Theor Appl Genet (2005) 111: 1430–1439
DOI 10.1007/s00122-005-0075-3

ORIGINAL PAPER

H. Katayama · C. Uematsu

Structural analysis of chloroplast DNA in *Prunus* (Rosaceae): evolution, genetic diversity and unequal mutations

Received: 15 October 2004 / Accepted: 3 August 2005 / Published online: 2 September 2005
© Springer-Verlag 2005

Abstract In order to understand the evolutionary aspects of the chloroplast DNA (cpDNA) structures in Rosaceous plants, a physical map of peach (*Prunus persica* cv. Hakuhou) cpDNA was constructed. Fourteen lambda phage clones which covered the entire sequence of the peach cpDNA were digested by restriction enzymes (*Sal*I, *Xho*I, *Bam*HI, *Sac*I, and *Pst*I) used singly or in combination. The molecular size of peach cpDNA was estimated to be about 152 kb. The gene order and contents were revealed to be equivalent to those of standard type of angiosperms by the localization of 31 genes onthe physical map. Eighteen accessions from 14 *Prunus* species (*P. persica*, *P. mira*, *P. davidiana*, *P. cerasis, P. cerasifera*, *P. domestica*, *P. insititia*, *P. spinosa*, *P. salicina*, *P. maritima*, *P. armeniaca*, *P. mume*, *P. tomentosa*, *P. zippeliana*, and *P. salicifolia*) and one interspecific hybrid were used for the structural analysis of cpDNAs. Seventeen mutations (16 recognition site changes and one length mutation) were found in the cpDNA of these 18 accessions by RFLP analysis allowing a classification into 11 genome types. Although the base substitution rate in the recognition site ($100p$ = 0.72) of cpDNA in *Prunus* was similar to that of other plants, i.e., *Triticum–Aegilops*, *Brassi ca*, and *Pisum*, it differed from *Pyrus* ($100p$ = 0.15) in Rosaceae. Seven mutations including one length mutation were densely located within a region of about 9.1 kb which includes *psb*A and *atp*A in the left border of a large single-copy region of *Prunus* cpDNAs. The length mutation was detected only in *P. persica* and consisted of a 277 bp deletion which occurred in a spacer region between the *trn*S and *trn*G genes within the 9.1 kb region. Additional fragment length mutations (insertion/deletion), which were not detected by RFLP analysis, were revealed by PCR and sequence analyses in *P. zippeliana* and *P. salicifolia*. All of these length mutations occurred within the 9.1 kb region between *psb*A and *atp*A. This region could be an intra-molecular recombinational hotspot in *Prunus* species.

Communicated by H. Nybom

H. Katayama (✉)
Food Resources Education and Research Center,
Kobe University, 675-2103 Hyogo, Japan
E-mail: hkata@kobe-u.ac.jp
Fax: +81-790-490343

C. Uematsu
Botanical Gardens, Graduate School of Science,
Osaka City University, 576-0004 Osaka, Japan

Introduction

It is well known that the structure and gene content of plastomes are conserved among divergent plant species (Downie and Palmer 1992). The complete sequencing of chloroplast DNA (cpDNA) from various plants, such as tobacco and some cereals (Shinozaki et al. 1986; Hiratsuka et al. 1989; Ogihara et al. 2002), has confirmed a high conservatism of chloroplast genomes during the course of plant evolution. This conservatism and uniparental heritance make it possible to compare phylogenetic relationships at various taxonomic levels (Palmer 1991). Although cpDNAs are strongly conserved across plant species, structural alterations such as inversions, translocations, insertions, and deletions have been found in certain plant lineages by comparing the structure of cpDNAs. Therefore, tracing the mutational events in chloroplast genomes enables us to reconstruct plant phylogeny at higher taxonomic levels (Downie and Palmer 1992; Doyle et al. 1992; Katayama and Ogihara 1996).

Three hypervariable regions, which may represent intra-molecular recombinational hotspots, have also been detected in the chloroplast genomes of gramineous plants by comparing those sequences to that of tobacco. These hypervariable regions are unevenly distributed within the cpDNA molecule. The first hotspot region having various length deletions is located within inverted repeats containing the large open reading frame *ycf*2 of tobacco (Hiratsuka et al. 1989; Maier et al. 1995). The

second one having various length polymorphisms is reported for wheat, maize, and rice in the large single-copy region (LSC) between the genes *rbc*L and *cem*A (Ogihara et al. 1988; Morton and Clegg 1993; Katayama and Ogihara 1996). The third one containing three major inversions in cereals is found in the LSC between *trn*S (UGA) and *rps*14 (Hiratsuka et al. 1989). Mutational events, such as length mutations, occurred frequently at the interspecific level in these hypervariable regions.

Phylogenetic classification in Rosaceae is still controversial. Schulze-Menz (1964) divided it into four subfamilies, i.e., Spiraeoideae, Rosoideae, Maloideae, and Amygdaloideae mainly according to their fruit types and chromosome numbers. Rehder (1940) recognized four genera (*Prunus*, *Maddenia* Hook. F. & Thomson, *Oemleria* Reichb., and *Prinsepia* Royle) in Amygdaloideae.

The genus *Prunus* contains more than 200 species distributed mainly in the temperate zones of the Northern hemisphere (Rehder 1940; Sauer 1993). Economically important species such as *P. persica* (L.) Batsch (peach), *P. domestica* L. (European plum), *P. salicina* Lindl. (Japanese plum), *P. armeniaca* L. (apricot), *P. avium* L. (sweet cherry), *P. cerasus* L. (sour cherry), and *P. dulcis* (Mill.) D.A. Webb. (almond) are included in this genus. As one of the most familiar genera of fruit trees, *Prunus* has been thoroughly studied and defined by morphological characters such as five sepals, five petals, a solitary carpel with a terminal style, fleshy drupe, and a solid branched pith (Rehder 1940).

For the purpose of phylogenetic study, a portion of cpDNA up to about 20 kb in length was analyzed by means of RFLP, PCR-RFLP or sequencing in *Prunus* (Badenes and Parfitt 1995; Brettin et al. 2000; Mohanty et al. 2000; Mohanty and Pedro 2001). For Rosaceous plants, with the exception of pear (*Pyrus ussuriensis* var. *hondoensis*), a physical map covering the entire cpDNA has, however, not yet been constructed (Katayama and Uematsu 2003). Therefore a comparison of characteristic features for the whole chloroplast genome structure has still not been accomplished except for a few reports concerning structural analysis such as RFLPs covering the entire chloroplast genome in *Prunus* (Kaneko et al. 1986; Uematsu et al. 1991).

Some researchers placed *Exochorda* Lindley (previously placed in the Spiraeoideae by Rehder 1940) in the Amygdaloideae based on its chromosome number and molecular data (Goldblatt 1976; Morgan et al. 1994; Lee and Wen 2001). Recent DNA sequencing analyses using cpDNA (*trn*L-*trn*F, *mat*K) and nuclear DNA (ITS and GBSSI) indicated that *Prunus* is not so strongly associated with *Oemleria*, *Prinsepia*, and *Exochorda* (Bortiri et al. 2001; Evans and Campbell 2002; Potter et al. 2002). In spite of intensive studies in Rosaceae, the phylogenetic position of *Prunus* in Amygdaloideae and the status of Amygdaloideae in Rosaceae are still uncertain and additional approaches will be required (Campbell et al. 1995; Potter et al. 2002).

