A. Mohanty · J.P. Martín · I. Aguinagalde Chloroplast DNA study in wild populations and some cultivars of Prunus avium L.

Received: 10 July 2000 / Accepted: 19 October 2000

Abstract The PCR-RFLP technique was used to detect chloroplast DNA diversity in wild populations of *Prunus avium* from five European deciduous forests and some cultivars. A study of 10.8% of the total chloroplast genome detected eight insertion-deletion (indel) mutations, distributed over 12 haplotypes. Six haplotypes (H1, H2, H3, H4, H5 and H6) were found in wild populations and eight (H2, H6, H7, H8, H9, H10, H11 and H12) in the cultivars. Only two haplotypes (H2 and H6) are shared by the wild populations and the cultivars. The mostabundant and frequent haplotype in wild populations is H2 (frequency=78%). The wider geographical distribution along with the high frequency reflects its ancient origin. Of the five populations, three are polymorphic. Populations GA (Scotland) and KE (Germany) have unique haplotypes. The total cpDNA diversity in wild populations is $h_T=0.40$, and a major portion of it is within populations $(h_s=0.37)$. The genetic differentiation among populations was low $(G_{STC}=0.08)$ and no genetic structure among wild populations was observed. A minimum-length spanning tree, demonstrating relationships among the haplotypes in wild populations, indicated two possible chloroplast lineages. The ten identified cultivars were represented by seven haplotypes; this result proposes the possible utilisation of the PCR-RFLP technique for the characterisation of sweet cherry cultivars. The cpDNA diversity in *P. avium* should be considered carefully for phylogenetic studies involving this species.

Keywords *Prunus avium* L. · cpDNA diversity · PCR-RFLP · Universal primers · Sweet cherry cultivars

Communicated by P.M.A. Tigerstedt

A. Mohanty (\mathbb{X})

Departamento de Biología Vegetal, Escuela Técnica Superior de Ingenieros Agrónomos, Universidad Politécnica de Madrid, Ciudad Universitaria, 28040-Madrid, Spain e-mail: jpmartin@bio.etsia.upm.es Fax: +34-91-3365656

Introduction

Studies on population structure and genetic diversity are essential for the conservation and management strategies of the genetic resources of any species. Chloroplast DNA (cpDNA) diversity has been successfully used for analysing the population structure and for phylogeographic studies of several species (Byrne and Moran 1994; Demesure et al. 1996; El Mousadik and Petit 1996; Levy et al. 1996; King and Ferris 1998; Mohanty et al. 2000). Also, considering the maternal inheritance of chloroplasts in most angiosperms [including *Prunus avium* (Heinze, personal communication)], assessment of cpDNA variation in wild populations and their cultivated forms provides good information on their maternal relationships and the extent of the cpDNA diversity shared between them (reviewed by Doebly 1992; Amane et al. 1999). The detection of intraspecific cpDNA variations has been facilitated using the PCR-RFLP (polymerase chain reaction – restriction fragment length polymorphism) technique, which employs universal cpDNA primers (Taberlet et al. 1991; Demesure et al. 1995; Dumolin-Lapegue et al. 1997). This technique has gained popularity for its easy accessibility and rapidity. Moreover, it detects a larger number of variations as at least 50% of the cpDNA variations are attributable to small insertions and deletions (Gielly and Taberlet 1994).

P. avium L. grows wild throughout the temperate regions of mainland Europe and into southern Russia; however, maximum concentration is between and south of the Caspian and Black seas (Beaver et al. 1995). It is a diploid (2n=2*x*=16), strictly allogamous, insect-pollinated species (Crane and Lawrence 1931; Crane and Brown 1937). Vegetative reproduction by suckering occurs naturally (Frascaria et al. 1993). Seeds are ingested and dispersed by animals and also by human activities. This species is economically important as a timber tree, fruit tree and as a rootstock species. Turok et al. (1996) has included this species in the list of noble hardwood trees. Intraspecific and interspecific hybridizations have been carried out in this species for genetic improvement by

