Molecular Ecology (2000) **9**, 1113–1122

Chloroplast DNA haplotype variation and population differentiation in *Sorbus aucuparia* L. (Rosaceae: Maloideae)

O. RASPÉ,*† P. SAUMITOU-LAPRADE,† J. CUGUEN† and A.-L. JACQUEMART*
**Unité d'Ecologie et de Biogéographie, Université Catholique de Louvain, Place Croix du Sud 4–5, B-1348 Louvain-la-Neuve, Belgium, and †Laboratoire de Génétique et Evolution des Populations Végétales, UPRESA-CNRS 8016, FR-CNRS-1818, Université de Lille 1, Bât. SN2, F-59655 Villeneuve d'Ascq Cedex, France*

Abstract

Intra-specific chloroplast DNA (cpDNA) variation was studied in *Sorbus aucuparia* L., an entomophilous, mid- or early successional tree producing fleshy fruits. Eight PCR-amplified fragments of the chloroplast genome were screened for restriction fragment length polymorphisms, using one or two 4 bp-cutter restriction endonucleases. cpDNA variation was investigated on two geographical scales: (1) among four regions in France and Belgium; and (2) within the Belgian region. A total of 150 individuals from six populations were analysed. Fourteen polymorphisms were detected in six of the cpDNA fragments. All polymorphisms probably resulted from insertions or deletions, and allowed the identification of 12 haplotypes. The level of genetic differentiation computed on the basis of haplotype frequencies was similar on the two geographical scales considered (G_{STc} = 0.286 among regions, G_{STc} = 0.259 among populations within the Belgian region). These values are much lower than those obtained in nine previously studied temperate tree species, which are all wind-pollinated, late-successional species producing dry fruits. These results might primarily be accounted for by the contrasting life history traits of *S. aucuparia*. In order to obtain insights into the relative contribution of pollen and seeds to gene flow, G_{STc} was also compared with previously obtained G_{ST} estimates based on allozyme data.

Keywords: cpDNA, gene flow, PCR–RFLP, rowan, seed dispersal

Received 2 December 1999; revision received 13 March 2000; accepted 13 March 2000

Introduction

During the last decade, chloroplast DNA (cpDNA) variation has received growing attention in the field of plant evolutionary biology. This is because of the markedly different characteristics of cpDNA compared to the nuclear genome, as well as the recent development of efficient molecular tools. The most striking and general differences between chloroplast and nuclear genomes are that the former show uniparental inheritance in most plant species (Reboud & Zeyl 1994; Birky 1995), a clonal mode of evolution and a slow rate of evolutionary change (Wolfe *et al*. 1987). These properties have been exploited in studies dealing with a variety of aspects of plant evolutionary biology such as phylogeny (e.g. Palmer *et al*. 1988; Gielly & Taberlet 1994; Wagstaff *et al*. 1998), inter-specific gene flow and hybridization (e.g. Allan *et al*. 1997) and plant phylogeography (e.g. El Mousadik & Petit 1996a; Petit *et al*. 1997; Soltis *et al*. 1997; King & Ferris 1998; Taberlet *et al*. 1998), as well as to get insights into the relative contribution of pollen and seed movement to intra-specific gene flow (e.g. McCauley 1995; El Mousadik & Petit 1996a; Tarayre *et al*. 1997). Because of the recent development of the two latter aspects, only a few empirical studies have been conducted so far, and more research is needed to allow relevant comparisons and generalizations to be made. In particular, investigations on species with contrasting life forms, reproductive biology or evolutionary history would allow the evaluation of the importance and predictive value of these factors.