We tried to determine structural alterations, which could be used to trace the evolution of the chloroplast genome in *Prunus*. Physical maps, genetic diversity and unequal mutations such as hotspot regions in *Prunus* cpDNAs are presented in this study.

Materials and methods

Peach (*P. persica* cv. Hakuhou) maintained at the Hyogo Prefectural Technology Center for Agriculture (HPTCA) was used for the construction of a physical

Table 1 The species of *Prunus* examined in the present study. Classification scheme follows Rehder (1940) and Mason (1913)

Code	Taxon	Species	Source	Accession	Name of cultivar or common name	Chloroplast genome type
1	Amygdalus	*Prunus persica* (L.) Batsch	HPTCA		Peach 'Hakuhou'	A
2	Amygdalus	*P. persica* (L.) Batsch	FRC-KU	2B0005K	Nectarine 'Reimei'	A
3	Amygdalus	*P. mira* Koehne	OCU*	4A0246	Koukakuto	B
4	Amygdalus	*P. davidiana* (Carriere) Franch.	OCU*	4A0248	David peach	C
5	Prunus	*P. cerasifera* Ehrh.	OCU	4A0340	Myrobalan plum	D
6	Prunus	*P. domestica* L.	OCU	4A0209	Stanley	D
7	Prunus	*P. insititia* L.	OCU*	4A0265	Krikon damson	D
8	Prunus	*P. spinosa* L.	OCU*	4A0264	Spinosa plum	E
9	Prunus	*P. salicina* Lindl.	OCU	4A0208	Santarosa	F
10	Prunus	*P. salicina* Lindl.	OCU	4A0205	Ohishiwase	F
11	Prunus	*P. maritima* Marshall	OCU*	4A0259	Beach plum	G
12	Prunus	*P. armeniaca* L.	OCU*	4A0255	Early orange	F
13	Prunus	*P. armeniaca* L.	OCU	4A0198	Heiwa	F
14	Prunus	*P. armeniaca* x *P. salicina*	OCU*	4A0256	Jinkyo	F
15	Prunus	*P. mume* Siebold & Zucc.	OCU	4A0051	Koshukoume	H
16	Cerasus	*P. tomentosa* Thunb.	OCU		Yusuraume	I
17	Laurocerasus	*P. zippeliana* Miquel	OCU	1E0541	Bakuchinoki	J
18	Padus	*P. salicifolia* Kunth	OCU	4A0247	Capulin cherry	K
19	Pyrus	*Pyrus ussuriensis* var. *hondoensis*	OCU	4A0134	Pear 'Aonashi'	–

HPTCA, Hyogo Prefectural Technology Center for Agriculture, Forestry and Fisheries; FRC-KU, Food research center, Kobe University; OCU, Osaka City University; OCU*, introduced from the National Institute of Fruit Tree Science (NIFTS) to Osaka City University in 1994

map of cpDNA. Eighteen accessions from 14 *Prunus* species and one interspecific hybrid originating from Asia, Europe, and America were used in this study (Table 1). These are maintained at the Botanical Gardens, Osaka City University (OCU). Six *Prunus* species and one hybrid indicated in Table 1 were introduced from the National Institute of Fruit Tree Science (NIFTS) to the Botanical Gardens of OCU in 1994 by grafting. Other *Prunus* species were from the original OCU collection or purchased from a nursery in 1994. Classification of these *Prunus* species followed Mason (1913) and Rehder (1940). Pear (*Pyrus ussuriensis* var. *hondoensis*) maintained at the Botanical Gardens of OCU was used as a representative of an outgroup.

Chloroplast DNA was isolated from green leaves of peach according to the method reported by Uematsu et al. (1991). Peach cpDNA was digested with *Sal*I, *Xho*I, *Bam*HI, *Sac*I, and *Pst*I alone or by combinations of two enzymes (Roche Diagnostics GmbH, Germany). The digested cpDNAs were fractionated by 0.7% agarose gel electrophoresis. Total DNAs for RFLP analysis and PCR were extracted from frozen leaves of *Prunus* according to Reiter et al. (1992) with a slight modification. A two times stock of CTAB extraction buffer was used and therefore three volumes of precipitation buffer were added in order to obtain DNA whilst avoiding co-precipitation of polysaccharide.

A peach cpDNA library was constructed according to the method of Katayama and Uematsu (2003). CpDNAs partially digested with *Sau3A*I were cloned into lambda FIX II vector (Stratagene, USA). After packaging using Gigapack Gold (Stratagene, USA), recombinant phages were selected by plaque hybridization with tobacco cpDNA clones (Shinozaki et al. 1986). Selected clones were screened by homologous hybridization to obtain overlapping clones which covered the entire region of peach cpDNA. Plaque and homologous hybridizations were carried out using ECL non-radioactive DNA labeling and detection kits (Amersham-Pharmacia Biotech, UK). Screened clones were connected to each other to generate a physical map of the entire peach cpDNA.

In order to determine the localization of genes on the physical map of peach cpDNA, 31 genes were prepared from tobacco cpDNA clones by PCR amplification or digestion using appropriate restriction enzymes. For PCR amplification, oligonucleotide primers were synthesized according to the published sequence data (Table 2). *rpl*20, *pet*B, *rpl*14, and *atp*A genes were isolated by digestion of tobacco cpDNA clones with appropriate restriction enzymes (Shinozaki et al. 1986; Katayama and Ogihara 1993).

The spacer region between *trn*S and *trn*G in the LSC region was amplified using the following primers; *trn*S-F: GTACGATTAACTCGTACAACGG and *trn*G-5′R: TTTACCACTAAACTATACCCG. Amplified fragments were fractionated by agarose gel electrophoresis. Sequencing was performed with an ABI310 Genetic Analyzer (Applied Biosystems, USA).

Table 2 List of the genes mapped on the physical map of peach cpDNA and the primer-pair forward/reverse sequences designed for PCR

