Tree Genetics & Genomes (2012) 8:313–326
DOI 10.1007/s11295-011-0442-y

ORIGINAL PAPER

Phylogenetic utility of structural alterations found in the chloroplast genome of pear: hypervariable regions in a highly conserved genome

Hironori Katayama · Miho Tachibana · Hiroyuki Iketani · Shao-Ling Zhang · Chiyomi Uematsu

Received: 19 March 2011 / Revised: 8 September 2011 / Accepted: 19 September 2011 / Published online: 26 November 2011
© Springer-Verlag 2011

Abstract The genome structure of pear chloroplast DNA (cpDNA) is extremely highly conserved in comparison with that of other angiosperms, and therefore, relatively few phylogenetic analyses for pear (*Pyrus* spp.) have been carried out using cpDNA as a marker. In this study, we identified two hypervariable regions in intergenic spacers of cpDNA from 21 species in *Pyrus*. One of these regions is 857 bp in length and lies between the *acc*D-*psa*I genes, and the other is a 904-bp region between the *rps*16-*trn*Q genes. The mutation rate of gaps for the two regions was 10 and 26 times higher, respectively, than the base change rate. Twenty-five haplotypes were revealed among 21 species in *Pyrus* by 36 mutations found in the two regions. These included 27 gaps and 9 base changes but excluded cpSSRs. Phylogenetic relationships between the 25 haplotypes were generated by haplotype network analysis. The 25 haplotypes represented three groups (types A–C) with two large deletions, one 228 bp in length between the *acc*D-*psa*I genes and the other 141 bp between the *rps*16-*trn*Q genes. Types A and B consisted mostly of pears native to East and South Asia. Type C contained mainly *Pyrus communis* and wild relatives native to Europe, West and Central Asia, Russia, and Africa. Type B might have diverged from primitives such as pea pears in type A. Phylogenetic utility of structural alterations (gaps) occurring in the hypervariable regions of *Pyrus* cpDNA is discussed.

Keywords *Pyrus* · Pear · Chloroplast DNA · Hypervariable region · Gaps · Indels

Communicated by E. Dirlewanger

H. Katayama (✉)
Food Resources Education and Research Center, Kobe University,
Kasai 675-2103 Hyogo, Japan
e-mail: hkata@kobe-u.ac.jp

M. Tachibana
Engineering Headquarters, OYO Corporation,
Saitama-shi 331-8688 Saitama, Japan

H. Iketani
National Institute of Fruit Tree Science,
National Agricultural and Food Research Organization,
Tsukuba 305-8605 Ibaraki, Japan

S.-L. Zhang
College of Horticulture, Nanjing Agriculture University,
Nanjing 210095, China

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

Introduction

The genus *Pyrus* in Pyrinae is widely distributed in Asia, Europe, and North Africa and has a long history of utilization of its fruits as food. The classification of *Pyrus* is often very difficult due to natural or artificial interspecific hybrids, which arise easily because of self-incompatibility and the lack of distinguishable characters between species. *Pyrus* can generate fertile progeny with ease, and repetitious cross hybridization generates numerous progeny harbouring different heterogeneity. Some of these could be given the rank of form, variety, or species (Rubtsov 1944; Browicz 1993). Rehder (1940) described 15 principal species, 6 varieties, and 5 related species in *Pyrus*. Rubstov (1944) considered that 14 species found in Central Asia should be added, and he proposed that the genus *Pyrus* consisted of c.a. 35 species in total. However, some of these are suspected of being hybrid so that Challice and Westwood (1973) tried to reconsider speciation in *Pyrus* and

recognized about 22 species according to numerical taxonomy using morphological characters and phenolic compounds. Bell (1990) recognized 22 primary species of pear, whose distribution covers most of Europe, temperate Asia, and mountain areas of northern Africa. In addition, he recognized at least nine natural or artificial interspecific hybrids between primary species; those were classified into different species.

Rubstov (1944) reported that European pear cultivars of *Pyrus communis* exhibit the characteristics of at least four species: *Pyrus elaegrifolia* Pall., *Pyrus salicifolia* Pall., *Pyrus korshinskyi* Litv., and *Pyrus syriaca* Boiss. Wild populations of *P. communis* var. *pyraster* and/or *Pyrus caucasica* Fed. are possible ancestors of cultivated *P. communis*, and there is also some evidence indicating an involvement of *Pyrus nivalis* Jacq. (Challice and Westwood 1973). However, it is still difficult to determine the origin of European pear cultivars phenotypically or even by DNA markers (Aldasoro et al. 1996; Volk et al. 2006).

Nakai (1919) described that four species in *Pyrus* native to Japan, i.e., *Pyrus pyrifolia* [Burm.] Nakai (Japanese pear), *Pyrus dimorphophylla* Makino (pea pear), *Pyrus hondoensis* [Nakai *et* Kikuchi] Rehder, and *Pyrus aromatica* [Nakai *et* Kikuchi] Rehder. *P. hondoensis* and *P. aromatica* are included in *Pyrus ussuriensis* Maxim. and treated as its varieties (*P. ussuriensis* var. *hondoensis* (Nakai et Kikuchi) Rehder and *P. ussuriensis* var. *aromatica* (Nakai et Kikuchi) Ohwi) by Rehder (1920) and Ohwi (1965), respectively. Although almost all modern Japanese pear cultivars are included in *P. pyrifolia*, the origin of *P. pyrifolia* is still unclear. Generally, Japanese pear cultivars were considered to have been domesticated from native *P. pyrifolia* occurring in Japan (Kikuchi 1948). However, candidate trees of native *P. pyrifolia* were only found near human habitation and therefore presumed to be escapes. The progenitor of Japanese pear cultivars may have come prehistorically from China (Shirai 1929; Kajiura 1983). Recently *P. ussuriensis* var. *aromatica* has been reported as native and prehistorically introduced into Northern Tohoku region in Japan by means of taxonomical and molecular analyses (Iketani et al. 2010).

In Chinese pear, 15 pear species have been identified (Gu and Sponberg 2003). Pear cultivars native to China consist of the following four groups: Chinese sand pear (*P. pyrifolia* Nakai), Ussurian pear (*P. ussuriensis* Maxim.), Chinese white pear (*Pyrus bretschneideri*), and Xinjiang pear (*Pyrus sinkiangensis* Yu) (Yu 1979). Kikuchi (1946) proposed that *P. bretschneideri* might be generated by hybridization between *P. ussuriensis* and *P. pyrifolia* according to their geographical distribution. A recent study using RAPD and SSR markers could not distinguish *P. bretschneideri* and *P. pyrifolia* (Teng et al. 2002; Bao et al. 2007). Teng et al. (2002) suggested that, from molecular data, *P. pyrifolia* might be a common progenitor of *P. bretschneideri* and Japanese pear cultivars.

Although it is well known that the structure and gene content of the chloroplast genome are conserved among divergent plant species, structural alterations such as insertions, deletions (gaps), inversions, and translocations have been found in certain plant lineages by comparing the genome structure of chloroplast DNA (cpDNA). Therefore, using mutational events in chloroplast genomes as DNA markers 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 intramolecular recombination hotspots, have been detected in the cpDNA of Poaceae (Ogihara et al. 1988; Hiratsuka et al. 1989; Morton and Clegg 1993; Maier et al. 1995; Katayama and Ogihara 1996). Tracing of indels in one hypervariable region led to an understanding of the phylogenetic relationships at interspecific level and at lower levels (Guo and Terachi 2005).

Although the family Rosaceae contains many commercially important fruit trees and ornamental trees, the chloroplast (cp) genome structure in rosaceous plants is still not well known. Recently, the cp genome structure was characterized at the physical map level for *Pyrus* and *Prunus* in Spiraeoideae (Katayama and Uematsu 2003, 2005). These reports indicated the extremely low cpDNA diversity in *Pyrus* and the existence of a hypervariable region in *Prunus* cpDNAs. At the DNA sequence level, Shaw et al. (2007) reported 21 variable noncoding chloroplast regions among angiosperms including *Prunus.*

Iketani et al. (1998) highlighted inconsistencies between the cpDNA types and the morphological classification when analysing relationships between Occidental and Oriental pears by RFLP in cpDNAs. They supposed hybridization or lineage sorting and proposed that a large number of materials including pure natural population would be necessary to solve this issue. Using comparative sequence analysis in six cpDNA regions such as *atp*B-*rbc*L intergenic spacer, *trn*L-Fspacer, *acc*D-*psa*I spacer, *ndh*A intron, *rpl*16 intron, and *rpo*C1 intron in cultivated pear varieties from Asia, Kimura et al. (2003) noted that the phylogenetic relationships among Asian pears was quite mingled and incongruent, because of lack of polymorphisms from the result of high conservation in the cp genome in *Pyrus*.

