimate cases of foreign DNA insertions into the plastome have been reported in four unrelated families/genus of angiosperms including *Daucus* (Iorizzo et al. 2012a), Apocynaceae (Straub et al. 2013), Bambusoideae (Ma et al. 2015), and *Anacardium* (Rabah et al. 2017). Identification of these rare events have been facilitated in part by the availability of complete mitogenome sequences. Given the wide distribution of these four families across four orders of land plants: Apiales (asterid II), Gentianales (asterid I), Sapindales (rosid II), and Poales (commelinid) combined with the lack of informative common features, suggested at least four independent events across all land plants, which likely occurred only once within each clade.

12.7 Inter-organelle DNA Transfer in the Apiaceae, a Story of First Discoveries

Goremykin et al. (2009), while analyzing the *Vitis vinifera* L. (grape) mitochondrial genome, detected two sequences of 74 and 126 nt which were similar to the carrot plastid genome (Ruhlman et al. 2006). The larger sequence has high similarity to the coding region of the mitochondrial cytochrome c oxidase subunit 1 gene (*cox1*), prompting the authors to suggest that its presence in the *Daucus* plastome might possibly represent a rare transfer of DNA from the mitochondrion into the plastid. These two sequences are contained within a large 1439-nt fragment of the *D. carota* inverted repeat at positions 99,309–100,747 and 139,407–140,845 (Ruhlman et al. 2006) that is a part of the 30rps12-trnV-GAC intergenic spacer region. This fragment, however, has no similarity to any other published plastid nucleotide region (Goremykin et al. 2009). Subsequently, Iorizzo et al. (2012a), in characterizing the entire carrot mitochondrial genome, verified the presence of this sequence in both plastid and mitochondrial genomes and designated this site as the *D. carota* mitochondrial-plastid (DcMP) region (Fig. 12.1a). The DcMP sequence is 1452 nt-long in the carrot plastome and is present as three noncontiguous, rearranged sequences in the mitochondrial genome of *D. carota* (Iorizzo et al. 2012a). In the plastome, however, the DcMP 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. Analysis of the plastid DcMP sequence identified three putative open reading frames (ORFs) with similarity to retrotransposon element domains (*gag* domain and reverse transcriptase) and a 6 nt direct repeat (CTTGAC), flanking the DcMP sequence, upstream of DcMP1, and downstream of DcMP4 (Fig. 12.1b) (Iorizzo et al. 2012b). These characteristics suggested that the DcMP might be a non-LTR retrotransposon and the direct repeats represent target site duplication (TSD) created because of the DcMP integration following its mobilization from a donor site localized in the mitochondrial genome. Overall, these two complementary studies demonstrated for the first time that DNA transfer from the mitochondrion to the plastid can occur in flowering plants and provided a hypothesis about its possible mode of integration.

Considering the stability of the plastid genome, it is legitimate to hypothesize that a mt-to-pt insertion within a phylogenetic clade is likely to have originated from a single event in a common ancestor, making this type of insertion useful to trace ancestry and genetic relationships within the Scandiceae tribe, which includes three subtribes Daucinae, Torilidinae, and Scandicinae. Analysis of 37 plastid genomes including members of the Daucinae and Torilidinae subtribes indicated that the DcMP region was detected in all 36 members of the Daucinae clade and in *C. platycarpos*, a member of the Torilidinae clade (Spooner et al. 2017). Comparative analysis of the DcMP region across the 37 plastid genomes revealed 21 structural variants (SVs) (insertions or deletions) (Fig. 12.4). Relative to the plastid phylogeny of Spooner et al. (2017), there is a direct cladistic relationship of these SVs with all accessions of *Daucus* and its immediate sister species

P. pumila, *R. polygama, and C. platycarpos* (Fig. 12.4). To expand the search for DcMP insertion within the Apiaceae, Downie and Jansen (2015) compared the plastomes of six Apiaceae species (*C. maritimum*, *D. carota*, *Hydrocotyle verticillata*, *Petroselinum crispum*, and *Tiedemannia filiformis* subsp. *greenmani)* including *Anthriscus cerefolium*, a member of the Scandicinae subtribe. Despite the observation that another putative insertion of mtDNA, unrelated to DcMP is present in the plastid genome of *P. crispum,* none of these six plastid genomes contain the DcMP sequence. Overall, these two studies indicated that the DcMP insertion is restricted to the Torilidinae subtribe (*C. platycarpos*) and Daucinae (36 species), which implies that within the Scandicinae tribe these two subtribes are genetically more closely related as compared with the Scandicinae subtribe where the insertion has not been detected. This hypothesis is supported by previous systematic and molecular marker work (Lee and Downie 2000; Lee et al. 2001) and confirms our hypothesis that detection of the DcMP sequence can be used as a marker to delineate relationships in this clade.