Gene	Forward primer	Reverse primer
*psb*A	5′-AGCTTCGATAGCAGGTAGGTCT-3′	5′-ATGACTGCAATTTTAGAGA-3′
*mat*K	5′-TGATACATAGTGCGATACAG-3′	5′-GCTAACTCAACGGTAGAGTAC-3′
*rps*16	5′-TAGGATTGATTAGGACGAAGTT-3′	5′-ACTTCGTTTGAAACGATGTGGT-3′
*rps*2	5′-TTTGGATTCCTATCTAATGAT-3′	5′-TTTACCACTAAACTATACCC-3′
*rpo*C1	5′-ACCTCGTAATCGTAAGAAAGT-3′	5′-TCGGATACGAAGGTATCAAAT-3′
*rpo*B	5′-TCAGGATAGCAAACATTCTCTA-3′	5′-TGGAAATGAGGGAATAT-3′
*psb*DC	5′-ATGACTATAGCCC TTGGTA-3′	5′-TTAAGATTGGGGTCATG-3′
*psa*A	5′-TTATCCTACTGCAATAA TTCTT-3′	5′-ATTTCTCATAGTTGGTGCTGC-3′
*rps*4	5′-CCTTTATATTGGAATGGGTGAAG-3′	5′-GTCGCGTTACCGAGGACCTCG-3′
*trn*L	5′-CCTTTATATTGGAATGGGTGAAG-3′	5′-GTCGCGTTACCGAGGACCTCG-3′
*psb*G	5′-TAATTCGGGGGAAGATACTGAAT-3′	5′-AATGAGTTTCGACGTATTGGGT-3′
*atp*BE	5′-AGCTTCTTGTGGATCAATGTC-3′	5′-CGTCCAAGCAGGATCCGAAG-3′
*rbc*L	5′-ATGTCACCACAAACTACTCC-3′	5′-TACCTTCCTGAGCAAGATCA-3′
*acc*D	5′- TCAATGGGCAGTCTTGGTC C-3′	5′-GCCGTTTAGTTGGCCTATGC-3′
*ycf*4	5′-CTGAGTGAGTTATT TA AGCTCCA-3′	5′-GACATCACGAGTCCTTGTG-3′
*cem*A	5′-CCACA AG GACTCGTGATGTC-3′	5′-GTATCTTGATTAATTGGATCG-3′
*pet*A	5′-AGATTACTCGATCTATTTCC-3′	5′-TAGAAATTCATTTCGGCCAATT-3′
*psb*B	5′-TGCCTTGGTATCGTGTTCATA-3′	5′-GCAATATGATGAGAGGCATA-3′
*rpl*23	5′-ATGGATGGAATCAAATATGCA-3′	5′-AATAGAATAACCCGGAAGAAG-3′
*ycf*2	5′-TTGATTTCTCGAACCGAG AG-3′	5′-ACCTTTTCGAACTCTGTCT-3′
*ndh*B	5′-AGGCTATCCTGAGCAATGCA-3′	5′-TACTTCGAAAGTAGCTGCTTCA-3′
16S	5′-CATGCTTAACACATGCAAGT-3′	5′-TACCGTACTCCAGCTTGGTAG-3′
23S	5′-CTAAGGTAGCGAAATTCCTTGT-3′	5′-CTTGGCTACCCAGCGTTTACC-3′
*ndh*F	5′-AGATAAGAAGAGATGCGTCC-3′	5′-TCATTCCACTTCCAGTCCCTAT-3′
*ndh*D	5′-TAGATAGAATAACTTCGACCTT-3′	5′-AGCTTTAACATCTCTTGGCCAA-3′
*ndh*A	5′-TCTTCAGCGGCTGCAATAGCT-3′	5′-ATTGACAACCTCGTCCCAACT-3′
*ycf*1	5′-TATTGATATGAAGATTGCCG-3′	5′-TGAAACCTTGGCATATATCT-3′

*atp*A , *rpl*20, *pet*B, and *rpl*14 genes were isolated by digestion of tobacco cpDNA clones (Katayama and Ogihara 1993)

The base substitution rate (100*p*) was calculated from restriction site changes in *Prunus* cpDNA after Brown et al. (1979).

The presence or absence of each restriction site was scored as binary data. An unweighted pair group mean cluster analysis (UPGMA) was performed (with the average distance option) with NTSYS-pc version 2.01 (Rohlf 1998) using a similarity matrix based on the shared restriction sites. A maximum parsimony analysis was done with PAUP (version 4.0: Swofford 1998) using a heuristic search with the option of random addition sequence (100 replications), TBR swapping algorithm, and MULPARS. Bootstrap analysis with 100 replications (heuristic; 100 random; TBR) was performed using PAUP.

Results

Cloning and physical map of peach cpDNA

The genome size of peach cpDNA was estimated to range from 150.5 kb for *Pst*I to 152.7 kb for *Sal*I by summation of fragments using five restriction enzymes; *Sal*I, *Xho*I, *Bam*HI, *Sac*I, and *Pst*I (Table 3). According to these estimates, the molecular size of peach cpDNA was considered to be approximately 152 kb. Fourteen overlapping clones, pPpcp-1 (12.0 kb), pPpcp-2 (12.7 kb), pPpcp-3 (15.4 kb), pPpcp-4 (14.0 kb), pPpcp-5 (15.1 kb), pPpcp-6 (13.8 kb), pPpcp-7 (13.1 kb), pPpcp-8 (14.4 kb), pPpcp-9 (18.0 kb), pPpcp-10 (12.8 kb), pPpcp-11 (16.0 kb), pPpcp-12 (14.0 kb), pPpcp-13 (14.0 kb), and pPpcp-14 (18.0 kb), covering the entire region of peach cpDNA were isolated by homologous hybridization (Fig. 1). Restriction sites were confirmed by digestion with *Sal*I, *Xho*I, *Bam*HI, *Sac*I, and *Pst*I singly or in combinations of two enzymes. Consequently a physical map using five restriction enzymes was constructed for peach. Thirty-one genes were located on it (Fig. 2).

Mutations revealed by RFLP analysis and cpDNA types

One fragment length mutation and 16 restriction site changes were detected in 18 accessions by RFLP analysis in the present study (Table 4). These were aligned on the physical map of peach cpDNA after identification of all fragments involved in these mutations by Southern hybridization using the corresponding peach clones as a probe. The positions of the insertion and the mutated restriction sites found in each cpDNA type are presented in Fig. 3. Based on these results, the 18 accessions of *Prunus* could be classified into 11 cpDNA types (Types

Table 3 Molecular size of restriction fragments of peach cpDNA generated with *Sal*I, *Xho*I, *Sac*I, *Bam*HI and *Pst*I

*Sac*I		*Xho*I		*Sac*I		*Bam*HI		*Pst*I	
Fragments	kb	Fragments	kb	Fragments	kb	Fragments	kb	Fragments	kb
Sl-1	43.7	X-1	24.8	Sc-1	27.8	B-1	14.0	P-1	41.8
Sl-2	22.0	X-2	14.0	Sc-2	25.4	B-2	10.3	P-2	21.6
Sl-3	20.2	X-3	12.5(2)	Sc-3	20.4	B-3	9.4	P-3	20.5
Sl-4	17.0	X-4	11.5(2)	Sc-4	19.6	B-4	8.0	P-4	16.7
Sl-5	16.3	X-5	10.7(2)	Sc-5	16.3	B-5	7.6	P-5	16.0
Sl-6	13.5	X-6	9.5	Sc-6	12.0	B-6	6.5 (3)	P-6	14.5
Sl-7	11.0	X-7	7.2	Sc-7	7.5	B-7	6.1	P-7	12.7
Sl-8	2.3	X-8	3.7	Sc-8	6.0	B-8	5.8	P-8	3.1
Sl-9	2.2 (2)	X-9	3.5 (3)	Sc-9	3.7	B-9	4.6	P-9	2.8
Sl-10	0.9 (2)	X-10	3.2 (2)	Sc-10	3.4 (2)	B-10	4.0	P-10	0.8
Sl-11	0.5	X-11	3.1	Sc-11	2.0 (2)	B-11	3.9		
		X-12	1.2	Sc-12	1.8	B-12	3.7 (2)		
		X-13	0.8 (2)			B-13	3.3		
		X-14				B-14	2.8 (3)		
		X-15				B-15	2.7 (2)		
						B-16	2.4		
						B-17	2.3		
						B-18	2.2		
						B-19	2.0 (2)		
						B-20	1.9		
						B-21	1.8 (3)		
						B-22	1.6		
						B-23	1.2 (2)		
						B-24	1.1 (5)		
						B-25	0.8 (4)		
						B-26	0.6 (2)		
						B-27	0.5 (3)		
						B-28	0.4 (3)		
Total	152.7		151.4		151.3		152.5		150.5