We tried to develop the hypervariable regions harbouring structural alterations as a cpDNA marker to trace the evolution of the cp genome and to understand the phylogenetic position at the inter and/or intraspecific level in *Pyrus*. In this report, the phylogenetic utility of hypervariable regions in the extremely conserved chloroplast genome is evaluated and effectiveness of haplotype network analysis in Occidental and Oriental pear cpDNAs,

using both gaps and base changes found in two hypervariable regions, is determined.

Materials and methods

Eighty-one accessions from 21 species in *Pyrus* originating from Asia, Europe, and Africa were selected for this study (Table 1). Total DNA of European, West Asian, and North African pears was obtained from the DNA bank in the Royal Botanic Gardens, Kew, UK. Twenty-one accessions of Chinese pear varieties are maintained at the Nanjing Agricultural University, China. Two accessions of *P. ussuriensis* Maxim. var. *hondoensis* (Nakai & Kikuchi) Rehder endemic to Japan are maintained at the National Institute of Fruit Tree Science. Other Japanese pear cultivars and wild pears in Japan were selected from the *Pyrus* germplasm collection at the Food Resources Education and Research Center, Kobe University, or selected from the collection of the Botanical Gardens, Osaka City University. Classification of these *Pyrus* followed Rehder (1940), Yu (1979), and Ohwi (1965).

Total DNA was isolated from young leaves of pear according to the method reported by Hosaka (1995). Oligonucleotide primers designed by Small et al. (1998) were used to detect a hypervariable region in the intergenic region between *acc*D-*psa*I genes in *Pyrus* cpDNA. Another hypervariable region between *rps*16-*trn*Q was amplified using primers *rps*16-F and *trn*Q-R (Table 2). PCR amplification of intergenic region between *rbc*L-*acc*D was performed by primers *rbc*L-F and *acc*D-R (Table 2). Amplified fragments were fractionated by agarose gel electrophoresis. Primers used for sequencing of the two hypervariable regions were shown in Table 2. Positions of primers corresponded to nucleotide numbers of tobacco and pear cpDNA complete sequences (accession numbers Z00044 and AP012207 in GenBank/EMBL/DDBJ).

DNA sequences of amplified fragments were determined using ABI3100 Genetic Analyser (Applied Biosystems, USA). The nucleotide sequences reported here have been assigned GenBank/EMBL/DDBJ accession numbers AB603891 to AB603936 for *rbc*L-*acc*D, AB604678 to AB604722 for *rps*16-*trn*Q, and AB604874 to AB604916 for *acc*D-*psa*I. The sequence of 'Chojuro' for *acc*D-*psa*I was previously assigned to GenBank accession no. AB204878 (Kimura et al. 2003).

DNA sequences were aligned using multiple alignment, GENETYX-MAC ver.12 (GENETYX CO. Japan) with minor manual adjustments. Mutation rates among haplotypes in different regions were calculated as the average number of nucleotide substitutions and length mutations per site according to Jukes and Cantor (1969) using the PAUP version 4.0b10 (Swofford 1998). A median-joining network (MJ) including potential median vectors was performed with Network. 4.5.1.6 (Bandelt et al. 1999) using DNA nucleotide data based on the gaps (indels) and base changes found in two hypervariable regions. Gaps and base changes were scored as independent single character.

Results

Two hypervariable regions in the conserved pear cp genome

A total of 40 mutations including 30 gaps (indels) and 9 base changes were found and characterized into three intergenic regions (*acc*D-*psa*I, *rps*16-*trn*Q, and *rbc*L-*acc*D) for 45 accessions. The length of an intergenic region between *rbc*L-*acc*D was determined for 45 accessions from 21 species in *Pyrus* by sequencing. A base change was found at an intergenic region between *rbc*L-*acc*D among 45 accessions. The base change and length mutation rates revealed with the Jukes and Cantor method among haplotypes in this region were quite low (ranging from 0.0009 for 3 accessions of *Pyrus pashia*, *P. korshinskyi*, and *P. ussuriensis* var. *hondoensis* to 0.001 for 21accessions with an average of 0.001, and from 0.0006 for 25 accessions to 0.0024 for an accession of *Pyrus gharbiana* with an average of 0.0009) for the 21 species in *Pyrus* (Table 3). The base change rate was almost equivalent to the length mutation rate in an intergenic region between *rbc*L-*acc*D. The length of an intergenic region between *acc*D-*psa*I was determined for 21 species in *Pyrus* by sequencing and found to vary from 585 bp for *P. korshinskyi* and *P. bretschneideri* 'Pingli' to 857 bp for *P. pyraster*. Sixteen haplotypes were determined by 15 gaps and 6 base changes found in this region. Although the base change rate among 45 accessions in this region was quite low (ranging from 0.0003 for 34 accessions to 0.0021 for an accession of *Pyrus cossonii* with an average of 0.0006) for the 21 species in *Pyrus*, the length mutation rate ranged from 0.0037 for seven accessions to 0.013 for an accession of *P. pyraster* with an average of 0.0056 and was approximately 9.3 times higher than the base change rate (Table 3). The length of the gaps varied from 1 to 228 bp (Table 4). Ten motifs of short direct repeats varied from 6 to 919 bp in length, and one inverted repeat motif of 15 bp in length was found in association with length variations in each repeat (Table 6).

The length of the intergenic region between *rps*16-*trn*Q genes varied from 662 bp for *P. cossonii* to 904 bp for *P. korshinskyi*. Fourteen haplotypes were determined by 12 gaps and 3 base changes found in this region (Table 5). The base change rates were also quite low, ranging from 0.00008 for 42 accessions to 0.0016 for 2 accessions of *P.*

Table 1 Plant materials used in this study and cpDNA haplotypes determined by two large gaps

Code	Species	Cultivar and/ or accession	Distribution	Source	*acc*D-*psa*I G4 (228 bp)	*rps*16-*trn*Q G3 (141 bp)	Haplotype
1	*Malus* x *domestica* Borkh.	Fuji		KU-FRC	−	−	A[a]
2	*P. nivalis* Jacq.	1982–5375	WC Europe,	RBG-KEW	−	+	C[a]
3	*P. amygdaliformis* Vill.	1973–11528	Med. Europe, Asia Minor	RBG-KEW	−	+	C[a]
4	*P. elaeagrifolia* Pall.	4A0126	SE Europe, Russia, Turkey	BG-OCU	−	+	C[a]
5	*P. cossonii* Rehd.	1986–1630	Algeria	RBG-KEW	−	+	C[a]
6	*P. gharbiana* Trab.	1986–1627	Morocco, W Algeria	RBG-KEW	−	+	C[a]
7	*P. mamorensis* Trab.	1993–3514	Morocco	RBG-KEW	−	+	C[a]
8	*P. salicifolia* Pall.		NW Iran, NE Turkry, S Russia	Market	−	+	C[a]
9	*P. regelii* Rehd.	1986–8384	SC Asia	RBG-KEW	−	−	A[a]
10	*P. pashia* D. Don.	1986–1616	Pakistan, India, Nepal	RBG-KEW	+	−	B[a]
11	*P. cordata* Desv.	Plym-16	W France, UK	Plymouth, UK	−	+	C[a]
12	*P. cordata* Desv.	Plym-17	W France, UK	Plymouth, UK	−	+	C[a]
13	*P. korshinskyi* Litv.	1949–15208	SC Asia, Afghanistan	RBG-KEW	+	−	B[a]
14	*P. balansae* Decne.	1969–16873	Turkey, Iran	RBG-KEW	−	+	C[a]
15	*P. pyraster* Burgsd.	1969–2432	SE Europe, Turkey	RBG-KEW	−	+	C[a]
16	*P. caucasica* Fed.	1986–8388	SE Europe, Turkey	RBG-KEW	−	+	C[a]
European pear cultivar							
17	*P. communis* L.	La France	West to SE. Europe	KU-FRC	−	+	C[a]
18	*P. communis* L.	Passe Crassane	West to SE. Europe	KU-FRC	−	+	C[a]
19	*P. communis* L.	Bartlett	West to SE. Europe	KU-FRC	−	+	C[a]
20	*P. communis* L.	Flemish Beauty	West to SE. Europe	KU-FRC	−	+	C[a]
21	*P. communis* L.	Starkrimson	West to SE. Europe	KU-FRC	−	+	C[a]
22	*P. communis* L.	General Leclerc	West to SE. Europe	KU-FRC	−	+	C[a]
23	*P. communis* L.	Le Lectier	West to SE. Europe	KU-FRC	−	+	C[a]
24	*P. communis* L.	Marguerite Marillat	West to SE. Europe	KU-FRC	−	+	C[a]
Pea pear							
25	*P. betulaefolia* Bunge	4A0155	CS China, S Manchuria	BG-OCU	−	−	A[a]
26	*P. fauriei* Scheid.	4A0118	Korea	BG-OCU	−	−	A[a]
27	*P. fauriei* Scheid.	4A0122	Korea	BG-OCU	−	−	A[a]
28	*P. dimorphophylla* Mak.	4A0159	Japan	BG-OCU	−	−	A[a]
Japanese pear cultivar							
29	*P. pyrifolia* (Burm.) Nak.	Choujuurou	Japan	KU-FRC	+	−	B[a]
30	*P. pyrifolia* (Burm.) Nak.	Housui	Japan	KU-FRC	+	−	B[a]
31	*P. pyrifolia* (Burm.) Nak.	Kousui	Japan	KU-FRC	+	−	B[a]
32	*P. pyrifolia* (Burm.) Nak.	Nijisseiki	Japan	KU-FRC	+	−	B[a]
33	*P. pyrifolia* (Burm.) Nak.	Imamuraaki	Japan	KU-FRC	+	−	B[a]
34	*P. pyrifolia* (Burm.) Nak.	Nangetsu	Japan	KU-FRC	−	−	A[a]
35	*P. pyrifolia* (Burm.) Nak.	Niitaka	Japan	KU-FRC	−	−	A[a]
36	*P. pyrifolia* (Burm.) Nak.	Aikansui	Japan	KU-FRC	+	−	B[b]
37	*P. pyrifolia* (Burm.) Nak.	Atago	Japan	KU-FRC	+	−	B[b]
38	*P. pyrifolia* (Burm.) Nak.	Chikusui	Japan	KU-FRC	+	−	B[b]
39	*P. pyrifolia* (Burm.) Nak.	Hougetsu	Japan	KU-FRC	+	−	B[b]
40	*P. pyrifolia* (Burm.) Nak.	Shinkou	Japan	KU-FRC	+	−	B[b]
41	*P. pyrifolia* (Burm.) Nak.	Shinsei	Japan	KU-FRC	+	−	B[b]
42	*P. pyrifolia* (Burm.) Nak.	Shinsetsu	Japan	KU-FRC	+	−	B[b]
43	*P. pyrifolia* (Burm.) Nak.	Syuugyoku	Japan	KU-FRC	+	−	B[b]
44	*P. pyrifolia* (Burm.) Nak.	Tama	Japan	KU-FRC	+	−	B[b]
45	*P. pyrifolia* (Burm.) Nak.	Yakumo	Japan	KU-FRC	+	−	B[b]