Table 1 Geographical origin and number of individuals in the wild populations of *P. avium* used in the study

Population	Abbrev.	Origin	No. of individuals
Montejo de la Sierra Chizé Parco Nazionale delle Foreste Casentinesi Kelheim Glen Affric	MS CН FC. KE GA	Spain France Italy Germany Scotland	10 10 10 10

Table 2 Preliminary results showing PCR conditions, size of amplified fragment and quality of amplification, using cpDNA universal primers in *P. avium*

^a Abbreviations are the same as in Dumolin-Lapegue et al. (1997)

^b NA: no amplification; +: faint amplification; ++: good amplification

introducing additional desirable characters (Tydeman and Garner 1966; Tobutt 1985; Nicoll 1993).

Most of the genetic variability studies in both wild and cultivated forms of this species have been carried out using nuclear DNA markers such as isozymes (Frascaria et al. 1993; Beaver et al. 1995; Mariette et al. 1997). The only study which used cpDNA markers was that of Heinze (1999), where populations of *P. avium* from eastern Europe were analysed and three haplotypes were found. In the present investigation we analysed wild populations of *P. avium* collected from deciduous forests of western Europe and found six haplotypes. The main objectives were: (1) to study the population structure of the chloroplast genome in wild *P. avium*, and (2) to compare the cpDNA diversity (haplotypes) between some sweet cherry cultivars and wild cherries. The PCR-RFLP technique was used for detecting variation in the chloroplast genome.

Material and methods

Plant material

Five wild populations of *P. avium* were sampled from five deciduous forests across western Europe (Table 1). The distance between individuals in every population was at least 200 m. Ten identified [Ambrunesa, Variety 3/13, Celeste, Somerset (origin: Spain), Barnard (origin: France), Bing, 46ED, 106EB325, 4E50 and Van (origin: USA)] and five unidentified cultivars (labelled as UN1 to UN5; collection site: Sierra de Ayllon, Madrid, Spain) were used for comparative study. Fresh leaves were collected from plants in the field, frozen and transferred to the freezer at –80°C.

DNA extraction, amplification and digestion

DNA was extracted from frozen leaves following the protocol of Torres et al. (1993). Extracted DNA was quantified and a working solution of DNA $(4 \nvert g/\mu l)$ was made.

Preliminary amplification of cpDNA was done using 16 pairs of universal primers (Table 2); all are described in Dumolin-Lapegue et al. (1997). The amplifications were performed in 30 µl of a reaction mixture consisting of 0.2 µM of each primer, 200 µM of each of the four dNTPs, 2 mM of $MgCl_2$, 0.5–1.0 U of *EcoTaq* DNA polymerase in the buffer provided by the manufacturers of the enzyme (ECOGEN, S.R.L.), and 12 ng of genomic DNA. The PCR amplifications were carried out in a PTC-100 thermal cycler (MJ Research, Inc.) with a heated lid, using an initial cycle of 4 min at 94°C, followed by 30 cycles of 45 s at 94°C, 45 s at 45°C to 62°C and 2 min to 5 min 30 s at 72°C (annealing temperature and extension time depending on the length of the fragment to be amplified; Table 2), and finally a 10 min extension at 72°C.

Three restriction enzymes (*Alu*I, *Hin*fI, *Taq*I; Amersham) were used for the digestion of PCR products (8.5 µl/enzyme). The reaction mixture (20 µl) was incubated for 18 h with two units of *Alu*I or *Hin*fI, or five units of *Taq*I. Restriction fragments were separated on 2.6% agarose gels cast in a Tris-borate-EDTA buffer (1×), run at 3 V/cm for 4 h, stained with ethidium bromide and visualized in UV light. The digests were also resolved on 8% polyacrylamide gels run in TBE buffer $(1\times)$ at 350 V for 4 to 8 h (the time depending

on the size of fragment to be resolved). These gels were silver stained (Bassam et al. 1991).

The size of the polymorphic bands was analysed with Kodak Digital Science 1D Image Analysis Software. A 50-bp ladder (Pharmacia) was used as a molecular-size marker.