Fragment copy number is given in parentheses

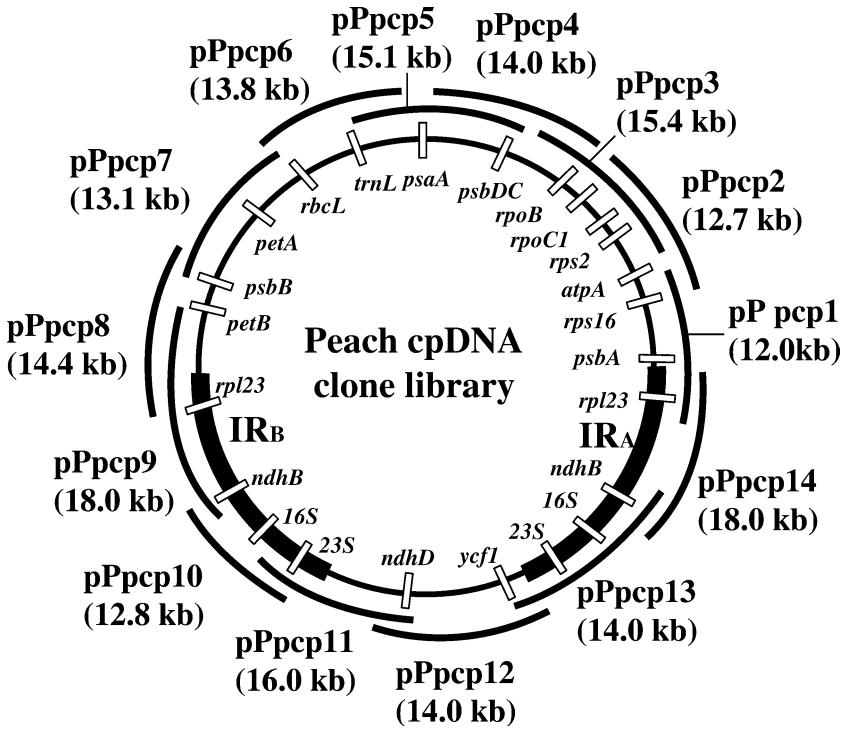

Fig. 1 Overlapping clones of peach cpDNA (*P. persica* cv. Hakuhou). The extent of the inverted repeats (IRA and IRB) and the position of the genes on the peach cpDNA are indicated *inside the circle*. The clone names and insert sizes of the 14 phage clones are drawn *outside the circle*

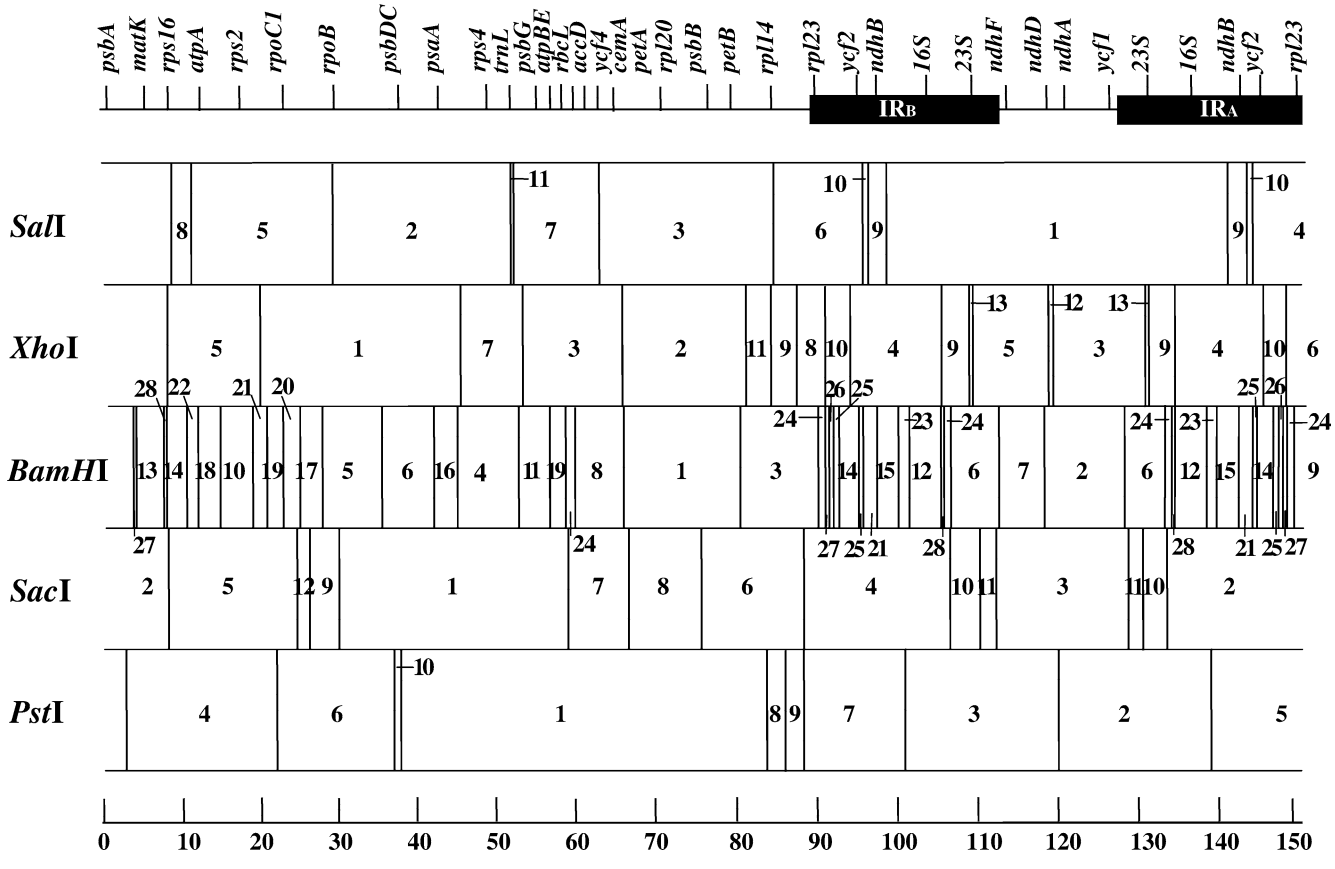

Fig. 2 Physical map of peach cpDNA in a linear form showing restriction sites of five restriction enzymes, *Sal*I, *Xho*I, *BamH*I, *Sac*I, and *Pst*I. IRA and IRB are the inverted repeat regions. The numbers given in the physical map correspond to those presented in Table 3. The location of 31 genes is indicated on the *upper line*