Correspondence: Olivier Raspé. Fax: +32 10 47 34 90; E-mail: raspe@ecol.ucl.ac.be

Sorbus aucuparia is a small, insect-pollinated, self-incompatible tree (Raspé 1998). It is widespread in Europe, from Iceland and Northern Russia to the mountains of central Spain, Portugal, Italy and the Caucasus, as well as in Northern Asia minor (Clapham *et al*. 1962). The species produces fleshy fruits which are eaten by birds, the main seed dispersers (Snow & Snow 1988). In many places (mainly at lower altitudes), *S. aucuparia* often behaves as a hardy pioneer or post-pioneer species (Kullman 1986; Rameau *et al*. 1989), populations of which are later replaced by late-successional tree populations. At high altitude, however, it is one of the few species which can maintain the tree habit and its populations may be part of the late-successional vegetation. The life history traits of *S. aucuparia* contrast to some extent with those of most of the previously studied tree species such as oaks, European beech and some gymnosperms, which are all wind-pollinated, late-successional species producing dry fruits.

A previous study of isozyme variation (Raspé & Jacquemart 1998) demonstrated weak genetic differentiation among *S. aucuparia* populations ($G_{ST} = 0.06$), suggesting high levels of gene flow, as in most other tree species (Hamrick *et al*. 1992). In this paper, we present the results of a study which was aimed at assessing the influence of the species life history traits on the population structure of cpDNA variation and on the relative contribution of pollen and seed movement to gene flow.

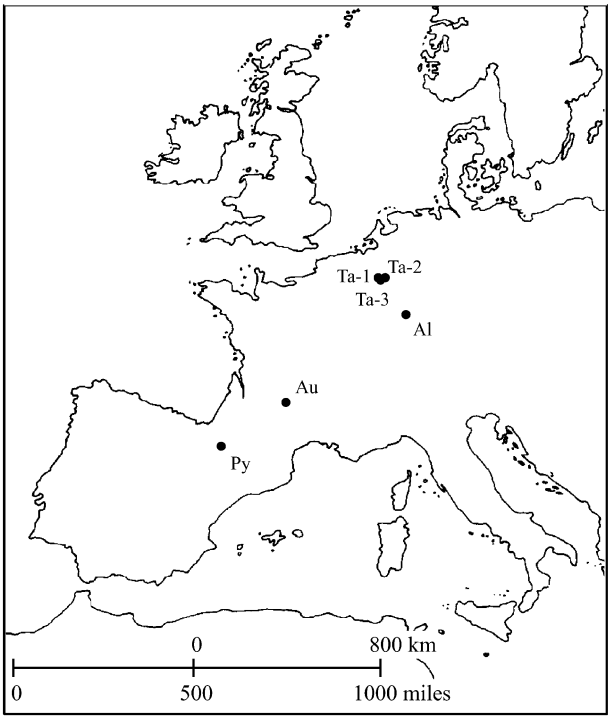

Fig. 1 Map of *Sorbus aucuparia* populations sampled. Py, Pyrénées; Au, Auvergne; Al, Alsace; Ta, Plateau des Tailles (Belgium).

Materials and methods

Sampled populations and DNA isolation

Sorbus aucuparia populations were sampled on two different scales. First, three populations were sampled within a single region ('Plateau des Tailles', Belgium). Second, three additional populations were sampled from three French regions (Pyrénées, Auvergne and Alsace) (Fig. 1 and Table 1). These populations, along with the three Belgian populations pooled together, constitute the inter-regional scale. All these populations had previously been studied with isozyme markers (Raspé & Jacquemart 1998). At each site, twigs bearing several dormant buds were collected from 30–50 trees, brought to the laboratory, and put in water until buds broke out. Part of the young tissues was used for enzyme extraction while the remainder was stored at −80 °C for DNA isolation. Total DNA was extracted following a rapid procedure adapted from Edwards *et al*. (1991). Two major modifications were made: approximately 20–40 mg of tissues were crushed in 200 µL of buffer and 200 µL were added afterwards, and PVP 40 (1% w:v) was added to the buffer described in Edwards *et al*. (1991). Overall, 150 individuals were studied (20 individuals from each French population and 30 individuals from each Belgian population).