Analysis of data

The HAPLODIV program (Pons and Petit 1995) was used to calculate the frequency of the haplotypes, an estimation of the parameters of cpDNA diversity (h_T =total diversity, h_S =average intrapopulation diversity, and G_{STC} =level of population subdivision) and their standard errors.

A matrix of frequencies of haplotypes in each of the populations was used for generating a matrix of distances between populations selecting the coefficient of "Euclidean distances squared". The latter matrix was employed to construct a dendrogram for assessing relatedness among the five populations by the UPGMA method, using the NTSYS-pc version 1.6 package (Rohlf 1992).

The number of mutational differences between the haplotypes of wild populations was calculated to produce a minimum-length spanning tree of haplotypes employing the programs from NTSYS-pc (Rohlf 1992). The procedure was used to connect points (haplotypes) by direct links having the smallest possible total length (Prim 1957). Minimum spanning networks are alternatives to Wagner parsimony trees, but better convey the connections between haplotypes (Excoffier and Smouse 1994).

Results

Of the 16 pairs of universal primers of cpDNA, five pairs (HK, K1K2, *rpo*CC, CD and VL) were chosen for the present investigation based on the degree of amplification (Table 2). A total of 16,700 bp (10.8%) of the chloroplast genome was analysed [cpDNA size in *P. avium* is 155,000 bp (Heinze, personal communication). Each of the five amplified fragments were digested by three restriction enzymes and, of the 15 combinations, seven (HK-*Hin*fI, K1K2-*Alu*I, K1K2-*Hin*fI, K1K2-*Taq*I, CD-*Hin*fI, CD-*Taq*I and VL-*Hin*fI) exhibited resolivable polymorphic patterns and hence were selected for the complete survey. The fragment amplified by the primer pair *rpo*CC showed monomorphic patterns with the three restriction enzymes. Out of a total of 51 resolvable fragments 16 (31.4%) were polymorphic. All the polymorphisms were due to insertion-deletion (indel) mutations, which ranged between 5 bp and 15 bp (Table 3). **Table 3** Majority patterns and variants obtained with different primer-restriction enzyme combinations in *P. avium*. Restriction patterns with insertions are indicated by (A) and with deletions by (B)

Figure 1 shows the indel mutation obtained in the combination K1K2-*Hin*fI in some individuals of four wild populations of *P. avium* (MS, CH, GA and KE). In all, 12 haplotypes, which resulted from different combinations of these indel mutations (details can be obtained from the authors), were found in wild and cultivated *P. avium*.

Chloroplast DNA diversity in wild populations

Six haplotypes were found in the five wild populations consisting of a total of 44 individuals (Table 4) and 13.6% of the individuals represent the six haplotypes. Populations CH (France) and FC (Italy) are monomorphic and are represented by the haplotype H2 only. Population KE (Germany) has the maximum number of haplotypes (four $-$ H1, H2, H3 and H4), of which two (H3 and H4) are unique to it (Table 4). Population GA (Scotland) contains three haplotypes (H2, H5 and H6) of which H5 and H6 are unique. No unique haplotypes were found in population MS (Spain). H2 is the most common haplotype, present in all five populations. As calculated by the HAPLODIV program, the frequency of this haplotype is highest (0.78) followed by H1 (0.08), and the frequency of the unique haplotypes is 0.02 or 0.05 (Table 4).

Table 4 Haplotype frequencies and composition of the five wild populations of *P. avium*

^a MS=Montejo de la Sierra; CH=Chizé; FC=Parco Nazionale delle Foreste Casentinesi; KE=Kelheim; GA=Glen Affric ^b Calculated by the HAPLODIV program (Pons and Petit 1995)

Fig. 2 Dendrogram of five populations of *P. avium* generated by UPGMA clustering analysis using the "Euclidean distances squared" coefficient. *MS*=Montejo de la Sierra; *CH*=Chizé; *FC*=Parco Nazionale delle Foreste Casentinesi; *KE*=Kelheim; *GA*=Glen Affric

The dendrogram obtained by the UPGMA method using the frequency of haplotypes in each population showed population GA (Scotland) separated from the rest of the populations by a maximum dissimilarity level of 30% (Fig. 2). Populations CH (France) and FC (Italy), which are identical, form one subgroup. The other subgroup consists of MS (Spain) and KE (Germany), which differ by 6%. The two subgroups are separated by a dissimilarity level of 15%. The total genetic diversity is $h_T=0.40$. A high proportion of this diversity is within populations ($h_S=0.37$). The genetic differentiation among populations is G_{STC} =0.08.