Table 4 Fragment size mutations and restriction site changes revealed in cpDNA of 18 *Prunus* species detected by five restriction enzymes

Mutation number[a]	Location	Size (kbp)	Carrier cytoplasms (Code number)
Insertion			
5	Sl8-B14[b]	0.3	All except 1, 2
Gain			
1	B9	4.6 → 3.2 + 1.4	8
2	B13	3.3 → 1.7 + 1.6	16, 17
3	B13	3.3 → 2.0 + 1.3	5, 6, 7, 8
6	B22	1.6 → 0.3 + 1.3	16
7	SC5	16.3 → 3.8 + 12.5	All except 1, 2, 18
9	SC9	3.7 → 0.9 + 2.8	15
10	B5	7.6 → 1.6 + 6.0	All except 1, 2, 17
11	B5	7.6 → 1.8 + 5.8	17
13	B3	9.4 → 2.0 + 7.4	17, 18
14	P8	3.1 → 1.1 + 2.0	16
15	X4	11.5 → 9.5 + 2.0	11
16	B7	6.1 → 4.8 + 1.3	4, 17, 18
Loss			
4	X6-X5	9.4 + 10.7 → 20.2	5, 6, 7, 8, 17
8	B10-B21	4.5 + 1.8 → 6.3	All except 1, 2
12	B5-B6	7.6 + 6.5 → 14.1	All except 1, 2, 3, 4
17	B7-B2	6.1 + 10.3 → 16.4	3, 17

[a]The mutation numbers given here correspond to those presented in Fig. 3. Insertion, site gains and site losses here are all with reference to *P. persca* cpDNA[b]Fragment designations in the location are given in Fig. 2Sl, *Sal*I; X, *Xho*I; B, *BamH*I; Sc, *Sac*I; P, *Pst*I

A–K; Fig. 3). Peach and nectarine, both *P. persica*, shared cpDNA type A. *Prunus cerasifera*, *P. domestica*, and *P. insititia* belonged to cpDNA type D. *P. salicina* (Japanese plum), *P. armeniaca* (European apricot 'Early orange' and Japanese apricot 'Heiwa') and one interspecific hybrid between *P. armeniaca* and *P. salicina* possessed cpDNA type F. The base substitution rate between each cpDNA type is given in Table 5. The average base substitution rate observed among the 11 cpDNA genotypes in *Prunus* was 0.72 substitutions per 100 base pairs.

Unequal distribution of mutational events

Based on the RFLP analysis with 16 restriction site changes in the LSC, SSC, and IR, seven mutations were localized in ca. 20.3 kb region of the LSC-1 including the *psb*A to *rps*2 gene (Table 6, Fig. 3). The mutation frequency for restriction sites in this region was higher than in other regions of *Prunus* cpDNA (Table 6). In particular, seven mutations including a length mutation were observed in a ca. 9.1 kb region between *psb*A and *atp*A (Fig. 3). Mutations were unevenly distributed in the cpDNA molecule of *Prunus*. A length mutation predicted to occur between *rps*16 and *atp*A according to RFLP analysis, was found by PCR analysis only in *P. persica* where it consisted of a deletion within the spacer region between *trn*S and *trn*G (Fig. 4a, b). This deletion was determined to be 277 bp in length by sequence analysis. In addition, PCR amplification of this region produced a fragment that was 118 bp shorter in *P. zippeliana* and *P. salicifolia* than in other accessions (except *P. persica*; Fig. 4b). Sequence analysis showed that this fragment length reduction resulted from a combination of several small deletions.

Relationships between cpDNA types

The UPGMA tree consisted of two major groups, one with chloroplast genome types A–I (*Amygdalus*, *Prunus*, and *Cerasus* sect. *Microcerasus*) and another with chloroplast genome types J and K (*Laurocerasus* and *Padus*; Fig. 5a). The strict consensus tree of the eight equally parsimonious trees, which is 68 steps long with a CI of 0.93 and an RI of 0.64, showed different topology (Fig. 5b). Although only weekly supported, cpDNA types D–K (*Prunus*, *Cerasus*, *Laurocerasus*, and *Padus*) formed a monophyletic group. Two clades including cpDNA types D and E, and J and K, respectively, were relatively well supported in this tree. The three cpDNA types A, B, and C, all belonging to *Amygdalus*, did not form a clade. Two clusters, i.e., cpDNA types D and E, and cpDNA types J and K were supported in both trees.

Discussion

Peach chloroplast genome size and gene order

In our study, the peach chloroplast genome was estimated to be ca. 152 kb, which is about 10 kb longer than previously reported from estimations based on a gel electrophoresis image (Kaneko et al. 1986; Uematsu et al. 1991). Using this method, it would be difficult to visualize very small fragments or to determine copy number correctly. In the present study, accurate sizes for