Sequence analysis of the DcMP regions detected in 36 species (Spooner et al. 2017) revealed other important aspects related to IGT in plants. Within the DcMP region, two large insertions were detected in the *C. platycarpos* plastid genome, named Cp MP5 (6663 nt) and Cp MP6 (360 nt). A large portion of the Cp MP5 sequence (KX832334 from 102,567 to 105,470) shares a high similarity (91% identity) with DCAR_022437, a nuclear gene located on carrot Chr6 annotated as an auxin response factor (ARF). The alignment covers seven of the 14 DCAR_022437 predicted exons, and none of its flanking nuclear sequences shares similarity with other plastid sequences (Fig. 12.5a). These findings represent the first evidence of a known nuclear sequence inserted in a plastid genome. Either the plastid ARF DNA sequence found in *C. platycarpos* could be part of the ancestral mitochondrial DcMP sequence, or it could have been transferred directly from the nucleus or mitochondrion into the plastid after the mt-to-pt DcMP insertion occurred. The mechanism of transfer of this nuclear DNA relative to the insertion of DcMP in the plastid genome is unknown. However, the sequence covering the DcMP and CpMP regions documented in *C. platycarpos* contains an intact *cox1* copy and fragments of *ARF* gene. Indeed, the Cp MP5 3′ end and Cp MP6 5′ end are contiguous to the pt-DcMP2 sequence and the carrot mt-Dc MP2 flanking sequences and cover the full length of the mitochondrial *cox1* gene (Fig. 12.5b). These findings indicate that direct insertion of nDNA into the plastome at the very same locus as mtDNA insertion is implausible compared with its insertion along with the mtDNA, as mitogenomes of land plants contain abundant foreign DNA from both IGT and HGT events (Knoop 2004; Alverson et al. 2010; Park et al. 2014). In particular, an *ARF* gene (*ARF17*) has been transferred to the mitogenome in several genera of Brassicaceae (Qiu et al. 2014).

In higher plants, horizontally transferred DNA is generally not functional in the recipient genome (Bock 2015; Richardson and Palmer 2007). In contrast, in carrot the DcMP sequence integrated three new functional promoters (P1, P2, and P3) located 105-, 41-, and 16-nt upstream of trnV, respectively, at the 3′—DcMP insertion junction. According to Manna et al. (1994), all three promoters are expressed in carrot cells and were responsible for the differential expression of trnV during embryogenesis. Assuming that all three promoters have a functional role, we expect their sequences to be conserved. Across all the samples harboring the pt-DcMP insertion, SVs resulted in the deletion of the P1 or P2 promoter sequences in at least one species (Spooner et al. 2017). In contrast, despite the observation that multiple independent insertion or deletion events occurred in the DcMP-4 region near the P3 promoter, its sequence is conserved across all accessions harboring the DcMP insertion (Fig. 12.4). Considering correct the hypothesis proposed by Manna et al. (1994) that the P3