The minimum-length spanning tree showing the phylogenetic relationships between the six haplotypes is given in Fig. 3. Haplotypes H1 and H2 form two nodes of the tree. H3 and H5 are ends of the branches related to H1, and H4 and H6 are closely related to the node H2.

Chloroplast DNA diversity in cultivars

The nature of the mutations in the cultivars was the same as that of the individuals of the wild populations. In all, eight haplotypes were found in the ten identified cultivars and five unidentified cultivars. Three haplotypes are common between some pairs of cultivars: H9 in 'Bing' and 'Celeste', H11 in '46ED' and '106EB325', and H12 in '40E50' and 'Van'. Of the five unidentified cultivars, two individuals have the same haplotype (H6), two have H7 which correspond to the cultivar 'Ambrunesa', and one individual has the haplotype H9, corresponding to 'Bing' and 'Celeste'.

Comparison of haplotypes between wild populations and cultivars

Out of a total of 12 haplotypes, four (H1, H3, H4 and H5) are present only in wild populations, six (H7, H8, H9, H10, H11 and H12) only in cultivars, and the remaining two haplotypes (H2 and H6) are common between them. Considering each of the two common haplotypes: H2 is shared by all wild populations and the cultivar 'Barnard'; H6 is common between GA (Scotland) and two unidentified cultivars.

Discussion

From the present investigation, only one of the six haplotypes, i.e. H2, dominates all the wild populations and has the highest frequency. Its wider geographical distribution along with its high frequency reflects its ancient origin, whereas low frequencies of unique haplotypes indicate that they may be the result of recent mutations.

The UPGMA dendrogram demonstrates that the southern European populations are more closely related to each other and are clearly separated from the northern European population GA (Scotland). This is in agreement with our results on *Prunus spinosa* (Mohanty et al. 2000).

The genetic differentiation among populations in *P. avium* $(G_{STC}=0.08)$ is less than in *P. spinosa* $(G_{STC}=0.19;$ Mohanty et al. 2000), and is much lower compared to other forest species such as *Quercus petraea* (G_{STC}=0.90; Petit et al. 1993a), *Fagus sylvatica* (G_{STC}=0.83; Demesure et al. 1996), *Argania spinosa* (G_{STC}=0.60; El Mousadik and Petit 1996) and *Alnus glutinosa* (G_{STC} =0.87; King and Ferris 1998). In all the

Fig. 3 Minimum-length spanning tree of six cpDNA haplotypes of *P. avium* from five European deciduous forests. The *asterisks* with Roman numerals represent restriction pattern differences between haplotypes for each primer-restriction enzyme combination. *I*=HK-*Hin*fI, *II*=K1K2-(*Alu*I)1, *III*=K1K2-(*Alu*I)2, *IV*=K1K2-*Hin*fI, *V*=K1K2-*Taq*I, *VI*=CD-*Hin*fI, *VII*=CD-*Taq*I, *VIII*=VL-*Hin*fI

latter species the G_{STC} is much higher than the G_{STN} (genetic differentiation among populations using nuclear markers) which, as predicted by Petit et al. (1993b), indicates that maternally inherited cytoplasmic genomes are always comparatively much more structured than the nuclear genome since cytoplasmic gene flow is limited to seed dispersal only.