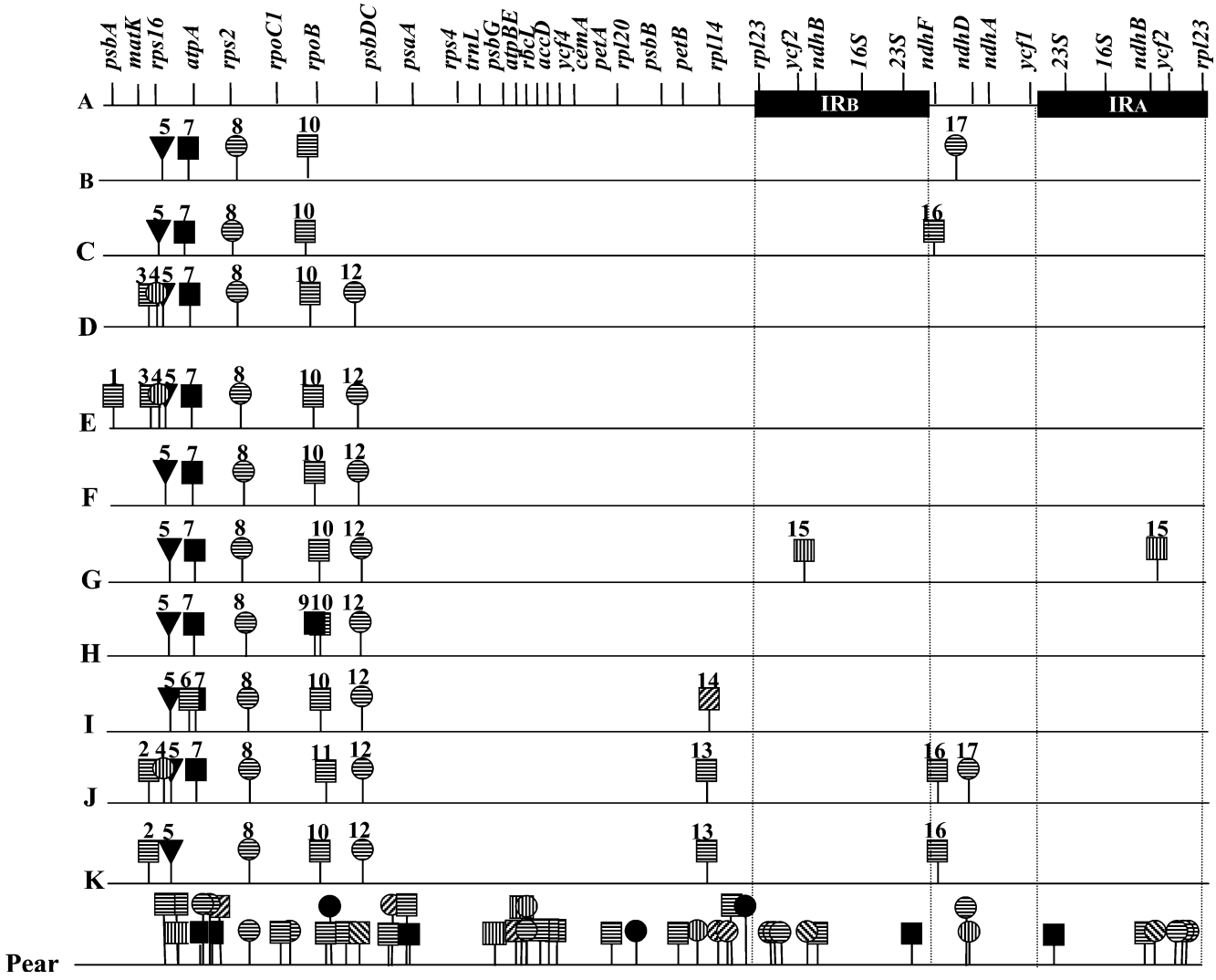

Fig. 3 Locations of the length and restriction site mutations observed among 11 cpDNA types presented on the physical map of peach cpDNA (*A type*). An *inverted triangle* indicates insertion compared with *P. persica*. The symbols for restriction site changes are as follows: ▀ = *Sac*I, ▥ = *Xho*I, ▤ = *BamH*I, ▨ = *Pst*I site. *Square and circle* symbols represent a site gain and loss, respectively, for each enzyme

Table 5 Restriction site differences observed and base substitution rates estimated among 11 *Prunus* chloroplast genome types

Cp genome type	A	B	C	D	E	F	G	H	I	J	K	Average in *Prunus*	Pear
A		4	4	6	7	4	6	5	6	9	6	5.7	48
B	0.63		2	4	5	2	4	3	4	7	6	4.1	48
C	0.61	0.30		4	5	2	4	3	4	7	4	3.9	50
D	0.94	0.63	0.61		1	2	4	3	4	7	6	4.1	52
E	0.78	0.77	0.77	0.15		3	5	4	5	8	7	5.0	53
F	0.63	0.32	0.30	0.32	0.47		2	1	2	7	4	2.9	50
G	0.93	0.61	0.61	0.61	0.77	0.30		3	4	9	6	4.7	52
H	0.77	0.47	0.15	0.47	0.61	0.15	0.46		3	8	5	3.8	51
I	0.93	0.61	0.61	0.61	0.77	0.3	0.61	0.46		9	6	4.7	52
J	1.39	1.08	1.07	1.08	1.23	1.08	1.39	1.23	1.39		5	7.4	53
K	0.93	0.93	0.61	0.93	1.07	0.61	0.93	0.77	0.93	0.61		5.5	52
Average in *Prunus*	0.85	0.64	0.56	0.64	0.74	0.45	0.72	0.55	0.72	1.16	0.83	0.72	
Pear	8.18	8.18	8.43	8.85	8.96	8.51	8.77	8.63	8.77	8.96	8.63		

The no. of restriction site differences between every pair of cpDNA is shown in the upper right half of this table. The matrix given in the lower left half of the table is the base substitution rate per 100 base pairs in the recognition sites which were calculated after Brown et al. (1979)

Table 6 Number of mutations in six regions of cpDNA observed among 18 *Prunus* species

Region	*BamH*I fragment involved	Size (kb)	Number of restriction site observed	Number of mutated site	Deletion	Mutations per restriction site
LSC-1	B9-B21	20.3	13	7	1	0.53
LSC-2	B19-B6	21.4	12	4	0	0.33
LSC-3	B16-B8	23.2	13	0	0	0.00
LSC-4	B1-B3	23.4	12	2	0	0.17
IR	B24-B6	23.8 (2)	25	1	0	0.04
SSC	B7-B2	16.4	4	2	0	0.50
Total		152.2	106	16	1	0.25

LSC, Large single-copy region; IR, inverted repeats; SSC, small single-copy region

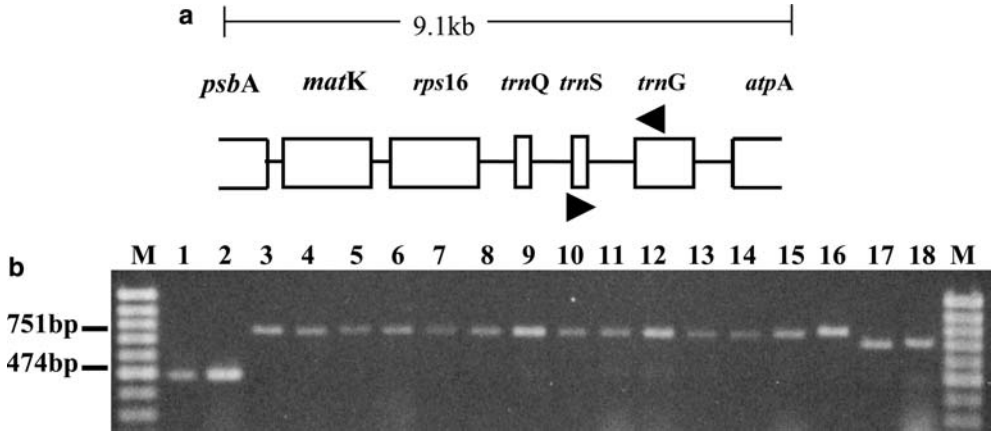

Fig. 4 Gene structure of the 9.1 kb region between *psb*A and *atp*A and the position of synthesized primers for PCR (**a**). PCR patterns obtained from the spacer region between *trn*S and *trn*G in *Prunus* species (**b**). Lane number corresponds to code number for each *Prunus* species in Table 1

all fragments shown in Fig. 2 were confirmed by Southern hybridization. Thus, we could conclude that the genome size is 152 kb. This value indicates that the genome of peach cpDNA is equivalent in size to that of tobacco, *Arabidopsis*, and pear (Shinozaki et al. 1986; Sato et al. 1999; Katayama and Uematsu 2003).

The gene order of the peach chloroplast genome was equivalent to that of tobacco, *Arabidopsis*, spinach, and pear in Rosaceae, which are known as standard type angiosperms, but differed from that of *Lotus* and *Oenothera*, which possess large inversions, and also differed from cpDNAs of cereals, which harbor large deletions and inversions (Shinozaki et al. 1986; Hiratsuka et al. 1989; Hachtel et al. 1991; Maier et al. 1995; Sato et al. 1999; Hupfer et al. 2000; Schmitz-Linneweber et al. 2001; Ogihara et al. 2002; Katayama and Uematsu 2003). In conclusion, the peach cpDNA structure belongs to the standard type of angiosperm cpDNA in terms of molecular size and gene order.