Table 1 Description of sampled *Sorbus aucuparia* populations

Population	Latitude	Longitude	Altitude (m)	Soil type
Pyrénées (Py)	42°48′ N	00°32′ W	1600	M + S
Auvergne (Au)	45°30′ N	02°55′ E	1270	M
Alsace (Al)	48°05′ N	07°08′ E	950	M
Tailles 1 (Ta-1)	50°14′ N	05°42′ E	630	M
Tailles 2 (Ta-2)	50°15′ N	05°47′ E	565	P
Tailles 3 (Ta-3)	50°14′ N	05°47′ E	590	M + P

M, mineral; P, peaty; S, skeletal.

Table 2 List of the chloroplast primers and PCR parameters used in this study

Primer 1	Primer 2	Abbreviation	Reference	Annealing temperature (°C)	Elongation time (min)	Restriction endonucleases
trnH	trnK1	HK	1	62.0	2	*Hin*fI
trnK (exon 1)	trnK (exon 2)	K1K2	1	60.0	3	*Hin*fI
trnQ	trnR	QR	2	58.0	4	*Hin*fI
trnC	trnD	CD	1	57.5	4	*Hin*fI/*Dpn*II
trnD	trnT	DT	1	48.0	2	*Hin*fI
psbC	trnS	CS	1	58.0	2	*Hin*fI
psaA	trnS	AS	1	62.0	4	*Hin*fI
trnV	rbcL	VL	2	58.5	2.5	*Alu*I

References: 1, Demesure *et al*. (1995); 2, Dumolin-Lapègue *et al*. (1997).

DNA amplification and digestion

All individuals were characterized by restriction analysis of eight cpDNA fragments amplified by PCR with pairs of universal primers described in Demesure *et al*. (1995) and Dumolin-Lapègue *et al*. (1997) (Table 2). Diluted template DNA (final dilution 1:100) was amplified in 25 μL of reaction mixture containing 3.5 mM $MgCl_2$, 200 μg mL^{-1} BSA, 200 μM of each of the four dNTPs, 0.2 μM of each primer, 0.625 U of AmpliTaq Standard (Perkin Elmer) and 2.5 μL of the AmpliTaq buffer. Amplifications were also performed in 15 μL reaction volume without a significant decrease in yield, except for the CD primers. PCR reactions were carried out in Perkin Elmer thermocyclers (model 2400 or 9600), using one cycle of 5 min at 95 °C, 30 cycles of 45 s at 92 °C, 45 s at 48–62 °C (depending on the primers, see Table 2) and 2–4 min at 72 °C (depending on the fragment to be amplified, see Table 2), followed by one cycle of 10 min at 72 °C. PCR products (5 μL) were digested for 2–4 h with 1–2 units of restriction enzyme (reaction volume of 10 μL). As a first step, 20 individuals (five from each region) were screened for variation using two different 4 bp cutter restriction enzymes (*Alu*I and *Hin*fI) in order to ascertain whether the observed polymorphisms were due to insertions or deletions (indels), or to point mutations in restriction sites. Restriction fragments were separated by electrophoresis on both 1.6% agarose gels (fragments > 300 bp) and 8% polyacrylamide gels (prepared with a 19:1 acrylamide–bisacrylamide solution; fragment size between 80 and 300 bp) in 0.5 × TBE buffer. Agarose gels were stained with ethidium bromide and photographed with the BioImage system (Bioprobe) under UV light, whereas polyacrylamide gels were silver-stained.

For the fragment CD, polymorphic restriction fragments were observed in both agarose and polyacrylamide gels. In order to speed the scoring of variation for these fragments, the PCR products were co-digested with two restriction endonucleases (*Hin*fI and *Dpn*II), which allowed us to separate all polymorphic fragments on 6% polyacrylamide gels alone. For routine analysis of the remaining 130 individuals, the amplified fragments AS, CS, DT, HK, K1K2 and QR were digested with *Hin*fI, and the VL fragment was digested with *Alu*I. Restriction products of fragments AS, CS, DT, HK, K1K2 were electrophoresed on 1.6% agarose gels, and restriction products of fragment QR were separated on 6% polyacrylamide gels.