In *P. avium*, Frascaria et al. (1993) and Mariette et al. (1997) used isozyme markers to study four and six populations respectively, and found low genetic differentiation among the populations $(G_{STN}=0.05$ and $G_{STN}=0.06$, respectively). No significant genetic structure among the populations was observed in these studies. This was in agreement with the low mean G_{STN} for trees with ingested seeds (Hamrick et al. 1992). From our present investigation, even cpDNA markers (which are predicted to be genetically more structured) did not show genetic structuring of the populations (G_{STC} =0.08) in terms of the studied geographical scale. However, it may be noted that the comparison of G_{STC} and G_{STN} is from different populations, and may therefore be treated with caution. One of the main reasons for the absence of genetic structure appears to be cytoplasmic gene flow among populations (i.e. migration of seeds). The migration rate of seeds may be higher than expected, considering that human activities have a greater influence on seeds produced by an edible species.

From the phylogenetic tree of wild haplotypes obtained in the present survey, one possible inference is that there exist two chloroplast lineages represented by haplotypes H1 and H2. H1 is related to H3 and H5 by one and two mutations respectively. H4 and H6 are related to the other lineage (H2) by one mutation each. H4 forms a link between H2 and H1; four mutations separating H4 from H1. Analysis of more populations may provide some missing haplotypes (in addition to H4) which can further relate H1 and H2. In wild populations of *P. avium* from eastern Europe, three possible chloroplast lineages have been described by Heinze (personal communication).

The sweet cherry cultivars showed eight haplotypes. Two (H2 in Barnard and H6 in two unidentified cultivars) of the eight haplotypes are in the wild populations and the rest (H7, H8, H9, H10, H11 and H12) are exclusively in the cultivars. This indicates that the origin (in the maternal line) of the studied cultivars (Ambrunesa, Variety 3/13, Bing, Celeste, Somerset, 40E50, Van, and three unidentified cultivars) was different from the investigated wild individuals. This result is different from most comparative studies between cultivars and their wild relatives, which have mostly demonstrated that the haplotype composition of cultivars is very homogeneous and that cultivars have a common maternal origin with any one, or more, wild haplotypes (Doebly 1992; Amane et al. 1999). In *P. avium*, Heinze (1999) found three haplotypes in wild populations from eastern Europe of which two were shared by the cultivars he analysed. Since we have used primer and restriction enzymes different from Heinze (1999), it is not possible to compare the haplotypes from the two studies. It is probable that some of the cultivars in the present study share haplotypes with wild individuals of eastern European populations, which are closer to the centre of origin of *Prunus*. Keeping in mind the present result, which distinguishes seven haplotypes in the ten identified cultivars, we propose the possible utilisation of the PCR-RFLP technique for classifying cultivars into groups based on their haplotypes or even for the identification of cultivars. Also, it is suggested that phylogenetic studies involving *P. avium* should take into account the intraspecific cpDNA variation existing in wild as well as in cultivated cherries.

Acknowledgements We thank to Dr. Remy J. Petit for providing us with the universal primers and the HAPLODIV program. This research was supported by the EC research program FAIR5- CT97–3795. The experiments comply with the current laws of the country in which the experiments were carried out.