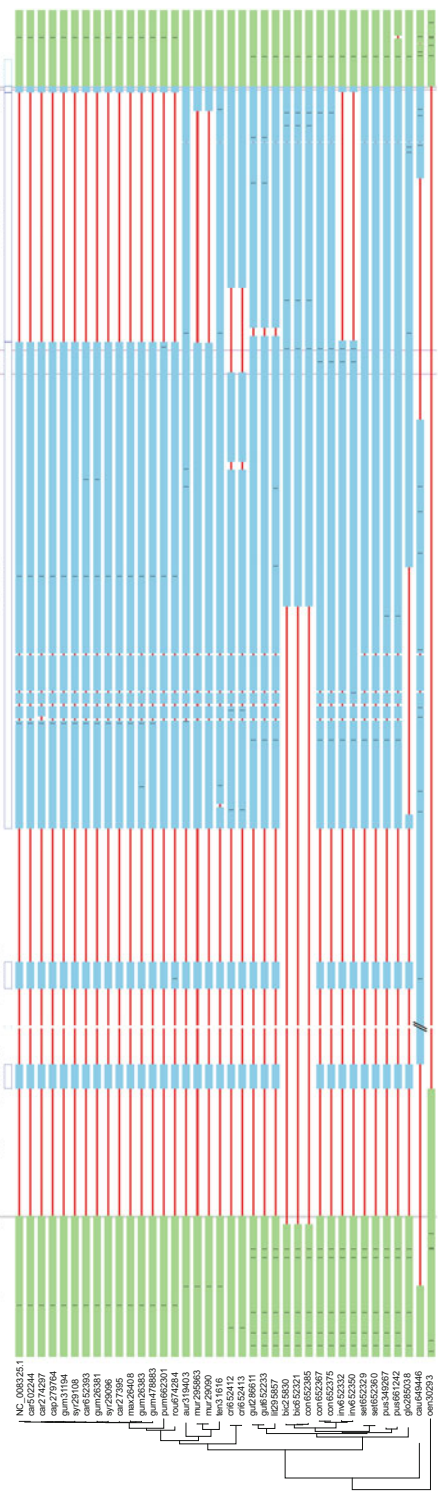

Fig. 12.4 Phylogenetic distribution and sequence comparison of plastid sequences spanning the mitochondrial-to-plastid (mt-to-pt) insertion designated as DcMP across all species included in Spooner et al. (2017). Green segments represent plastid sequence, and blue segments represent sequence of mitochondrial origin. The green region "F" designates conserved plastid sequences flanking the mt-to-pt insertion. The green region "A" designates a 339-nt region containing the ancestral promoter P4 and P5 (Tohdoh et al. 1981). DcMP1-2-3-4 (blue) designates the regions spanning the original mt-to-pt insertion described in Iorizzo et al. (2012a). CpMP5 and CpMP6 denote the two large insertions (6663 and 360 nt) identified Spooner et al. (2017) in *C. platycarpos*; trnV (green) represents the region coding for the trnV-GAC gene in the carrot plastid genome; P1, P2, and P3 indicate the location of the three putative promoters of the *D. carota* trnV (Manna et al. 1994). The vertical gray lines indicate the location of the 6-nt direct repeat flanking the DcMP insertion in *D. carota* and described in Iorizzo et al. (2012b). The double slash designates the masked portion of the 6663-nt insertion identified in *C. platycarpos*; this portion of DcMP5 insertion was masked to fit the figure in one panel. Vertical black lines indicate single-nucleotide polymorphisms (SNPs). Red lines indicate deletions identified based on the sequence alignment