Data analysis

Variation among haplotypes. Parsimony analysis was used to derive the phylogenetic relationships among haplotypes. The molecular data were scored as unordered multi-state characters and the equally most-parsimonious trees were inferred using the branch-and-bound algorithm of PAUP 3.1 (Swofford 1993). The resulting trees were used (i) to analyse the relationships among haplotypes, and (ii) to classify the information produced by the different primers, in order to identify fragments subject to high levels of homoplastic mutations.

Analysis of the genetic differentiation. The total diversity (h_T), the average intra-population diversity (h_S), and the level of gene differentiation among populations (G_{STc}) and their standard errors were estimated following Pons & Petit (1995). The differentiation among populations was further investigated using UPGMA analysis to infer a tree from pairwise Cavalli-Sforza & Edwards (1967) genetic distances between populations.

Ennos (1994) showed that the levels of differentiation among populations estimated from nuclear markers (G_{STn}) and chloroplast markers (G_{STc}) could be used to derive the pollen to seed migration ratio, using the following formula:

$$r = \frac{m_p}{m_s} = \frac{\left(\frac{1}{G_{STn}} - 1\right) - 2\left(\frac{1}{G_{STc}} - 1\right)}{\left(\frac{1}{G_{STc}} - 1\right)} \qquad \text{(Ennos 1994)}$$

Table 3 Size of the polymorphic restriction fragments (in base pairs)

	Size (in bp)				
Polymorphic fragment	1	2	3	4	5
DT1	383	365			
CD2	311	308			
CD3	318	308			
CD4	297	292			
CD6	260	257	250		
CD7	240	229			
AS2	511	488			
QR4	247	202			
HK3	256	251			
VL3	675	574	564	554	544

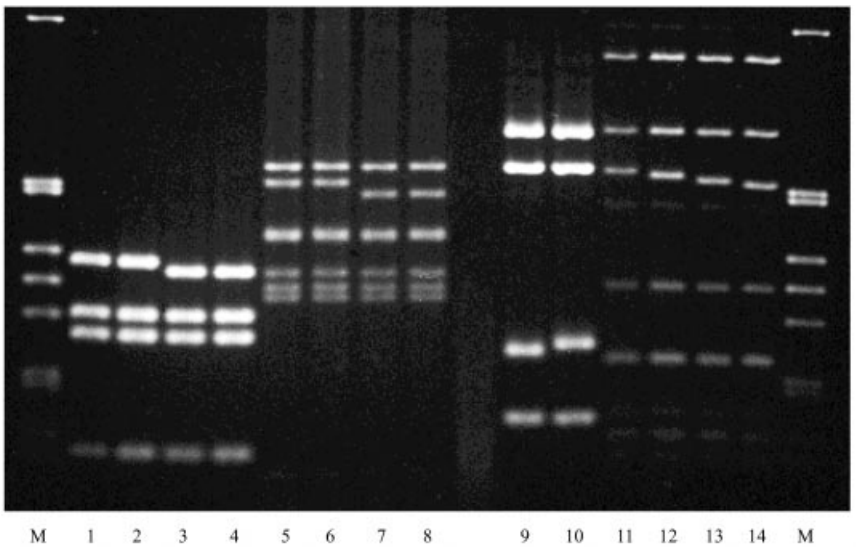

Fig. 2 cpDNA polymorphisms observed in *Sorbus aucuparia*. Lanes 1–4, CD fragment digested with *Hin*fI; lanes 5–8, AS fragment digested with *Hin*fI; lanes 9 & 10, HK fragment digested with *Hin*fI; lanes 11–14, VL fragment digested with *Alu*I (four alleles); M, molecular weight marker X (Boehringer).

where m_p and m_s are pollen and seed migration rates, respectively. This equation relies on a set of assumptions. In particular, the island model of population structure and equilibrium between drift and migration, as well as a strictly maternal mode of inheritance of the chloroplast genome are assumed. Although the inheritance of cpDNA has not yet been investigated in *S. aucuparia*, we could reasonably assume it to be maternal, as has been demonstrated in the majority (but not all) of the angiosperms studied so far (Reboud & Zeyl 1994; Birky 1995). We therefore used the above formula to compare the levels of genetic differentiation between nuclear and chloroplastic markers, with G_{STn} estimated on the basis of allozyme data previously obtained for the populations studied here (Raspé & Jacquemart 1998).