Table 1 (continued)

Code	Species	Cultivar and/ or accession	Distribution	Source	*acc*D-*psa*I G4 (228bp)	*rps*16-*trn*Q G3 (141bp)	Haplotype
46	*P. pyrifolia* (Burm.) Nak.	Yasato	Japan	KU-FRC	+	−	B[b]
47	*P. pyrifolia* (Burm.) Nak.	Tsukushiinunashi	Japan	BG-OCU	+	−	B[b]
Chinese sand pear cultivar							
48	*P. pyrifolia* (Burm.) Nak.	Jimi	China	CH-NAU	+	−	B[b]
49	*P. pyrifolia* (Burm.) Nak.	Cangxi xueli	China	CH-NAU	−	−	A[b]
50	*P. pyrifolia* (Burm.) Nak.	Xinya	China	CH-NAU	+	−	B[b]
51	*P. pyrifolia* (Burm.) Nak.	Deshengxian	China	CH-NAU	+	−	B[b]
Chinese white pear cultivar							
52	*P. bretschneideri* Rehd..	Yali	China	BG-OCU	+	−	B[a]
53	*P. bretschneideri* Rehd..	Pingli		BG-OCU	+	−	B[a]
Chinese ussurian pear cultivar							
54	*P. ussuriensis* Maxim.	Zhuili	China	BG-OCU	−	−	A[a]
55	*P. ussuriensis* Maxim.	Jianbali	China	BG-OCU	−	−	A[b]
56	*P. ussuriensis* Maxim.	Tzuli	China	BG-OCU	−	−	A[b]
57	*P. ussuriensis* Maxim.	Balixiang	China	CH-NAU	−	−	A[a]
58	*P. ussuriensis* Maxim.	Gongchuanli	China	CH-NAU	+	−	B[a]
59	*P. ussuriensis* Maxim.	Nanguoli	China	CH-NAU	−	−	A[b]
60	*P. ussuriensis* Maxim.	Quizi	China	CH-NAU	−	−	A[b]
61	*P. ussuriensis* Maxim.	Xiehuatian	China	CH-NAU	+	−	B[b]
62	*P. ussuriensis* Maxim.	Jingbai	China	CH-NAU	+	−	B[b]
63	*P. ussuriensis* Maxim.	Wuxiang	China	CH-NAU	−	−	A[b]
64	*P. ussuriensis* Maxim.	Fuwuxiang	China	CH-NAU	−	−	A[b]
65	*P. ussuriensis* Maxim.	Hongbalixiang	China	CH-NAU	−	−	A[b]
66	*P. ussuriensis* Maxim.	Xinyali	China	CH-NAU	+	−	B[b]
67	*P. ussuriensis* Maxim.	Taianchangba	China	CH-NAU	+	−	B[b]
68	*P. ussuriensis* Maxim.	Gaopingdahuang	China	CH-NAU	+	−	B[b]
69	*P. ussuriensis* Maxim.	Lijiangbaili	China	CH-NAU	−	−	A[b]
70	*P. ussuriensis* Maxim.	Daxiangshui	China	CH-NAU	+	−	B[b]
71	*P. ussuriensis* Maxim.	Qinan changba	China	CH-NAU	+	−	B[b]
72	*P. ussuriensis* Maxim.	Binxian Dayisheng	China	CH-NAU	−	−	A[b]
73	*P. ussuriensis* Maxim.	Baipisu	China	CH-NAU	+	−	B[b]
Wild ussurian pear in Japan							
74	*P. ussuriensis* var. *aromatica* (Nak. et Kik.) Rehd.	4A0128	Japan	BG-OCU	−	−	A[a]
75	*P. ussuriensis* var. *aromatica* (Nak. et Kik.) Rehd.	4A0143	Japan	BG-OCU	−	−	A[a]
76	*P. ussuriensis* var. *aromatica* (Nak. et Kik.) Rehd.	Sotoorihime	Japan	NIFTS	+	−	B[a]
77	*P. ussuriensis* var. *hondoensis* (Nak. et Kik.) Rehd.	4A0134	Japan	BG-OCU	−	−	A[a]
78	*P. ussuriensis* var. *hondoensis* (Nak. et Kik.) Rehd.	Mt-6[c]	Japan	NIFTS	−	−	A[a]
79	*P. ussuriensis* var. *hondoensis* (Nak. et Kik.) Rehd.	Kr-1[c]	Japan	NIFTS	−	−	A[a]
80	*P. ovoidea* Rehd.	4A0151	Japan	BG-OCU	−	−	A[b]
81	*P.*x *phaeocarpa* Rehd.	4A0165	N. China	BG-OCU	−	−	A[b]

Adapted from Bell (1990). *KU-FRC* Food Resources Center, Kobe University, *RBG-KEW* Royal Botanical Garden-KEW, *BG-OCU* Botanical Gardens, Osaka City University, *CH-NAU* College of Horticulture, Nanjing Agriculture University, *NIFTS* National Institute of Fruit Tree Science, + and − means the presence or absence of large gaps (deletions) in two intergenic regions

[a] Haplotypes determined by sequencing in two hypervariable regions

[b] Haplotypes determined by PCR amplification in two hypervariable regions

[c] Refer to Iketani et al. (2010) for more information on two accessions (Mt-6 and Kr-1) of *P. ussuriensis* var. *hondoensis*