Genetic diversity

The average base substitution rate (100*p*) in *Prunus* cpDNA of 0.72 is not high when compared with other taxa, i.e., 0.66 in wheat–*Aegilops*, 0.38 in *Pisum*, and 0.87 in *Brassica* (Palmer et al. 1983, 1985; Ogihara and Tsunewaki 1988). However, the base substitution rate differs within the family Rosaceae, i.e., 0.72 in *Prunus* versus 0.15 in *Pyrus* (Katayama and Uematsu 2003).

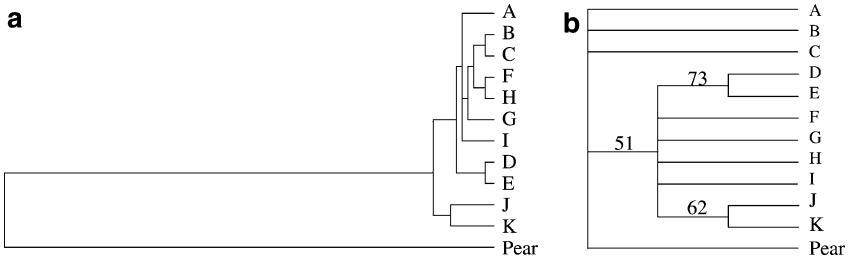

Fig. 5 Unweighted pair group mean cluster analysis (**a**) and strict consensus of the eight equally most parsimonious trees by the heuristic search (**b**) for 11 chloroplast genome types of *Prunus* species and pear. Numbers above branches give the percent occurrence of the clade in 100 bootstrap replications. Chloroplast genome types A–K are given in Table 1

This may reflect a difference in characteristic features between *Prunus* cpDNA and *Pyrus* cpDNA. In *Pyrus*, only three chloroplast genome types were observed in 16 species investigated (Katayama and Uematsu 2003; Katayama and Uematsu, unpublished data). These two types were distinguished by only one restriction site change. Thus the base substitution rate (100*p*) in *Pyrus* was quite low. On the other hand, at least 11 genome types were detected in 18 accessions of *Prunus* investigated in this study. *Prunus* cpDNA is relatively divergent in comparison with *Pyrus* cpDNA.

Hyper variable region in *Prunus* cpDNAs

Almost half of the mutational events, i.e., one length mutation and seven restriction site changes revealed by RFLP analysis were located within the LSC-1 region of 20.3 kb. The deletion of 277 bp between *trn*S and *trn*G region was observed only in *P. persica*. Sequence analysis of the same region revealed that *P. zippeliana* and *P. salicifolia* also had length mutations. This highly variable region in the *Prunus* cpDNA may contain a recombinational hotspot region between *psb*A and *atp*A genes, since short direct repeats of an 8 bp specific motif (GTTATTTA) and an 11 bp length T-stretch (microsatellite sequence) were found inside the break point of the 277 bp deletion in all *Prunus* species (except *P. persica*) examined. In wheat species intra-molecular recombination mediated by short direct repeats has been proposed in the recombinational hotspot region between *rbc*L and *psa*I (Ogihara et al. 1988). Therefore, the 277 bp deletion in *P. persica* could be considered to have resulted from intra-molecular recombination. In order to reveal whether this 9.1 kb region between *psb*A and *atp*A is a recombinational hotspot region or not in *Prunus* cpDNA, further sequence analyses should be carried out to find a unique sequence for intra-molecular recombination.

Evolutionary aspects of cpDNAs in *Prunus*

In the present study, *P. tomentosa* (section *Microcerasus* subgenus *Cerasus*) belongs to the cluster of *Amygdalus* and *Prunus* in the UPGMA tree. This result agrees with the finding that some species including *P. tomentosa* in *Microcerasus* were more closely related to the *Prunus–Amygdalus* group than the *Cerasus–Padus–Laurocerasus* group (Bortiri et al. 2001; Lee and Wen 2001).

Some of the presently analyzed species in subgenera *Amygdalus* and *Prunus* were included also in a previous, RFLP-based study (Uematsu et al. 1991) which, however, lacked species of *Cerasus*, *Laurocerasus* and *Padus*. The major difference between the two studies is that *P. mira* (type B) and *P. davidiana* (type C) cluster closely in the present study but not in the previous. *P. armeniaca* (type F) and *P. mume* (type H) clustered together in both studies, whereas *P. persica* (type A) did not cluster with the other *Amygdalus* species in either study. This may suggest that the cpDNA structure of *P. persica* is rather different from other *Amygdalus* species. The topology of the UPGMA tree obtained by this study was considered to be more accurate and more reliable than the previous one because of the higher resolution of the experimental data obtained.

On the other hand, the maximum parsimonious tree gave a different topology compared with the UPGMA tree. Due to the small number of informative characters, even the clade consisting of cpDNA types D–K was not so strongly supported. Relationships among subgenera in *Prunus* remain uncertain and relationships among species that belong to the subgenus *Amygdalus* are still not clear.

Data concerning length mutations, such as deletions and insertions, were not reflected in the phylogenetic analysis. However, the 277 bp deletion found only in *P. persica* could serve as a useful molecular marker to find the origin of *P. persica* cpDNA. Other length mutations within the 9.1 kb region between *psb*A and *atp*A genes would also be a useful tool for studying cpDNA evolution in *Prunus*.

Acknowledgements We wish to thank the anonymous reviewers for their critical comments which allowed us to improve the original manuscript. Thanks also to Dr. K. Hamada and Dr. H. Araki, HPTCA, and Dr. M. Yamaguchi, NIFTS, for providing the dormant shoots of *Prunus* used in the work. Thanks are also due to Mr. M. Miyake, Mr. S. Kakehi, Mr. H. Hashizume, Mr. S. Kitanishi, and Mr. M. Fujimatsu, FRC-KU, for maintenance of plant materials. Sincere appreciation is expressed to Dr. Anne Edwards, John Innes Centre, UK, and Dr. M. N. Tamura, OCU, for their useful suggestions.

References

Badenes ML, Parfitt DE (1995) Phylogenetic relationships of cultivated *Prunus* species from an analysis of chloroplast DNA variation. Theor Appl Genet 90:1035–1041

Bortiri E, Oh S, Jiang J, Baggett S, Granger A, Weeks C, Buckingham M, Potter D, Parfitt D (2001) Phylogeny and systematics of the genus *Prunus* L. (Rosaceae) as determined by sequence analysis of ITS and the chloroplast *trn*L-*trn*F spacer DNA. Syst Bot 26:797–807

Brettin TS, Karle R, Crowe EL, Iezzoni AF (2000) Chloroplast inheritance and DNA variation in sweet, sour, and ground cherry. J Hered 91:75–79

Brown WM, George M Jr, Wilson AC (1979) Rapid evolution of animal mitochondrial DNA. Proc Natl Acad Sci USA 76:1967–1971

Campbell CS, Donoghue MJ, Baldwin BG, Wojciechowski MF (1995) Phylogenetic relationships in Maloideae (Rosaceae): Evidence from sequences of the internal transcribed spacers of nuclear ribosomal DNA and its congruence with morphology. Am J Bot 82:903–918

Downie SR, Palmer JD (1992) Use of chloroplast DNA rearrangements in reconstructing plant phylogeny. In: Soltis PS, Soltis DE, Doyle JJ (eds) Molecular systematics of plants. Chapman and Hall, New York, pp 14–35