Results

A total of 51 restriction fragments obtained with the eight cpDNA PCR fragments were screened for variation in 150 individuals. Monomorphic patterns were observed with two pairs of primers (K1K2 and CS). The six remaining cpDNA PCR fragments generated 10 polymorphic restriction fragments (Table 3; Fig. 2), which together displayed 14 polymorphisms and allowed the separation of 12 haplotypes (Table 4). In addition to these 10 fragments, two fragments from the CD restriction profile were highly polymorphic. The use of these fragments in separating haplotypes resulted in a disproportionate increase of haplotype number, along with a high level of homoplasy. We could show that these two fragments were actually part of a PCR by-product of approximately 345 bp which was consistently amplified in all individuals. Because the origin of this secondary PCR fragment is not known and its abnormal polymorphism, it will not be considered any further in this study. All the detected mutations were apparently indels, but this could be confirmed only for the polymorphisms at fragments DT1, CD3 and VL3, which were observed with two different restriction

Table 4 Description of the 12 haplotypes found in *Sorbus aucuparia*. For all haplotypes, named A–L, the combination of alleles for all polymorphic fragments is given

	Polymorphic fragments									
Haplotypes	DT1	CD2	CD3	CD4	CD6	CD7	AS2	QR4	HK3	VL3
A	1	2	1	2	3	2	2	1	1	5
B	1	2	2	2	3	2	2	1	1	5
C	1	2	2	2	3	2	1	1	1	5
D	2	1	2	2	3	1	2	1	1	4
E	2	2	2	1	3	2	2	1	1	2
F	2	2	2	2	3	2	2	1	1	4
G	2	2	2	1	2	2	2	1	1	4
H	2	2	2	1	3	2	2	2	1	3
I	1	2	2	2	1	2	2	1	2	5
J	2	2	2	1	3	2	2	1	1	1
K	2	2	2	1	3	2	2	1	1	3
L	2	2	2	1	3	2	2	1	1	5

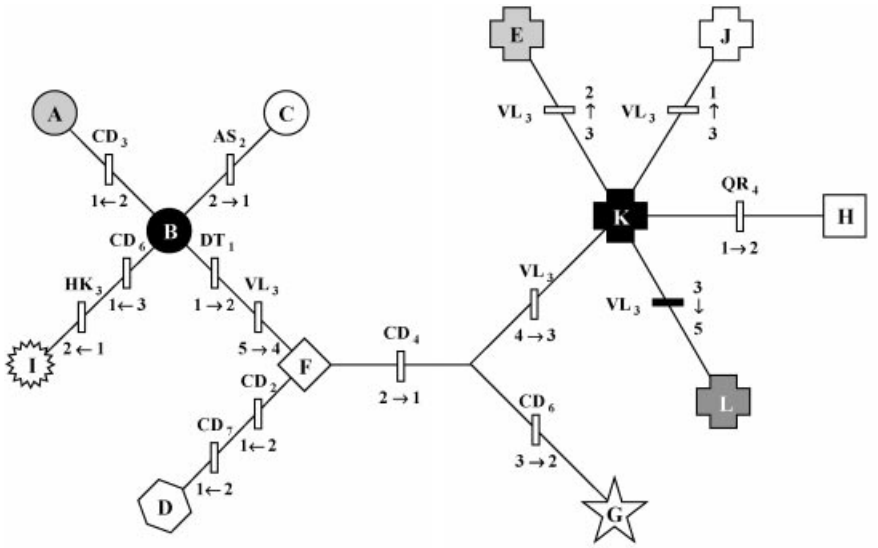

Fig. 3 One of the most parsimonious phylogenetic trees obtained with the 12 cpDNA haplotypes found in *Sorbus aucuparia*, with the position and direction of all character changes. The black rectangle indicates the homoplastic mutation.

endonucleases (*Hin*fI and *Alu*I). Although the remaining polymorphisms were observed only with *Hin*fI or after co-digestion by *Hin*fI and *Dpn*II (fragment CD), they might still result from indels. Indeed, some mutations were detected in individuals which were not tested with *Alu*I. Furthermore, the variation might have gone unseen after restriction with *Alu*I if the restriction fragment containing the mutation was either too small to be detected on the gels, or too long, compared to the size of the indel, to match the resolving power of the gels.