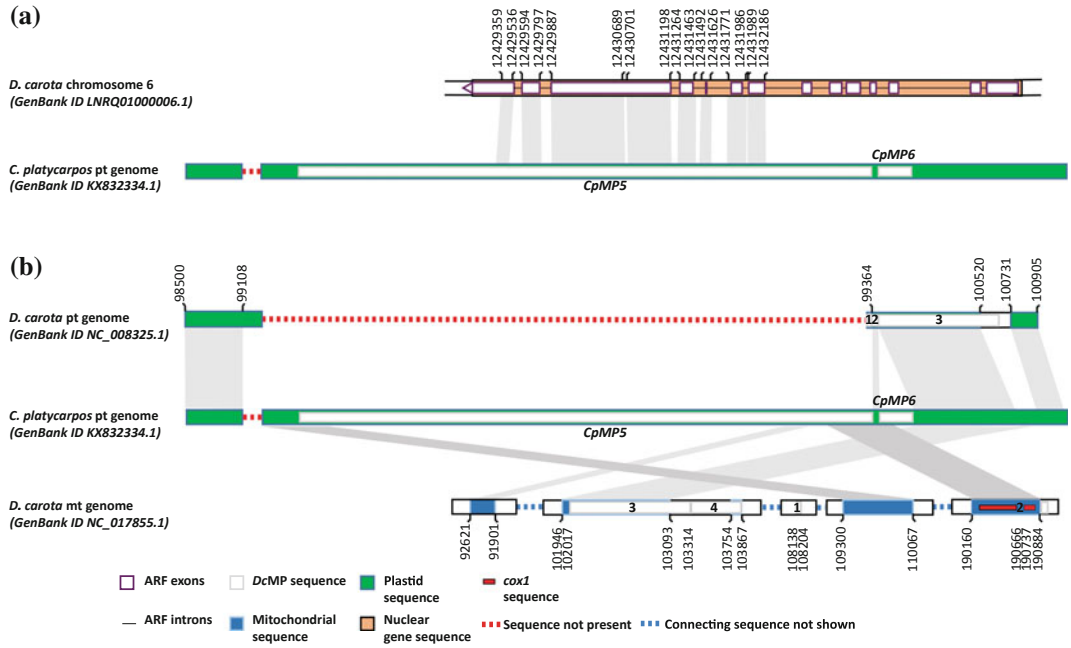

Fig. 12.5 DcMP comparative analysis. **a** Comparison between the *Daucus carota* nuclear genome region containing auxin response factor (ARF) gene DCAR_022437 in the antisense orientation, and *Caucalis platycarpos* plastid sequence spanning the DcMP region. Gray shading linking sequences indicate regions with >92% nucleotide similarity. **b** Comparison between the *C. platycarpos* plastid sequence spanning the DcMP region and *D. carota* plastid and mitochondrial genomes. Red dashed lines indicate deletions of the sequence in the corresponding genome. Mitochondrial sequences are not directly contiguous, which are represented by gaps and blue dashed lines. Regions labeled with single digits 1 through 4 correspond to DcMP regions 1 through 4

promoter plays a functional and advantageous role on the expression of trnV, the comparative studies suggest that natural selection has maintained its sequence intact promoting the retention of the ancestral DcMP sequence in the plastid genome after its first integration.

References

Adams KL, Qiu Y-L, Stoutemyer M, Palmer JD (2002) Punctuated evolution of mitochondrial gene content: high and variable rates of mitochondrial gene loss and transfer to the nucleus during angiosperm evolution. Proc Natl Acad Sci USA 99:9905–9912

Alvarado JS, López DH, Torres IM, Meléndez MM, Batista RA, Raxwal VK, Berríos Juan AN, Arun A (2017) Sequencing and de novo assembly of the complete chloroplast genome of the Peruvian carrot (*Arracacia xanthorrhiza* Bancroft). Genome Announc 5(7):e01519

Alverson AJ, Wei X, Rice DW, Stern DB, Barry K, Palmer JD (2010) Insights into the evolution of mitochondrial genome size from complete sequences of *Citrullus lanatus* and *Cucurbita pepo* (Cucurbitaceae). Mol Biol Evol 27:1436–1448

Alverson AJ, Zhuo S, Rice DW, Sloan DB, Palmer JD (2011) The mitochondrial genome of the legume *Vigna radiata* and the analysis of recombination across short mitochondrial repeats. PLoS ONE 6: e16404

Arrieta-Montiel MP, Shedge V, Davila J, Christensen AC, Mackenzie SA (2009) Diversity of the Arabidopsis mitochondrial genome occurs via nuclear-controlled recombination activity. Genetics 183:1261–1268

Bach IC, Olesen A, Simon PW (2002) PCR-based markers to differentiate the mitochondrial genomes of petaloid and male fertile carrot (*Daucus carota* L.). Euphytica 127:353–365

Banasiak Ł, Wojewódzka A, Baczyński J-P, Reduron M, Piwczński M, Kurzyna-Młynik R, Gutaker R, Czarnocka-Cieciura A, Kosmala-Grzechnik S, Spalik K (2016) Phylogeny of Apiaceae subtribe Daucinae and the taxonomic delineation of its genera. Taxon 65:563–585

Bergthorsson U, Adams KL, Thomason B, Palmer JD (2003) Widespread horizontal transfer of mitochondrial genes in flowering plants. Nature 424:197–201

Bock R (2015) Engineering plastid genomes: methods, tools, and applications in basic research and biotechnology. Annu Rev Plant Biol 66:211–241

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. BMC Bioinform 10:421

Chang S, Yang T, Du T, Huang Y, Chen J, Yan J, He J, Guan R (2011) Mitochondrial genome sequencing helps show the evolutionary mechanism of mitochondrial genome formation in *Brassica*. BMC Genom 12:497