Table 2 PCR and sequencing primers used in this study

Region	Primer name	Primer (5′–3′)	Length (bp)	Position	Reference
*rbc*L-*acc*D	F	TGATCTTGCTCAGGAAGGTA	20	58891–58910[a]	In this study
	R	CGCGTTTAGTTGGCCTATGC	20	60798–60779[a]	In this study
*acc*D-*psa*I					
For PCR	F	GGAAGTTTGAGCTTTATGCA	20	60955–60974[a]	Small et al. (1998)
	R	TTTCCGGCAATTGCAATGGCTTCT	24	62162–62139[a]	Small et al. (1998)
For sequencing	F	GGAAGTTTGAGCTTTATGCA	20	60955–60974[a]	In this study
	2F	CTGAGTGAGTTATTTAAGCTCC	22	61258–61279[a]	In this study
	2R	GGACGTTAAATAACTCACTCAG	22	61279–61258[a]	In this study
	RS4	GGCTTAGTATTTCCGGCAATTGC	23	62771–62749[b]	In this study
*rps*16-*trn*Q					
For PCR	*rps*16-F	ACCACATCGTTTCAAACGAAG	21	6182–6202[a]	In this study
	*trn*Q-R	GCTATTCGGAGGTTCGAATCC	21	7447–7427[a]	In this study
For sequencing	16F	ACCACATCGTTTCAAACGAAG	21	6182–6202[a]	In this study
	F2	CTCTTAAACGGAAGACTGGTCG	22	7076–7097[b]	In this study
	R1	CCTATCACGTTACTTGAAGATAGA	24	6626–6603[b]	In this study
	R5	CAGAACATACCTGACCCACGATCAT	25	7248–7224[b]	In this study

[a] Tobacco cpDNA complete sequences to which positions of primers correspond (accession number Z00044 in DNA data bank)

[b] Pear cpDNA complete sequences to which positions of primers correspond (accession number AP012207 in DNA data bank)

gharbiana and *P. ussuriensis* var. *hondoensis* with an average of 0.0001 for the 21 species in *Pyrus* (Table 3). The length mutation rate ranged from 0.0018 for eight accessions to 0.0052 for an accession of *P. cossonii* with an average of 0.0028 and was approximately 28 times higher than the base change rate. The length of the 12 gaps varied from 1 to 141 bp (Table 5). This region also contained six direct repeats ranging from 16 to 42 bp in each gap (Table 6; Fig. 1). Two chloroplast simple sequence repeat loci (repetition of the single nucleotide ‘A’ between 9 and 21 times) were found adjacent to the borders of the 141-bp large gap in the region between the *rps*16-*trn*Q genes (Table 7). These two SSR loci were removed from the data set for phylogenetic analysis, because the length of the SSR loci was too variable among accessions. Finally, 25 haplotypes, H1-H25, derived from the two intergenic regions *acc*D-*psa*I and *rps*16-*trn*Q, were identified among 21 species in *Pyrus* by 36 mutations including 27 gaps and 9 base changes (Tables 4 and 5). The mutation rates including both base changes and length mutations for two intergenic regions *acc*D-*psa*I and *rps*16-*trn*Q among 21 species in *Pyrus* were estimated ranging from 0.004 to 0.0132 with an average of 0.0062 and from 0 to 0.0044 with an average of 0.0027. Two regions evolved 1.4–3.2 times faster than another intergenic region between the *rbc*L-*acc*D genes ranging 0 to 0.0035 with an average of 0.0019.

Haplotype network inferred from gap characters

Analysis of genetic relatedness of 21 species in *Pyrus* was based on the 36 mutations of 27 gaps and 9 base changes from 25 haplotypes (H1-H25) with the MJ network (Bandelt et al. 1999). Validity of gaps coded as phylogenetic characters was verified. The MJ network enabled the identification of three major groups (types A–C) from two large deletions, (the 228-bp deletion in *acc*D-*psa*I and the 141-bp deletion in *rps*16-*trn*Q) (Fig. 2). The 141-bp

Table 3 Mutation rates of three intergenic regions among cpDNA haplotypes in *Pyrus*

Region	Length without gaps (bp)	No. of sequenced accessions	No. of base change	Base substitution rate		No. of gaps	Rate of length mutations	
				Range	Average		Range	Average
*acc*D-*psa*I	548	45	3[a]	0.0003–0.0021	0.0006	15	0.0040–0.0132	0.0062
*rps*16-*trn*Q	608	45	1[b]	0.00008–0.0016	0.0001	9[b]	0.0018–0.0052	0.0026
*rbc*L-*acc*D	537	45	1	0.0006–0.0024	0.0009	2	0.0009–0.0016	0.0010

Mutation rates were calculated by the Jukes–Cantor method (Jukes and Cantor 1969)

[a] Three base changes (BC2～BC4) within G4 were not used in the calculation of mutation rates in *acc*D-*psa*I

[b] Three indels (G4～G6) and two base changes (BC2 and BC3) within G3 were not used in the calculation of mutation rates in *rps*16-*trn*Q

Table 4 Characterization of 16 haplotypes by insertions (+)/deletions (−) and base changes found in the 857-bp intergenic region between *acc*D and *psa*I in *Pyrus* species by DNA sequencing

Haplo type	Code[a]	Base change (bc) and gap (g)																																
		–[b]	–[b]	–[b]	–[b]	–[b]	–[b]	–[b]	–[b]	bc1 137[c]	–[b]	g1 255[c]	g2 262[c]	g3 278[c]	g4 294[c]	–[b]	bc2 355[c]	bc3 400[c]	bc4 410[c]	–[b]	g5 514[c]	g6 514[c]	g7 540[c]	bc5 568[c]	g8 571[c]	g9 573[c]	g10 590[c]	g11 607[c]	g12 624[c]	g13 697[c]	g14 719[c]	bc6 740[c]	g15	–[b]
h1	3, 11, 12, 22	c	g	a	g	c	t	a	g	g	t					t	c	t	a	a	−10[d]	-g		g				+20[d]				c		
h2	14, 17~21	c	g	a	g	c	t	a	g	g	t					t	c	t	a	a	-10[d]	-g		g				+20[d]				c		
h3	23	c	g	a	g	c	t	a	g	g	t					t	c	t	a	a	-10[d]	-g		g				+20[d]				c		
h4	24	c	g	a	g	c	t	a	g	g	t					t	c	t	a	a				g				+20[d]				c		
h5	15	c	g	a	g	c	t	a	g	g	t					t	c	t	a	a	-10[d]	-g	-13[d]	g-	-g	+tt	+15[d]	+20[d]	+17[d]			c		
h6	2,8	c	g	a	g	c	t	a	g	g	t					t	c	t	a	a				g				+20[d]				c		
h7	16	c	g	a	g	c	t	a	g	g	t					t	c	t	a	a	-10[d]	-g	13[d]	g		+tt		+20[d]	+17[d]	+23[d]		c		
h8	4	c	g	a	g	c	t	a	g	g	t					t	c	t	a	a	-10[d]	-g		g				+20[d]			+6[d]	c		
h9	5	c	g	a	g	c	t	a	g	g	t	-23[d]				t	c	t	a	a	-10[d]	-g		t				+20[d]				c		
h10	6	c	g	a	t	c	t	a	g	t	t					t	c	t	a	a				g				+20[d]				c		
h11	7	c	g	a	g	c	t	a	g	g	t			+23[d]		t	c	t	a	a				g				+20[d]				c		
h12	53	c	g	a	g	c	t	a	g	g	t				-228	nd‡	nd‡	nd‡	nd‡	nd‡				g								c		
h13	13,	c	g	a	g	c	t	a	g	g	t				-228	nd‡	nd‡	nd‡	nd‡	nd‡				”								c		
h14	10	c	g	a	g	c	t	a	g	g	t				-228	nd‡	nd‡	nd‡	nd‡	nd‡				g								c	+22[d]	
h15	29, 30, 32, 33 52, 58,	c	g	a	g	c	t	a	g	g	t		+t		-228	nd‡	nd‡	nd‡	nd‡	nd‡				g								c		
h16	76	c	g	a	g	c	t	a	g	g	t				-228	nd‡	nd‡	nd‡	nd‡	nd‡				g								c		
h17	31,	c	g	a	g	c	t	a	g	g	t				-228	nd‡	nd‡	nd‡	nd‡	nd‡				g								c		
h18	34, 35, 54, 57	c	g	a	g	c	t	a	g	g	t					t	t	g	a	a				g		+tt		+20[d]				g		
h19	25~27, 74,	c	g	a	g	c	t	a	g	g	t					t	t	g	c	a				g								c		
h20	75	c	g	a	g	c	t	a	g	t	t					t	t	g	a	a				g								c		
h21	78	c	g	a	g	c	t	a	g	g	t					t	t	g	a	a				g								c		
h22	79	c	g	a	g	c	t	a	g	t	t					t	t	g	a	a				g								c		
h23	77	c	g	a	g	c	t	a	g	g	t					t	t	g	c	a				g								c		
h24	28	c	g	a	g	c	t	a	g	g	t					t	t	g	a	a				g								c		
h25	9	c	g	a	g	c	t	a	g	g	t					t	c	t	a	a	-10[d]	-g		g				+20[d]				c		
	apple	t	a	t	t	a	a	t	a+t	g	c					a	c	g	a	t				g								c		+6[d]

nd The inside region of Gap4

[a] Numbers refer to the code number in Table 1

–[b] Eleven base changes and 2 gaps from apple were used as out of group for haplotype network analysis