Genetic relationship between haplotypes

Phylogenetic analysis on the 12 haplotypes yielded 16 065 most-parsimonious trees of length 15. Since 14 polymorphisms were used in the analysis, this minimum length indicates that a homoplastic mutation was inferred. The homoplastic mutation apparently occurred either in fragment CD4 or in fragment VL3. Because of the high level of polymorphism of VL3, we consider it more likely that the homoplastic mutation occurred in this fragment. This homoplasy, along with the high level of polymorphism of the fragment VL3 (four polymorphisms), could account for the large number of most-parsimonious trees obtained. We therefore conducted a second analysis but did so after excluding VL3 from the data. A single most-parsimonious tree of length 10 was then obtained (i.e. with no homoplasy). Next, we repeated the analysis taking all characters into account, but enforcing the tree obtained without VL3 as a constraint. In this way, a total of 224 most-parsimonious trees were constructed (out of the 16 065 previously inferred). These trees notably differ in the nature (reversal versus 'parallel' mutation) and place of the homoplastic mutation. One of these trees is represented in Fig. 3.

Table 5 Haplotype composition of the six *Sorbus aucuparia* populations

	Populations						
Haplotypes	Py	Au	Al	Ta-1	Ta-2	Ta-3	Total
A		6			1		7
B	1	1		15	2	15	34
C		2	17	2	19	1	41
D	2			6	2	14	24
E	8	3					11
F	2						2
G	1						1
H			1				1
I				2			2
J			1				1
K	6	8		5	6		25
L			1				1
Total	20	20	20	30	30	30	150

Py, Pyrénées; Au, Auvergne; Al, Alsace; Ta, Plateau des Tailles.

Table 6 Components of haplotype diversity (standard errors are given in brackets)

	Inter-regions	Intra-region
h_T	0.892 (0.053)	0.816 (0.047)
h_S	0.637 (0.118)	0.605 (0.046)
G_{STc}	0.286 (0.146)	0.259 (0.117)
G_{STn} (allozymes)	0.025	0.045
r	—	5.5

Chloroplast DNA diversity and population differentiation

The genetic composition of all populations is given in Table 5. All populations contained at least three different haplotypes, with a mean of 4.7 ± 1.03 (SD). Four out of 12 haplotypes (B, C, D and K) represented 83% of the total sample, whereas six other haplotypes (F, G, H, I, J and L) were found only once or twice and were restricted to a single population. On both geographical scales considered, total chloroplast diversity was high and was mainly distributed within populations (Table 6). It is interesting to note that there is a high differentiation of the haplotypes, even at a local scale among the neighbouring populations of the Plateau des Tailles. Haplotype B is the most frequent in populations Ta-1 and Ta-3 (along with haplotype D in the latter) but is found at low frequency elsewhere, whereas haplotype C is found in more than half of the individuals in Ta-2. In addition, all populations except Ta-2 seem to cluster according to their geographical location in the UPGMA tree (Fig. 4).

The ratio between pollen and seed migration rates (r), estimated at the intra-regional scale, was 5.5.