References

- Amane M, Lumaret R, Hany V, Ouazzani N, Debain E, Vivier G, Deguilloux MF (1999) Chloroplast-DNA variation in cultivated and wild olive (*Olea europaea* L.). Theor Appl Genet 99: 133–139
- Bassam BJ, Caetano-Anollés G, Gresshoff PM (1991) Fast and sensitive silver staining of DNA in polyacrylamide gels. Anal Biochem 80:81–84
- Beaver JF, Iezzoni AF, Ramm CW (1995) Isozyme diversity in sour, sweet and ground cherry. Theor Appl Genet 90:847–852
- Byrne M, Moran GF (1994) Population divergence in the chloroplast genome of *Eucalyptus nitens*. Heredity 73:18–28
- Crane MB, Brown AG (1937) Incompatibility and sterility in the sweet cherry, *Prunus avium* L. J Pomol Hort Sci 15:80–116
- Crane MB, Lawrence WJC (1931) Sterility and incompatibility in diploid and polyploid fruits. J Genet 24:97–107
- Demesure B, Comps B, Petit RJ (1996) Chloroplast DNA phylogeography of the common beech (*Fagus sylvatica* L.) in europe. Evolution 50:2515–2520
- Demesure B, Sodzi N, Petit RJ (1995) A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. Mol Ecol 4:129–131
- Doebley J (1992) Molecular systematics and crop evolution. In: Soltis PS, Soltis DE, Doyle JJ (eds) Molecular systematics of plants. Chapman and Hall, New York, pp 202–222
- Dumolin-Lapegue S, Pemonge M-H, Petit RJ (1997) An enlarged set of consensus primers for the study of organelle DNA in plants. Mol Ecol 6:393–397
- El Mousadik A, Petit RJ (1996) Chloroplast DNA phylogeography of the argan tree of Morocco. Mol Ecol 5:547–555
- Excoffier L, Smouse PE (1994) Using allele frequencies and geographic subdivision to reconstruct gene trees within a species: molecular variance parsimony. Genetics 136:343–359
- Frascaria N, Santi F, Gouyon PH (1993) Genetic differentiation within and among populations of chestnut (*Castanea sativa* Mill.) and wild cherry (*Prunus avium* L.). Heredity 70: 634–641
- Gielly L, Taberlet P (1994) The use of chloroplast DNA to resolve plant phylogenies: non-coding vs RbcL sequences. Mol Biol Evol 11:769–777
- Hamrick JL, Godt MJW, Sherman-Broyles SL (1992) Factors influencing levels of genetic diversity in woody plant species. New For 6:95–124
- Heinze B (1999) Molecular genetic investigation in wild and cultivated *Prunus avium* in Austria and beyond. In: International Congress: Applications of Biotechnology to Forest Genetics. 22–25 September, Vitoria-Gasteiz, Spain, pp 77–80
- King RA, Ferris C (1998) Chloroplast DNA phylogeography of *Alnus glutinosa* (L.) Gaertn. Mol Ecol 7:1151–1161
- Levy F, Antonovics J, Boyton JE, Gilham NW (1996) A population genetic analysis of chloroplast DNA in *Phacelia.* Heredity 76: 143–155
- Mariette S, Lefranc M, Legrand P, Ttaneyhill D, Frascaria-Lacoste N, Machon N (1997) Genetic variability in wild cherry populations in France. Effects of colonizing processes. Theor Appl Genet 94:904–908
- Mohanty A, Martín JP, Aguinagalde I (2000) Chloroplast DNA diversity within and among populations of the allotetraploid *Prunus spinosa* L. Theor Appl Genet 100:1304–1310
- Nicoll FJ (1993) Genetic improvement of cherry for farm woodlands. Quart J For 87:187–194
- Petit RJ, Kremer A, Wagner DB (1993a) Geographic structure of chloroplast DNA polymorphisms in European oaks. Theor Appl Genet 87:122–128
- Petit RJ, Kremer A, Wagner DB (1993b) Finite island model for organelle and nuclear genes in plants. Heredity 71:630–641
- Pons O, Petit RJ (1995) Estimation, variance and optimal sampling of gene diversity. 1. Haploid locus. Theor Appl Genet 90:462–470
- Prim RC (1957) Shortest connection networks and some generalizations. Bell Labs Tech J 36:1389–1401
- Rohlf FJ (1992) NTSYS-PC: mumerical taxonomy and multivariate analysis system. Exeter Software, New York
- Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol Biol 17:1105–1109
- Tobutt KR (1985) New approaches to breeding sweet cherry scion varieties at East Malling, with particular reference to small tree size. Acta Hort 169:43–50
- Torres AM, Weeden NF, Martín A (1993) Linkage among isozyme, RFLP and RAPD markers in *Vicia faba*. Theor Appl Genet 85:937–945
- Turok J, Eriksson G, Kleinschmit J (1996) Noble hardwoods network. Report 1st Meeting, 24–27 March, Escherode, Germany. International Plant Genetic Resources Institute, Rome, Italy
- Tydeman HM, Garner RJ (1966) Breeding and testing rootstocks for cherries. Report East Malling Research Station 1965, pp 130–154