[c] Numbering starts from the first base of *P. pyraster* harbouring the longest sequence in the intergenic region between *acc*D and *psa*I. Apple was used as out of group for haplotype network analysis

[d] Gaps harbouring direct repeat motifs

Table 5 Characterization of 14 haplotypes by insertions (+)/deletions (−) and base changes found in the 904-bp intergenic region between *rps*16 and *trnQ* in *Pyrus* species by DNA sequencing

Haplo type	Code[a]	Base change (BC) and gap (G)																					
		–[b]	BC1 216[c]	G1 241[c]	–[b]	G2 343[c]	–[b]	–[b]	G3 491[c]	BC2 504[c]	G4 534[c]	BC3 612[c]	G5 627[c]	G6 634[c]	–[b]	G7 676[c]	–[b]	G8 746[c]	–[b]	G9 763[c]	G10 786[c]	G11 855[c]	G12 884[c]
H1	3, 11, 12, 22	C	C		G			T	−141	nd‡	nd‡	nd‡	nd‡	nd‡	A		T		C				
H2	14, 17~21	C	C		G			T	−141	nd‡	nd‡	nd‡	nd‡	nd‡	A		T		C				+16[d]
H3	23	C	C		G			T	−141	nd‡	nd‡	nd‡	nd‡	nd‡	A		T		C				
H4	24	C	C		G			T	−141	nd‡	nd‡	nd‡	nd‡	nd‡	A		T		C				
H5	15	C	C		G			T	−141	nd‡	nd‡	nd‡	nd‡	nd‡	A		T		C				
H6	2,8	C	C		G			T	−141	nd‡	nd‡	nd‡	nd‡	nd‡	A	−24[d]	T		C				
H7	16	C	C		G			T	−141	nd‡	nd‡	nd‡	nd‡	nd‡	A		T		C				+16[d]
H8	4	C	C		G			T	−141	nd‡	nd‡	nd‡	nd‡	nd‡	A		T		C			+42[d]	
H9	5	C	C	−28[d]	G			T	−141	nd‡	nd‡	nd‡	nd‡	nd‡	A		T	+C	C				
H10	6	C	T		G			T	−141	nd‡	nd‡	nd‡	nd‡	nd‡	A		T		C				
H11	7	C	C		G			T	−141	nd‡	nd‡	nd‡	nd‡	nd‡	A		T		C				
H12	53	C	C		G			T		T		A			A		T		C	+24[d]			
H13	13,	C	C		G			T		T		A			A		T		C	+24	+23[d]		
H14	10	C	C		G			T		T		A			A		T		C	+24[d]			
H15	29, 30, 32, 33, 52, 58	C	C		G			T		T		A			A		T		C	+24[d]			
H16	76	C	C		G			T		T	-T	A	+A		A		T		C	+24[d]			
H17	31,	C	C		G			T		T		A			A		T		C				
H18	34, 35, 54, 57	C	C		G			T		T		A			A		T		C				
H19	25~27, 74,	C	C		G			T		G		A			A		T		C				
H20	75	C	C		G			T		G		A		-T	A	−24[d]	T		C				
H21	78	C	C		G			T		G		A			A	−24[d]	T		C				
H22	79	C	C		G			T		G		A			A	−24[d]	T		C				
H23	77	C	C		G			T		G		A			A	−24[d]	T		C				
H24	28	C	C		G			T		T		A			A		T		C				
H25	9	C	C		G	−12		T		G		C			A		T		C				
	Apple	A	C	−28[d]	A		+A	G		G		A			G		C		A				

Apple was used as out of group for haplotype network analysis

nd The insides region of Gap3

[a] Numbers refer to the code number in Table 1

–[b] Six base changes and a gap from apple were used as out of group for haplotype network analysis

[c] Numbering starts from the first base of *P. korshinskyi* harbouring the longest sequence in intergenic region between *rps*16 and *trn*Q

[d] Gaps including direct repeat motifs

Table 6 Direct and inverted repeats detected in a gap

Gap	Sequence	Length of a repeat unit (bp)	No. of repeats	Code
*acc*D-*psa*I				
G1	AAAGAGTGAATTCTTTC	17	1	5
G3	AAAGAGTGAATTCTTTC	17	3	7
	AAAGAGTGAATTCTTTC	17	2	All except for 5 and 7
G5,	ATGTAGAAAC	10	4	3~5, 9, 11, 12, 14, 15, 17~22
G7	ATGTAGAAAC	10	2	16, 23
	ATGTAGAAAC	10	3	All except for the above
G10	TTATTTATATATTTAT[a]	16	2	15
G10, G12	TTTTATTAATTAATATA	17	5	15
G11	TTTTATTAATTAATATA	17	3	2~9, 11, 12, 14~16, 17~24, 34, 35, 54, 57
G13	TTTTATTAATTAATATA	17	2	All except for the above
G14	TATCCT	6	2	4
	TATCCT	6	1	All except for the above
G15	AGGTATACTTAGTATAATA	19	2	10
	AGGTATACTTAGTATAATA	19	1	All except for the above
*rps*16-*trn*Q				
G1	ATCTTCAAGTAACGTGATAGGATAAATT	28	1	5
	ATCTTCAAGTAACGTGATAGGATAAATT	28	2	All except for 5
G7	AGAAATAAGAATCACATTCTATAA	24	1	2, 8, 75, 77~79
	AGAAATAAGAATCACATTCTATAA	24	2	All except for the above
G9, G10	ATCTTTATTTTGATCTATTTTATT	23	3	13
	ATCTTTATTTTGATCTATTTTATT	23	2	10, 13, 29, 30, 32, 33, 52, 53, 58, 76
	ATCTTTATTTTGATCTATTTTATT	23	1	All except for the above
G11	GATATAGATTATGATATAGATATATATTTTATTTTTTATTA	42	2	4
	GATATAGATTATGATATAGATATATATTTTATTTTTTATTA	42	1	All except for 4
G12	TATAGATTAATAAAAT	16	2	14, 16~21
	TATAGATTAATAAAAT	16	1	All except for the above

[a] Inverted repeats between *acc*D-*psa*I

deletion allowed Occidental pears to be categorised as type C and distinguished them from oriental pears (categorised as types A and B). Type A was grouped by seven haplotypes, H18–H25, including wild ussurian pears (*P. ussuriensis*), pea pears (*Pyrus betulaefolia*, *Pyrus fauriei*, and *P. dimorphophylla*), two Japanese pear cultivars (*P. pyrifolia*), and two Chinese cultivars (*P. ussuriensis*), and *P. regelii*. H18 and H19 were the most common haplotypes in type A. Apple, H19, and H20–H23 were branched out from median vector 1 (mv1). H18 was branched out from mv2, and H25 branched out from mv3, 4, and 5. The origin of haplotypes derived from median vectors is still unclear. Finally, 16 accessions (62%) from ussurian pears including wild and cultivated ones and all of the pea pears were included in type A by sequencing and/or PCR analyses (Table 1).

Type B was composed of six haplotypes, H12–H16, including Japanese pear cultivars (*P. pyrifolia*), Chinese sand pear and white pear (*P. pyrifolia* and *P. bretschneideri*), *P. pashia*, *P. korshinskyi* (H13), and *P. ussuriensis* var. *aromatica* (H16) (Fig. 2). Although *P. korshinskyi* belonged to Occidental pears according to their distribution, the 141-bp deletion typical of Occidental pears was not detected although the 228-bp deletion was. Thus, *P. korshinskyi* was involved in the type B. H15 was the most common in type B. H16, a local variety 'Sotoorihime' known as a derivative of *P. ussuriensis* var. *aromatica*, was branched out from H15. Type B was separated from type A by the large deletion of 228 bp in *acc*D-*psa*I. There was no median vector in type B. Therefore, type B might be derived from type A, because H17 in type B connected to H24 in type A.

Type C was composed of 11 haplotypes, H1–H11, and mainly contained European pear cultivars and wild relatives native to WC Asia, Europe, and Russia (Fig. 2). The main haplotypes in European pear cultivars (*P. communis*) were H1 and H2 in type C. H2 included most of the European pear cultivars (*P. communis*), but some cultivars dispersed in H1, H3, and H4. H1 was composed from *Pyrus cordata*, *Pyrus amygdaliformis*, and *P. communis* 'General leclerc'.