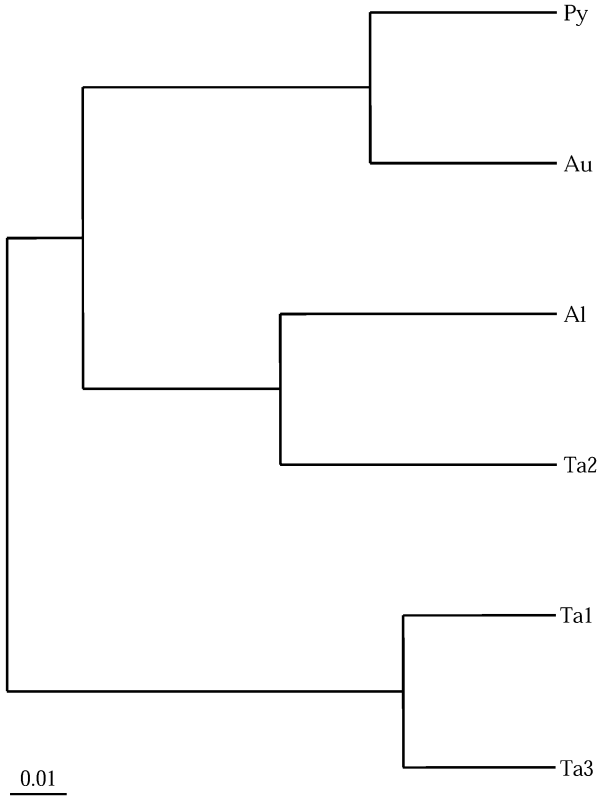

Fig. 4 UPGMA tree of *Sorbus aucuparia* populations, based on Cavalli-Sforza & Edwards (1967) genetic distances between populations. Py, Pyrénées; Au, Auvergne; Al, Alsace; Ta, Plateau des Tailles (Belgium).

Discussion

cpDNA polymorphism in Sorbus aucuparia

The majority, if not all, polymorphisms detected in this study are likely to result from short indels of 3–45 bp

(except one indel longer than 100 bp in fragment VL3). This is in accordance with the review by Clegg *et al*. (1994) which reported a higher frequency of indels relative to point mutations in the chloroplast genome. It should be noted, however, that the experimental approach used in this study (i.e. the screening of non-coding regions with one or two 4 bp cutter restriction endonucleases) is more likely to detect indels than point mutations, because the latter can be detected only when they affect a restriction site. The amount of variation detected in this study, despite the limited number of populations investigated, confirms the efficiency of the PCR–RFLP methodology in detecting polymorphism in the chloroplast genome (e.g. Demesure *et al*. 1996; El Mousadik & Petit 1996a; King & Ferris 1998).

As in other species, the number of mutations detected differed among PCR fragments. The CD and VL regions, in particular, yielded a very high number of mutations compared to other regions. Size differences among fragments are not a sufficient explanation, since large fragments such as AS and QR both showed only one mutation. High mutation rates in particular regions of the chloroplast genome have been explained in various ways, including (1) the presence of simple sequence (mostly mononucleotide) repeats (e.g. Powell *et al*. 1995; Weising & Gardner 1999); (2) direct or inverted repeat mediated deletions or insertions in A+T-rich regions (Ogihara *et al*. 1992); and (3) recurrent duplications of multi-nucleotide direct repeats (Hipkins *et al*. 1995). For VL3, the apparently constant size of length mutations (10 bp) and their concentration in a single restriction fragment are consistent with the last process. Precise determination of the nature of the mutations involved, however, was beyond the scope of this paper. Nevertheless, it is noteworthy that, because of the repetitive nature of the mutations involved in the evolution of simple or multi-nucleotide repeats, homoplasy is likely to occur. This seemed to be the case for the VL3 fragment in *Sorbus aucuparia*. A similar observation has been made for the VL region in a study on linkage disequilibrium between cpDNA and mtDNA haplotypes in *Beta vulgaris* subsp. *maritima* L. populations (Desplanque *et al*. 2000), suggesting a high level of mutation in this particular region. Consequently, while such regions may be very useful for studies of population structure or gene flow at a relatively restricted geographical scale (because of the high number of variants), their usefulness in phylogeny or phylogeography is lowered by the higher probability of homoplasy.