Chen Q, Feng X, Li M, Yang B, Gao C, Zhang L, Tian J (2016) The complete chloroplast genome sequence of *Fatsia japonica* (Apiales: Araliaceae) and the phylogenetic analysis. Mitochondr DNA Part A 27:3050–3051

Choi SA, Kim Y, Kim K-Y, Kim JH, Seong RS (2016a) The complete chloroplast genome sequence of the medicinal plant, *Angelica gigas* (Apiaceae). Mitochondr DNA Part B 1:280–281

Choi SA, Kim YJ, Lee WK, Kim KY, Kim JH, Seong RS (2016b) The complete chloroplast genome of the medicinal plant *Angelica decursiva* (Apiaceae) in Peucedani Radix. Mitochondr DNA Part B 1:210–211

Choi SA, Lee WK, Kim Y, Kim KY, Kim JH, Seong RS (2016c) The complete chloroplast genome sequence of *Ostericum koreanum* (Apiaceae). Mitochondr DNA Part B 1:252–253

Chumley TW, Palmer JD, Mower JP, Fourcade HM, Calie PJ, Boore JL, Jansen RK (2006) The complete chloroplast genome sequence of *Pelargonium* × *hortorum*: organization and evolution of the largest and most highly rearranged chloroplast genome of land plants. Mol Biol Evol 23:2175–2190

Corriveau JL, Coleman AW (1988) Rapid screening method to detect potential biparental inheritance of plastid DNA and results for over 200 angiosperm species. Am J Bot 75:1443–1458

Davila JI, Arrieta-Montiel MP, Wamboldt Y, Cao J, Hagmann J, Shedge V, Xu Y-Z, Weigel D, Mackenzie SA (2011) Double-strand break repair processes drive evolution of the mitochondrial genome in Arabidopsis. BMC Biol 9:64

Deng Y-Q, Wen J, Yu Y, He X-J (2017) The complete chloroplast genome of *Angelica nitida*. Mitochondr DNA Part B 2:694–695

Dong W, Liu H, Xu C, Zuo Y, Chen Z, Zhou S (2014) A chloroplast genomic strategy for designing taxon specific DNA mini-barcodes: a case study on ginsengs. BMC Genet 15:138

Downie SR, Jansen RK (2015) A comparative analysis of whole plastid genomes from the Apiales: expansion and contraction of the inverted repeat, mitochondrial to plastid transfer of DNA, and identification of highly divergent noncoding regions. Syst Bot 40:336–351

Dufay M, Touzet P, Maurice S, Cuguen J (2007) Modelling the maintenance of male-fertile cytoplasm in a gynodioecious population. Heredity (Edinb) 99:349–356

Ge L, Shen L, Chen Q, Li X, Zhang L (2017) The complete chloroplast genome sequence of *Hydrocotyle sibthorpioides* (Apiales: Araliaceae). Mitochondr DNA Part B 2:29–30

Goremykin VV, Salamini F, Velasco R, Viola R (2009) Mitochondrial DNA of *Vitis vinifera* and the issue of rampant horizontal gene transfer. Mol Biol Evol 26:99–110

Gualberto JM, Mileshina D, Wallet C, Niazi AK, Weber-Lotfi F, Dietrich A (2014) The plant mitochondrial genome: dynamics and maintenance. Biochimie 100:107–120

Guo W, Zhu A, Fan W, Mower JP (2017) Complete mitochondrial genomes from the ferns *Ophioglossum californicum* and *Psilotum nudum* are highly repetitive with the largest organellar introns. New Phytol 213:391–403

Gupta R, Ryzhikov M, Koroleva O, Unciuleac M, Shuman S, Korolev S, Glickman MS (2013) A dual role for mycobacterial RecO in RecA-dependent homologous recombination and RecA-independent single-strand annealing. Nucleic Acids Res 41:2284–2295

Han Z-J, Li W, Liu Y, Gao L-Z (2016) The complete chloroplast genome of North American ginseng, *Panax quinquefolius*. Mitochondr DNA Part A 27:3496–3497

Iorizzo M, Grzebelus D, Senalik D, Szklarczyk M, Spooner D, Simon P (2012a) Against the traffic: the first evidence for mitochondrial DNA transfer into the plastid genome. Mobile Genet Elem 2:261–266