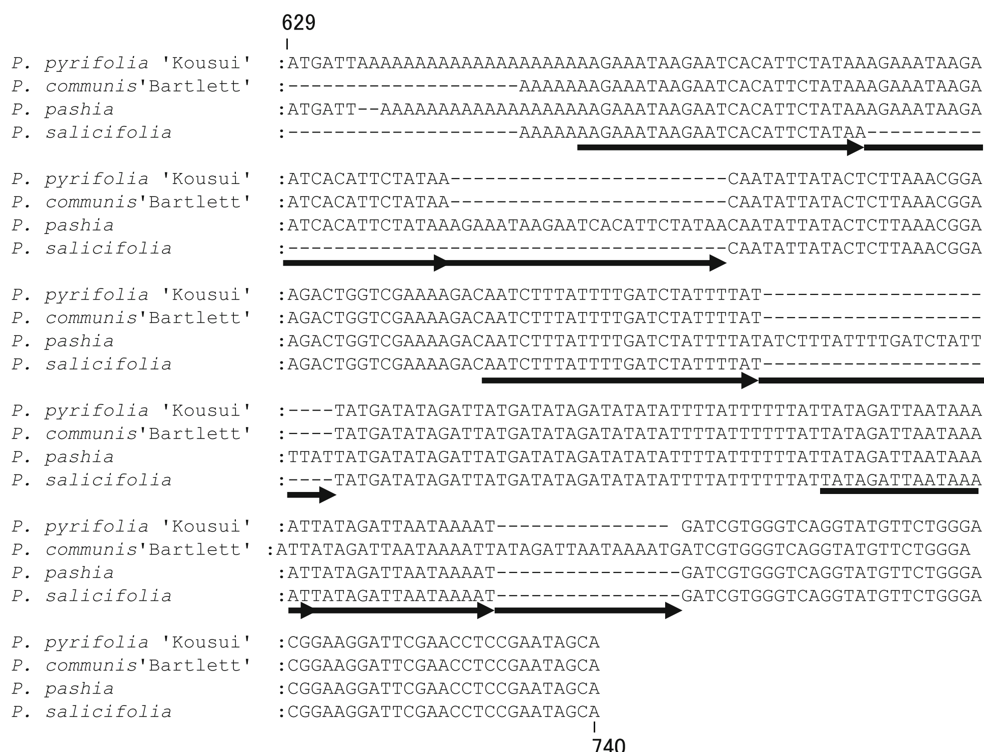

Fig. 1 Sequence alignment of intergenic region between *rps*16-*trn*Q among four species in *Pyrus*. Numbering starts from the first base of *P. korshinskyi*. *Arrowheads* and *dashed lines* mean short direct repeat regions and gaps

H3 (*P. communis* 'Le lectier'), H5 (*P. pyraster*), H8 (*P. elaeagrifolia*), and H9 (*P. cossonii*) were branched out from H1. Therefore, H1 might be ancestral to these haplotypes. *P. caucasica* (H7) was branched out from H2 (*Pyrus balansae* and five cultivars of *P. communis*). H6 (*P. nivalis* and *P. salicifolia*), H10 (*P. gharbiana*), and H11 (*P. mamorensis*) were branched out from H4 (*P. communis* 'Marguerite Marillat'). The haplotype of *P. nivalis* was coincident with that of *P. salicifolia.* There was no median vector within type C. Type C connected to type A via mv3, 4, and 2.

Table 7 Length of mononucleotide found in breakpoints of G3 between *rps*16 and *trn*Q in *Pyrus*

Code[a]	*rps*16-*trn*Q 5′-G3, 483[b]	Code[a]	*rps*16-*trn*Q 3′-G3 635[b]
9, 10, 13, 29~33, 52~54, 58, 76	A9	25	A14
75, 77	A10	26, 27, 74	A16
74	A11	75, 77	A18
		10, 13	A19
		9, 32, 52	A20
		29~31, 33~35, 53, 54, 57, 58, 76	A21

[a] Numbers refer to the code number in Table 1

[b] Numbering starts from the first base of *P. korshinskyi* harbouring the longest sequence in intergenic region between *rps*16 and *trn*Q

Discussion

Two hypervariable regions in the conserved pear cp genome

In the previous study, we reported high conservatism of pear cp genome in comparison with those of other angiosperms (Katayama and Uematsu 2003). In contrast to the previous result, our present study revealed the two intergenic regions of *acc*D-*psa*I and *rps16-trnQ* were extremely hypervariable. The intergenic region between *rbc*L-*cem*A (ca. 3 kb for Poaceae and ca. 5 kb for tobacco) which includes *acc*D-*psa*I has been reported as a hyper-variable region and might represent an intra-molecular recombinational hotspot mediated by short direct repeats in the cpDNA of Poaceae (Ogihara et al. 1988; Hiratsuka et al. 1989; Maier et al. 1995; Guo and Terachi 2005). The 228-bp largest deletion, occurring in the intergenic region between *acc*D-*psa*I, was reported by RFLP analysis based on physical mapping and DNA sequencing in Japanese pears (Katayama and Uematsu 2003; Kimura et al. 2003). A 13-bp direct repeat was detected only in the left border of the 228-bp deletion in this study. This deletion might be a result of intra-molecular recombination mediated by the short direct repeat.

The recombination frequencies of large gaps were lower than those of short gaps (Hamilton et al. 2003; Ingvarsson et al. 2003; Yamane et al. 2006). The 228-bp deletion was highly stable and seemed to arise at the interspecific level.

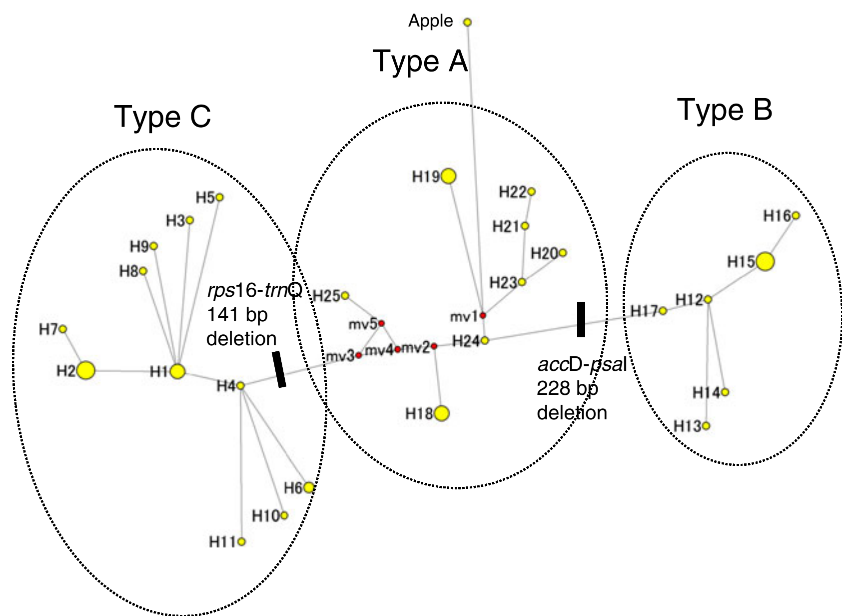

Fig. 2 Median-joining network for 25 cpDNA haplotypes in *Pyrus*. The haplotypes are indicated by the *circles*, the size of each circle being proportional to the observed frequency of each haplotype. Each node of haplotype and median vectors are labelled as H and mv. Two large deletions found in intergenic regions of *acc*D-*psa*I and *rps*16-*trn*Q are indicated by *black rectangles*. *Three dashed circles* (types A, B, and C) are grouped by two large deletions