Doyle JJ, Davis JI, Soreng RJ, Garvin D, Anderson MJ (1992) Chloroplast inversion and the origin of the grass family (Poaceae). Proc Natl Acad Sci USA 89:7722–7726

Evans RC, Campbell CS (2002) The origin of the apple subfamily (Maloideae; Rosaceae) is clarified by DNA sequence data from duplicated GBSSI genes. Am J Bot 89:1478–1484

Goldblatt P (1976) Cytotaxonomic studies in the tribe Quillajae (Rosaceae). Ann Missouri Bot Garden 63:200–206

Hachtel W, Neuss A, Stein JV (1991) A chloroplast DNA inversion marks an evolutionary split in the genus *Oenothera*. Evolution 45:1050–1052

Hiratsuka J, Shimada H, Whittier R, Ishibashi T, Sakamoto M, Mori M, Kondo C, Honji Y, Sun CR, Meng BY, Li YQ, Kanno A, Nishizawa Y, Hirai A, Shinozaki K, Sugiura M (1989) The complete sequence of the rice (*Oryza sativa*) chloroplast genome: intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. Mol Gen Genet 217:185–194

Hupfer H, Swiatek M, Hornung S, Herrmann RG, Maier RM, Chiu W-L, Sears B (2000) Complete nucleotide sequence of the *Oenothera elata* plastid chromosome, representing plastome I of the five distinguishable Euoenothera plastomes. Mol Gen Genet 263:581–585

Kaneko T, Terachi T, Tsunewaki K (1986) Studies on the origin of crop species by restriction endonuclease analysis of organellar DNA. II. Restriction analysis of ctDNA of 11 *Prunus* species. Jpn J Genet 61:157–168

Katayama H, Ogihara Y (1993) Structural alterations of the chloroplast genome found in grasses are not common in monocots. Curr Genet 23:160–165

Katayama H, Ogihara Y (1996) Phylogenetic affinities of the grasses to other monocots as revealed by molecular analysis of chloroplast DNA. Curr Genet 20:572–581

Katayama H, Uematsu C (2003) Comparative analysis of chloroplast DNA in *Pyrus* species: physical map and gene localization. Theor Appl Genet 106:303–310

Lee S, Wen J (2001) A phylogenetic analysis of *Prunus* and the Amygdaloideae (Rosaceae) using ITS sequences of nuclear ribosomal DNA. Am J Bot 88:150–160

Maier RM, Neckermann K, Igloi GL, Kossel H (1995) Complete sequence of the maize chloroplast genome: gene content, hotspots of divergence and fine tuning of genetic information by transcript editing. J Mol Biol 251:614–628

Mason SC (1913) The pubescent-fruited species of *Prunus* of the Southwestern States. J Agri Res 1:147–179

Mohanty A, Martin JP, Aguinagalde I (2000) Chloroplast DNA diversity within and among populations of the allotetraploid *Prunus spinosa* L. Theor Appl Genet 100:1304–1310

Mohanty A, Pedro J (2001) A population genetic analysis of chloroplast DNA in wild populations. Heredity 87:421–427

Morgan DR, Soltis DE, Robertson KR (1994) Systematic and evolutionary implications of *rbc*L sequence variation in Rosaceae. Am J Botany 81:890–903

Morton BR, Clegg MT (1993) A chloroplast DNA mutational hotspot and gene conversion in a noncoding region near *rbc*L in the grass family (Poaceae). Curr Genet 24:357–365

Ogihara Y, Tsunewaki K (1988) Diversity and evolution of chloroplast DNA in *Tritucum* and *Aegilops* as revealed by restriction fragment analysis. Theor Appl Genet 76:321–332

Ogihara Y, Terachi T, Sasakuma T (1988) Intramolecular recombination of chloroplast genome mediated by short direct-repeat sequences in wheat species. Proc Natl Acad Sci USA 85:8573–8577

Ogihara Y, Isono K, Kojima T, Endo A, Hanaoka M, Shiina T, Terachi T, Utsugi S, Murata M, Mori N, Takumi S, Ikeo K, Gojobori T, Murai R, Murai K, Matsuoka Y, Ohnishi Y, Tajiri H, Tsunewaki K (2002) Structural features of a wheat plastome as revealed by complete sequencing of chloroplast DNA. Mol Genet Genomics 266:740–746

Palmer JD (1991) Plastid chromosome: structure and evolution. In: Bogorad L, Vasil IK (eds) The molecular biology of plastids. Academic Press, San Diego, pp 5–53

Palmer JD, Shields CR, Cohen DB, Orton TJ (1983) Chloroplast DNA evolution and the origin of amphidiploid *Brassica* species. Theor Appl Genet 65:181–189

Palmer JD, Jorgensen RA, Thompson WF (1985) Chloroplast DNA variation and evolution in *Pisum*; patterns of change and phylogenetic analysis. Genetics 109:195–213

Potter D, Gao F, Bortiri PE, Oh SH, Baggett S (2002) Phylogenetic relationships in Rosaceae inferred from chloroplast *mat*K and *trn*L-*trn*F nucleotide sequence data. Plant Syst Evol 231:77–89

Rehder A (1940) A manual of cultivated trees and shrubs hardy in North American exclusive of the subtropical and warmer temperate regions, 2nd edn. Macmillan, New York, USA

Reiter RS, Young RM, Scolnik PA (1992) Genetic linkage of the Arabidopsis genome: methods for mapping with recombinant inbreds and random amplified polymorphic DNAs. In: Koncz C, Chua NH, Schell J (eds) Methods in Arabidopsis research. World Scientific, Singapore, pp 170–190

Rohlf FJ (1998) NTSYS, numerical taxonomy and multivariate analysis system, version 2.01. Exeter Publishing Ltd., Setauket, New York

Sato S, Nakamura Y, Kaneko T, Asamizu E, Tabata S (1999) Complete structure of the chloroplast genome of *Arabidopsis thaliana*. DNA Res 6:283–290

Sauer JD (1993) Historical geography of crop plants. CRC Press, Boca Raton, Florida

Schmitz-Linneweber C, Maier RM, Alcaraz JP, Cottet A, Herrmann RG, Mache R (2001) The plastid chromosome of spinach (*Spinacia oleracea*): complete nucleotide sequence and gene organization. Plant Mol Biol 45:307–315

Schulze-Menz GK (1964) Rosaceae. In: Melchior H (ed) Engler's syllabus der pflanzenfamilien II, 12th edn. pp 209–218

Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Yamaguchi-Shinozaki K, Ohto C, Torazawa K, Meng BY, Sugita M, Deno H, Kamogashira T, Yamada K, Kusuda J, Takaiwa F, Kato A, Tohdoh N, Shimada H, Sugiura M (1986) The complete nucleotide sequence of tobacco chloroplast genome; its gene organization and expression. EMBO J 5:2043–2049

Swofford DL (1998) Phylogenetic analysis using parsimony, version 4. Sinauer, Sunderland, Massachusetts, USA

Uematsu C, Sasakuma T, Ogihara Y (1991) Phylogenetic relationships in the stone fruit group of *Prunus* as revealed by restriction fragment analysis of chloroplast DNA. Jpn J Genet 66:59–69