Iorizzo M, Senalik D, Szklarczyk M, Grzebelus D, Spooner D, Simon P (2012b) De novo assembly of the carrot mitochondrial genome using next generation sequencing of whole genomic DNA provides first evidence of DNA transfer into an angiosperm plastid genome. BMC Plant Biol 12:61

Jansen RK, Ruhlman TA (2012) Plastid genomes of seed plants. In: Bock R, Knoop V (eds) Genomics of chloroplasts and mitochondria. Springer, Netherlands, Dordrecht, pp 103–126

Jansen RK, Cai Z, Raubeson LA, Daniell H, dePamphilis CW, Leebens-Mack J, Müller KF, Guisinger-Bellian M, Haberle RC, Hansen AK, Chumley TW, Lee S-B, Peery R, McNeal JR, Kuehl JV, Boore JL (2007) Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. Proc Natl Acad Sci USA 104:19369–19374

Judd WS, Campbell CS, Kellogg EA, Stevens PF, Donoghue MJ (2016) Plant systematics: a phylogenetic approach, 4th edn. Sinauer Associates, Sunderland

Kim K-J, Lee H-L (2004) Complete chloroplast genome sequences from Korean ginseng (*Panax schinseng*

Nees) and comparative analysis of sequence evolution among 17 vascular plants. DNA Res 11:247–261
Kim K, Lee J, Lee S-C, Kim N-H, Jang W, Kim S, Sung S, Lee J, Yang T-J (2016a) The complete chloroplast genome of *Eleutherococcus gracilistylus* (W.W.Sm.) S.Y.Hu (Araliaceae). Mitochondr DNA Part A 27:3741–3742
Kim K, Lee S-C, Lee J, Kim N-H, Jang W, Yang T-J (2016b) The complete chloroplast genome sequence of *Panax quinquefolius* (L.). Mitochondr DNA Part A 27:3033–3034
Kim K, Nguyen VB, Dong J, Wang Y, Park JY, Lee S-C, Yang T-J (2017) Evolution of the Araliaceae family inferred from complete chloroplast genomes and 45S nrDNAs of 10 *Panax*-related species. Sci Rep 7:4917
Knoop V (2004) The mitochondrial DNA of land plants: peculiarities in phylogenetic perspective. Curr Genet 46:123–139
Krzywinski M, Schein J, Birol İ, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA (2009) Circos: an information aesthetic for comparative genomics. Genome Res 19:1639–1645
Kubo T, Newton KJ (2008) Angiosperm mitochondrial genomes and mutations. Mitochondrion 8:5–14
Lee BY, Downie SR (2000) Phylogenetic analysis of cpDNA restriction sites and rps16 intron sequences reveals relationships among Apiaceae tribes Caucalideae, Scandiceae and related taxa. Pl Syst Evol 221:35–60
Lee B-Y, Levin GA, Downie SR (2001) Relationships within the spiny-fruited Umbellifers (Scandiceae subtribes Daucinae and Torilidinae) as assessed by phylogenetic analysis of morphological characters. Syst Bot 26:622–642
Lee HO, Kim K, Lee S-C, Lee J, Lee J, Kim S, Yang T-J (2016a) The complete chloroplast genome sequence of *Ledebouriella seseloides* (Hoffm.) H. Wolff. Mitochondr DNA Part A 27:3498–3499
Lee S-C, Oh Lee H, Kim K, Kim S, Yang T-J (2016b) The complete chloroplast genome sequence of the medicinal plant *Glehnia littoralis* F.Schmidt ex Miq. (Apiaceae). Mitochondr DNA Part A 27:3674–3675
Li R, Ma P-F, Wen J, Yi T-S (2013) Complete sequencing of five Araliaceae chloroplast genomes and the phylogenetic implications. PLoS ONE 8:e78568
Lilly JW, Havey MJ, Jackson SA, Jiang J (2001) Cytogenomic analyses reveal the structural plasticity of the chloroplast genome in higher plants. Plant Cell 13:245–254
Ma PF, Zhang YX, Guo ZH, Li DZ (2015) Evidence for horizontal transfer of mitochondrial DNA to the plastid genome in a bamboo genus. Sci Rep 5:11608
Manna F, Massardo DR, Wolf K, Luccarini G, Carlomagno MS, Rivellini F, Alifano P, Del Giudice L (1994) A tRNA gene mapping within the chloroplast rDNA cluster is differentially expressed during the development of *Daucus carota*. Nucleic Acids Res 22:1712–1718