The other gaps found in this region were rearranged in a unit of repeat sequences (Table 6). In contrast to the length mutation rate, ranging from 0.004 to 0.0132 with an average of 0.0062, the base change rate was relatively low ranging from 0.0003 to 0.0021 with an average of 0.0006 for the 21 species in *Pyrus* (Table 3). Although the intergenic region between *acc*D-*psa*I was hypervariable with respect to length mutation, the intergenic region between *rbc*L-*acc*D was highly conservative in *Pyrus* unlike that in Poaceae (Ogihara et al. 1991). The base change rate, ranging from 0.0006 to 0.0024 with an average of 0.0009, in the region between *rbc*L-*acc*D was quite low and equivalent to those of the other two intergenic regions, but the length mutation rate, ranging from 0.0009 to 0.0016 with an average of 0.001, was extremely low (Table 3). It is consistent with results indicating low genetic diversity of cpDNA in *Pyrus* by our previous study (Katayama and Uematsu 2003). It was lower than those of other taxa, i.e. 0.046–0.09 of *rbc*L-pseudo *rpl*23 hypervariable region in *Triticum–Aegilops*, 0.0055–0.333 of *trn*L-*trn*F intergenic region in Crassulaceae, and 0.0193 of *rpl*20-*rps*18 intergenic region in *Pinus* (Ogihara et al. 1991; Van Ham et al. 1994; Wang et al. 1999). Therefore, the hypervariable region that was observed in Poaceae may exist in the intergenic region between *acc*D-*psa*I in *Pyrus*. An intergenic region between *rps*16-*trn*Q genes was reported as one of the hypervariable regions in cpDNA of angiosperms (Shaw et al. 2007). Earlier we reported a hypervariable region in cpDNAs of *Prunus* (Rosaceae), which was about 9.1 kb and included *psb*A to *atp*A near the left border of LSC (Katayama and Uematsu 2005). This highly variable region in the *Prunus* cpDNA may contain a recombinational hotspot region between the *psb*A-*atp*A genes, since short direct repeats of an 8-bp specific motif (GTTATTTA) and an 11-bp length T-stretch (microsatellite sequence in cpDNA) were found inside the breakpoint of the 277-bp deletion found in all *Prunus* species examined except for *P. persica*. As a result of a survey for the same region in *Pyrus*, a number of gaps were detected in the intergenic region between *rps*16-*trn*Q (Table 5). Mononucleotide repeats (A stretches from 9 to 21 bp) were detected in the left and right borders of the 141-bp largest deletion (Table 7). The large deletion might have been mediated through recombination across mononucleotide repeats. The 141-bp deletion was highly stable and seemed to arise at the interspecific level. Twelve gaps of various lengths in this region seemed to have resulted from rearrangement in a unit of repeat sequences (Table 6, Fig. 1). In contrast with the high rate of length mutation (average of 0.0026), the base change rate was quite low at an average of 0.0001 in this region (Table 3). As the hypervariable region in cpDNA, Van Ham et al. (1994) found a total of 50 small gaps (partly due to mononucleotide repeat variation) in a intergenic *trn*L-*trn*F spacer of 15 species belonging to the families Crassulaceae, Saxifragaceae, and Solanaceae. Twenty-seven gaps were found in two regions in 21 species in *Pyrus* in this study. Both of the two intergenic regions (*acc*D-*psa*I and *rps*16-*trn*Q) were concluded to be hypervariable regions in *Pyrus* cpDNAs.

Haplotype network inferred from gap characters

After haplotype network analysis using all gap and base change characters in two regions, all cpDNA haplotypes were divided into three groups of A, B, and C by two large deletions, one being 228 bp in *acc*D-*psa*I and the other

being 141 bp in *rps*16-*trn*Q. These two deletions were stable and thus could be applicable to phylogenetic utility (Fig. 2).

Group A was composed basically of pea pears (*P. dimorphophylla*, *P. fauriei*, *P. betulaefolia*) and wild ussurian pears (*P. ussuriensis* Maxim. native to East Asia and in Japan). Exceptionally, two Japanese pear cultivars 'Niitaka' and 'Nangetsu' from *P. pyrifolia* were also included in type A. This could be explained by an interspecific hybridization between *P. pyrifolia* and *P. ussuriensis* such as the Chinese ussurian pear. Katayama et al. (2007) have already reported that three Japanese pear cultivars 'Nansui', 'Niitaka', and 'Nangetsu' have type A cytoplasm which might have resulted from interspecific hybridization between *P. pyrifolia* and *P. ussuriensis.*

Type B basically included pear cultivars in East Asia and India (*P. pyrifolia*, *P. bretschneideri*, and *P. pashia*). *P. pashia* was described in the literature as an intermediate in morphology between the Occidental and Oriental pear groups (Ghora and Panigrahi 1995). In the present study, *P. pashia* was found to have a close phylogenetic relationship to Oriental pears. Exceptionally, *P. korshinskyi* (H13) native to SC Asia and Afghanistan and *P. ussuriensis* var. *aromatica* 'Sotoorihime' (H16) were also included. Rehder (1940) considered that *P. korshinskyi* (synonym *P. bucharica* Litv.) was a related species of *P. communis* native to Central Asia. However, in the present study, an accession from *P. korshinskyi* belonged to type B; *P. communis* belonged to type C. Therefore, the cytoplasmic donor for *P. korshinskyi* remains unclear. Population structure analysis of cpDNA haplotype for *P. korshinskyi* will be required, because haplotype determination using a limited number of accessions from species or population has often led incorrect results.

Almost all the *P. ussuriensis* accessions belonged to type A, but 'Sotoorihime' was an exception. 'Sotoorihime' (H16) is known as a local pear variety derived from wild pear (*P. ussuriensis* var. *aromatica*) in the Northern Tohoku Region in Japan. Previously, we reported that about 80% of the 58 accessions of pear, including *P. ussuriensis* var. *aromatica* collected from Northern Tohoku region in Japan, had a 228-bp deletion like the haplotype in type B (Katayama et al. 2007). This suggests the coexistence of *P. ussuriensis* var. *aromatic* and hybrid progeny with *P. pyrifolia.* Recently, Iketani et al. (2010) represented the introgression from *P. pyrifolia* to *P. ussuriensis* var. *aromatica* by population structure analysis using SSR markers. CpDNA in *P. ussuriensis* var. *aromatica* in type A might be changed to that of *P. pyrifolia* in type B by chloroplast capture through interspecific crosses.

Type C included Occidental pears native to Europe, West and Central Asia, Russia, and Africa (*P. communis* including European pear cultivars, *P. nivalis*, *P. amygdaliformis*, *P. elaeagrifolia*, *P. cossonii*, *P. gharbiana*, *P. mamorensis*, *P. salicifolia*, *P. cordata*, *P. balansae*, *P. pyraster*, and *P. caucasica*). The main haplotypes in European pear cultivars (*P. communis*) were H1 and H2. *P. pyraster* (H5) and *P. caucasica* (H7), candidates for the ancestor of European pear cultivars, branched out from H1 and H2 individually. Aldasoro et al. (1996) reported that, phenotypically, *P. pyraster* was similar to *P. caucasica*. But the present study reveals that even though *P. pyraster* (H5) and *P. caucasica* (H7) share a large 228-bp gap (G12 for *acc*D-*psa*I), there are still four gap differences (G7, G8, G10, G13 for *acc*D-*psa*I, and G12 for *rps*16-*trn*Q) between the two accessions. Thus, in view of cpDNA structure, *P. pyraster* and *P. caucasica* might be distantly related. According to Challice and Westwood (1973), *P. cordata* native to the UK and France was grouped with *P. cossonii* and East Asian pears such as pea pears. But in the present study, *P. cordata* (native to Plymouth, UK) was included in H1 with *P. communis* 'General leclerc' and *P. amygdaliformis* both native to Europe. *P. cordata* may have diverged from a common ancestor of *P. communis* such as 'General leclerc' and *P. amygdaliformis*. At least it seems geographically congruent.

The haplotype network revealed that two long deletions were key to distinguishing three types, A, B, and C. The141-bp-long deletion in *rps*16-*trn*Q, particularly, could divide all pear accessions into two groups, i.e. Occidental pears and Oriental pears. Oriental pears were then divided into types A and B by another long deletion of 228 bp in *acc*D-*psa*I. In type A, pea and ussurian pears were involved. In contrast, type B included more domesticated ones, i.e. local or modern cultivars. Therefore, pears in type A are considered to have been more wild and primitive compared to pears in type B. This result agrees with the consideration described by Challice and Westwood (1973) that pea pears were phylogenetically primitive in *Pyrus* based on a chemotaxonomical study using phenolics. Although group B (*P. pyrifolia*, *P. bretschneideri*, and *P. pashia*) may have diverged from ancestral primitives harbouring haplotype H24 in type A via a 228-bp deletion, the origin of type C (Occidental pear) is still unclear because there are many median vectors (mv2, 3, 4, and 5) between type A and C, indicating there are still unknown or disappearing intermediate types. Genetic variations between Occidental and Oriental pears were quite clear (Iketani et al. 1998; Oliveira et al. 1999; Kimura et al. 2003; Zheng et al. 2008). Occidental pears are geographically and genetically distinct from Asian pears and might have evolved and diversified independently. Recurrent mutations such as gaps found in cpDNA easily result in homoplasy and often lead to erroneous phylogenetic relationships (Golenberg et al. 1993; Graham et al. 2000). Hence, it would be better to evaluate the gaps in comparison to the base change in order

to understand whether gaps found in the hypervariable regions are homoplasious mutations or not. In general, longer gaps seemed to be more stable and less homoplasious. This is only a preliminary report that suggests the phylogenetic utility of gaps in the hypervariable region as a cpDNA marker in *Pyrus* which has an extremely conserved chloroplast genome. These gaps could be used as a powerful tool to estimate population structure and gene flow in *Pyrus*.