Manzanilla V, Kool A, Nguyen Nhat L, Van Nong H, Le Thi ThuH, de Boer HJ (2018) Phylogenomics and barcoding of *Panax*: toward the identification of ginseng species. BMC Evol Biol 18:44
Mower JP, Sloan DB, Alverson AJ (2012) Plant mitochondrial genome diversity: the genomics revolution. In: Wendel JF, Greilhuber J, Dolezel J, Leitch IJ (eds) Plant genome diversity, vol 1. Plant genomes, their residents, and their evolutionary dynamics. Springer, Vienna, pp 123–144
Ohtani K, Yamamoto H, Akimitsu K (2002) Sensitivity to *Alternaria alternata* toxin in citrus because of altered mitochondrial RNA processing. Proc Natl Acad Sci USA 99:2439–2444
Palmer JD (1985) Comparative organization of chloroplast genomes. Ann Rev Genet 19:325–354
Park S, Ruhlman TA, Sabir JS, Mutwakil MH, Baeshen MN, Sabir MJ, Baeshen NA, Jansen RK (2014) Complete sequences of organelle genomes from the medicinal plant *Rhazya stricta* (Apocynaceae) and contrasting patterns of mitochondrial genome evolution across asterids. BMC Genom 15:405
Qiu Y, Filipenko SJ, Darracq A, Adams KL (2014) Expression of a transferred nuclear gene in a mitochondrial genome. Curr Plant Biol 1:68–72
Rabah SO, Lee C, Hajrah NH, Makki RM, Alharby HF, Alhebshi AM, Sabir JSM, Jansen RK, Ruhlman TA (2017) Plastome sequencing of ten nonmodel crop species uncovers a large insertion of mitochondrial DNA in cashew. Plant Genome. https://doi.org/10.3835/plantgenome2017.03.0020
Raubeson LA, Jansen RK (2005) Chloroplast genomes of plants. In: Henry RJ (ed) Plant diversity and evolution: phenotypic variation in higher plants. CABI Publishing, Wallingford, UK, pp 45–68
Richardson AO, Palmer JD (2007) Horizontal gene transfer in plants. J Exp Bot 58:1–9
Robison MM, Wolyn DJ (2002) Complex organization of the mitochondrial genome of petaloid CMS carrot. Mol Genet Genom 268:232–239
Rodríguez-Moreno L, González VM, Benjak A, Martí MC, Puigdomènech P, Aranda MA, Garcia-Mas J (2011) Determination of the melon chloroplast and mitochondrial genome sequences reveals that the largest reported mitochondrial genome in plants contains a significant amount of DNA having a nuclear origin. BMC Genom 12:424
Ross MG, Russ C, Costello M, Hollinger A, Lennon NJ, Hegarty R, Nusbaum C, Jaffe DB (2013) Characterizing and measuring bias in sequence data. Genome Biol 14:R51
Ruhlman T, Lee S-B, Jansen RK, Hostetler JB, Tallon LJ, Town CD, Daniell H (2006) Complete plastid genome sequence of *Daucus carota*: implications for biotechnology and phylogeny of angiosperms. BMC Genom 7:222
Samigullin TH, Logacheva MD, Terenteva EI, Degtjareva GV, Vallejo-Roman CM (2016) Plastid genome of *Seseli montanum*: complete sequence and comparison with plastomes of other members of the Apiaceae family. Biochem (Moscow) 81:981–985

Samigullin TH, Logacheva MD, Degtjareva GV, Terentieva EI, Vallejo-Roman CM (2017) Complete plastid genome of critically endangered plant *Prangos trifida* (Apiaceae: Apioideae). Conserv Genet Res

Shin D-H, Lee J-H, Kang S-H, Ahn B-O, Kim C-K (2016) The complete chloroplast genome of the Hare's Ear root *Bupleurum falcatum*: its molecular features. Genes 7:20

Sloan DB, Alverson AJ, Chuckalovcak JP, Wu M, McCauley DE, Palmer JD, Taylor DR (2012) Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. PLoS Biol 10:e1001241

Smith DR (2011) Extending the limited transfer window hypothesis to inter-organelle DNA migration. Genome Biol Evol 3:743–748

Soucy SM, Huang J, Gogarten JP (2015) Horizontal gene transfer: building the web of life. Nat Rev Genet 16:472–482

Spooner DM, Ruess H, Iorizzo M, Senalik D, Simon P (2017) Entire plastid phylogeny of the carrot genus (*Daucus*, Apiaceae): concordance with nuclear data and mitochondrial and nuclear DNA insertions to the plastid. Am J Bot 104:296–312

Straub SC, Cronn RC, Edwards C, Fishbein M, Liston A (2013) Horizontal transfer of DNA from the mitochondrial to the plastid genome and its subsequent evolution in milkweeds (Apocynaceae). Genome Biol Evol 5:1872–1885

Tan J, Yu Y (2018) The complete chloroplast genome of *Pimpinella rhomboidea* var. *tenuiloba*. Mitochondr DNA Part B 3:101–102

Tatum EL, Lederberg J (1947) Gene recombination in the bacterium *Escherichia coli*. J Bacteriol 53:673–684

Tilney-Bassett RAE (1978) The inheritance and genetic behavior of plastids. In: Kirk JTO, Tilney-Bassett RAE (eds) The plastids. Elsevier/North Holland, Amsterdam, pp 251–324

Timmis JN, Ayliffe MA, Huang CY, Martin W (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat Rev Genet 5:123–135

Tohdoh N, Shinozaki K, Sugiura M (1981) Sequence of a putative promoter region for the rRNA genes of tobacco chloroplast DNA. Nucleic Acids Res 9:5399–5406

Vetsigian K, Woese C, Goldenfeld N (2006) Collective evolution and the genetic code. Proc Natl Acad Sci USA 103:10696–10701

Vivek BS, Ngo QA, Simon PW (1999) Evidence for maternal inheritance of the chloroplast genome in cultivated carrot (*Daucus* carota L. ssp. *sativus*). Theor Appl Genet 98:669–672

Wang L, Du X-J, Li X-F (2016) The complete chloroplast genome sequence of the evergreen plant *Dendropanax dentiger* (Araliaceae). Mitochondr DNA Part A 27:4193–4194

Wolfe KH, Li WH, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc Natl Acad Sci USA 84:9054–9058

Wu Y, Zhang T-Z, Qiu D-Y, Chai Q, Fan W-B, Li Z-H, Fang M-F (2017) Complete plastid genome of *Bupleurum boissieuanum*, an endemic herb plant in western China. Conserv Genet Resour. https://doi.org/10.1007/s12686-017-0890-2

Yang J, Yue M, Niu C, Ma X-F, Li Z-H (2017) Comparative analysis of the complete chloroplast genome of four endangered herbals of *Notopterygium*. Genes 8:124

Yi D-K, Lee H-L, Sun B-Y, Chung MY, Kim K-J (2012) The complete chloroplast DNA sequence of *Eleutherococcus senticosus* (Araliaceae); comparative evolutionary analyses with other three asterids. Mol Cells 33:497–508

Yuan C, Zhong W, Mou F, Gong Y, Pu D, Ji P, Huang H, Yang Z Zhang C (2017) The complete chloroplast genome sequence and phylogenetic analysis of *Chuanminshen* (*Chuanminshen* violaceum Sheh et Shan). Physiol Mol Biol Plants 23:35–41

Zhang D, Li W, Gao C, Liu Y, Gao L (2016) The complete plastid genome sequence of *Panax notoginseng*, a famous traditional Chinese medicinal plant of the family Araliaceae. Mitochondr DNA Part A 27:3438–3439

Zhang Y-J, Gao H, Chen Y, Guo F-X, Bai G, Guo Y-Y, Yan F, Wang E-J (2018) Characterization of the complete plastid genome sequence of *Eleutherococcus brachypus* (Araliaceae), an endangered shrub in China. Conserv Genet Resour. https://doi.org/10.1007/s12686-018-1012-5

Zhao Y, Yin J, Guo H, Zhang Y, Xiao W, Sun C, Wu J, Qu X, Yu J, Wang X, Xiao J (2015) The complete chloroplast genome provides insight into the evolution and polymorphism of *Panax ginseng*. Frontiers Plant Sci 5:696

Zong X, Song J, Lv J, Wang S (2016) The complete chloroplast genome sequence of *Schefflera octophylla*. Mitochondr DNA Part A 27:4685–4686