Acknowledgements Thanks are due to Mr. M. Miyake, Mr. S. Kakehi, Mr. K. Masaki, Mr. T. Takemura, Mr. T. Tanikawa, and Ms. Y. Ohwa in Food Resources Research and Education Center, Kobe University, for maintenance of plant materials. Sincere appreciation is expressed to Dr. Anne Edwards, John Innes Centre, UK, for her English correction and useful suggestions. This work was supported in part by Grant-in-Aid (no. 17510196 and no. 19580031) for Scientific Research from the Ministry of Education, Science and Culture, Japan, and by a General Research Grant from the NIAS Gene Bank.

References

Aldasoro JJ, Aedo C, Garmendia FM (1996) The genus *Pyrus* L. (Rosaceae) in South-West Europe and North Africa. Bot J Linn Soc 121:143–158

Bandelt HJ, Forster P, Roehl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37–48

Bao L, Chen K, Zhang D, Cao Y, Yamamoto T, Teng Y (2007) Genetic diversity and similarity of pear (*Pyrus* L.) cultivars native to East Asia revealed by SSR (simple sequence repeat) markers. Genet Resour Crop Evol 54:959–971

Bell RL (1990) Pears (*Pyrus*). In: Moore JN, Ballington JR Jr (eds) Genetic resources of temperate fruit and nut crops I. International Society for Horticultural Science, Wageningen, pp 655–697

Browicz K (1993) Conspect and chorology of the genus *Pyrus* L. Arboretum Kornickie 38:17–33

Challice JS, Westwood MN (1973) Numerical taxonomic studies of the genus *Pyrus* using both chemical and botanical characters. Bot J Linn Soc 67:121–148

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

Ghora C, Panigrahi G (1995) The family Rosaceae in India. vol. 2. Dehra Dun India. p 481

Golenberg EM, Clegg MT, Durbin ML, Doebley J, Ma DP (1993) Evolution of a non-coding region of the chloroplast genome. Mol Phylogenet Evol 2:52–64

Graham S, Reeves W, Burns PA, Olmstead ACE (2000) Microstructural changes in noncoding chloroplast DNA: interpretation, evolution, and utility of indels and inversions in basal angiosperm phylogenetic inference. Int J Plant Sci 161:S83–S96

Gu C, Sponberg SA (2003) *Pyrus*. In: Wu Z (ed) Flora of China, vol 9. Science Press, Beijing, pp 173–179

Guo CH, Terachi T (2005) Variations in a hotspot region of chloroplast DNAs among common wheat and *Aegilops* revealed by nucleotide sequence analysis. Genes Genet Syst 80:277–285

Hamilton MB, Braverman JM, Soria-Hernanz DF (2003) Patterns and relative rates of nucleotide and insertion/deletion evolution at six chloroplast intergenic regions in New World species of the Lecythidaceae. Mol Biol Evol 20:1710–1721

Hiratsuka J, Shimada H, Whittier R, Ishibashi T, Sakamoto M, Mori M, Kondo C, Honji Y, Cr S, 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

Hosaka K (1995) Successive domestication and evolution of the Andean potatoes as revealed by chloroplast DNA restriction endnuclease analysis. Theor Appl Genet 90:356–363

Iketani H, Manabe T, Matsuta N, Akihama T, Hayashi T (1998) Incongruence between RFLPs of chloroplast DNA and morphological classification in east Asia pear (*Pyrus* spp.). Genet Resour Crop Evol 45:533–539

Iketani H, Yamamoto T, Katayama H, Uematsu C, Mase N, Sato Y (2010) Introgression between native and prehistorically naturalized (archaeophytic) wild pear (*Pyrus* spp.) populations in Northern Tohoku, Northeast Japan. Conserv Genet 11:115–126

Ingvarsson PK, Ribstein S, Taylor DR (2003) Molecular evolution of insertions and deletions in the chloroplast genome of *Silene*. Mol Biol Evol 20:1737–1740

Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HN (ed) Mammalian protein metabolism II. Academic, New York, p 21

Kajiura I (1983) Nihon nashi no kigen to chiriteki bunka. In: Akihama T (ed.), Ikushugaku Saikin no Shinpo. Youkendo, Tokyo. 25:3–13 (in Japanese)

Katayama H, Ogihara Y (1996) Phylogenetic affinities of the grasses to other monocots as revealed by molecular analysis of chlorolast DNA. Curr Genet 20:527–581

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

Katayama H, Uematsu C (2005) Structural analysis of chloroplast DNA in *Prunus* (Rosaceae): evolution, genetic diversity and unequal mutations. Theor Appl Genet 111:1430–1439

Katayama H, Adachi S, Yamamoto T, Uematsu C (2007) A wide range of genetic diversity in pear (*Pyrus ussuriensis* var. *aromatica*) genetic resources from Iwate, Japan revealed by SSR and chloroplast DNA markers. Genet Resour Crop Evol 54:1573–1585

Kikuchi A (1946) Classification of the strains and varieties of the Chinese pear. Studies from the Inst Hort. Kyoto Imper Univ 3:1–11 (in Japanese)

Kikuchi A (1948) Nashi. Kajuengeigaku Jyo-kan, Kajyusyurui kakuron, Yokendo, Tokyo. pp 64–116 (in Japanese)

Kimura T, Iketani H, Kotobuki K, Matsuta N, Ban Y, Hayashi T, Yamamoto T (2003) Genetic characterization of pear varieties revealed by chloroplast DNA sequences. J Hort Sci & Biot 78:241–247

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

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

Nakai T (1919) Notulae ad Plantas Japoniae et Koreae XXI. Bot Mag (Tokyo) 33:193–216

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, Terachi T, Sasakuma T (1991) Molecular analysis of the hot spot region related to length mutations in wheat chloroplast DNAs. I. Nucleotide divergence of genes and intergenic spacer regions located in the hot spot region. Genetics 129:873–884

Ohwi J (1965) Flora of Japan. Smithonian Institution, Washington DC, p 549

Oliveira CM, Mota M, Monte-Corvo L, Goulao L, Silva DM (1999) Molecular typing of *Pyrus* based on RAPD markers. Sci Hort 79:163–174

Rehder A (1920) New species, varieties and combinations. J Arn Arb 2:47–58

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, p 403

Rubtsov GA (1944) Geographical distribution of the genus *Pyrus* and trends and factors in its evolution. Am Nat 78:358–366

Shaw J, Lickey EB, Schilling EE, Small RL (2007) Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. Amer J Bot 94:275–288

Shirai K (1929) Shokubutu toraikou. Okashoten, Tokyo (in Japanese)

Small RL, Ryburn JA, Cronn RC, Seelanan T, Wendel JF (1998) The tortoise and the hare: choosing between noncoding plastome and nuclear *Adh* sequences for phylogeny reconstruction in a recently diverged plant group. Amer J Bot 85:1301–1315

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

Teng Y, Tanabe K, Tamura F, Itai A (2002) Genetic relationships of *Pyrus* species and cultivars native to east Asia revealed by randomly amplified polymorphic DNA markers. J Am Soc Hort Sci 127:262–270

Van Ham RCHJ, Hart H, Mes THM, Sandbrink JM (1994) Molecular evolution of noncoding regions of the chloroplast genome in the Crassulaceae and related species. Curr Genet 25:558–566

Volk GM, Richards CM, Henk AD, Reilley AA, Bassil NV, Postman JD (2006) Diversity of wild *Pyrus communis* based on microsatellite analysis. J Amer Soc Hort Sci 131:408–417

Wang XR, Tsumura Y, Yoshimaru H, Nagasaka K, Szmidt AE (1999) Phylogenetic relationships of Eurasian pines (*Pinus*, Pinaceae) based on chloroplast *rbc*L, *mat*K, *rpl*20-*rps*18 spacer, and *trn*V intron sequences. Amer J Bot 86:1742–1753

Yamane K, Yano K, Kawahara T (2006) Pattern and rate of indel evolution inferred from whole chloroplast intergenic regions in sugarcane, maize and rice. DNA Res 13:197–204

Yu TT (1979) Taxonomy of the fruit tree in China. Agriculture Press, Beijing (in Chinese)

Zheng X, Cai D, Yao L, Teng Y (2008) Non-concerted ITS evolution, early origin and phylogenetic utility of ITS pseudogenes in *Pyrus*. Mol Phyl Evol 48:892–903