Population differentiation and gene flow

It has been shown theoretically that, for many models of population structure, the level of genetic differentiation among populations is expected to be higher for maternally inherited cpDNA markers than for biparentally inherited nuclear genes (Birky *et al*. 1989; Petit *et al*. 1993a; Ennos 1994; Hu & Ennos 1997). Several factors contribute to this increase of genetic differentiation for the chloroplast genome. (1) Effective gene flow is limited to seed dispersal for maternally inherited DNA. (2) Drift is expected to be twice as strong for a haploid genome compared to a diploid one. (3) It has been shown in some hermaphrodite tree species (such as oaks) that the flowering, seed dispersal and recruitment patterns resulted in an effective number of trees contributing to the next generation as females that is much reduced compared to the effective number of trees acting as male parents (Dow & Ashley 1996). The latter factor, however, is not as 'absolute' as the first two. Indeed, differences in species' pollination ecology or seed dispersal mechanisms might have important consequences on the balance between female and male reproductive success of individual trees.

As expected, the level of differentiation among populations of *S. aucuparia* was higher for cpDNA markers (G_{STc} = 0.286) than for nuclear (isozyme) markers (G_{STn} = 0.025). These values of G_{ST}, as well as r, may be compared to estimates obtained in other tree species (Table 7). Although the assumptions of the island model of population structure on which the estimation of r is based may be questioned, the values in Table 7 are given in order to compare our results to those obtained for other species in a comparable geographical range. Whereas values of G_{STn} and r obtained in *S. aucuparia* were individually similar to some species, G_{STc} was less than half the value usually observed in other tree species. *S. aucuparia* did not share a similar combination of values of G_{STn}, G_{STc} and r with any of the tree species given in Table 7. Given the low level of cpDNA differentiation observed in *S. aucuparia*, the pollen to seed migration ratio is much lower than in species such as *Fagus sylvatica* and *Quercus* spp. in particular. As pointed out by Ennos (1994), pollen to seed migration ratios estimated by r should be considered not as absolute estimates, but as 'realistic' figures which are most useful in a comparative context. This is because the island model of population structure on which the estimation of r is based makes a large number of assumptions that are not upheld in nature (Ennos 1994). In particular, caution needs to be exercised in using data from species with long generation times, especially where disturbance of populations is known to have occurred in the recent past (Ennos 1994), which is precisely the case in this study (see below). We will therefore keep to the most striking differences among species, without drawing any inference on the actual relative contributions of pollen and seed migration rates to the effective gene flow in *S. aucuparia*.

The discrepancies mentioned above might be accounted for by the contrasting life history traits of *S. aucuparia*. First, this species produces fleshy fruits, which results in an efficient bird-mediated seed dispersal (Snow & Snow

Table 7 Comparison of genetic differentiation for nuclear (G_{STn}) and maternally inherited (G_{STc}) markers, and pollen to seed migration ratio (r), in various tree species

Species	G_{STn}	Reference	G_{STc}	Reference	r
Sorbus aucuparia	0.025	Raspé & Jacquemart 1998	0.29	This study	13.7†
Argania spinosa	0.25	El Mousadik & Petit (1996b)	0.60	El Mousadik & Petit (1996a)	2.5
Eucalyptus nitens	0.30	Moran (1992)	0.62	Byrne & Moran (1994)	1.8
Pinus contorta	0.09	Wheeler & Guries (1982)	0.72	Dong & Wagner (1993)	24
Pinus flexii	0.016§	Latta & Mitton (1997)	0.68‡	Latta & Mitton (1997)	128
Pseudotsuga menziesii	0.24	Li & Adams (1988)	0.73‡	Aagaard *et al*. (1995)	7
Alnus glutinosa	0.20	Prat *et al*. (1992)	0.87	King & Ferris 1998	23
Fagus sylvatica	0.054	Comps *et al*. (1990)	0.83	Demesure *et al*. 1996	84
Quercus petraea	0.024	Zanetto *et al*. (1994)	0.90	Petit *et al*. (1993b)	500
Quercus robur	0.032	Zanetto *et al*. (1994)	0.92	Petit *et al*. (1993b)	286

†Estimated on the inter-regional scale (see text for discussion). ‡Based on maternally inherited mitochondrial markers. §Median value over seven allozyme loci.

1988) and is likely to generate high levels of gene flow through seed movement. Second, the pioneer or post-pioneer habit of *S. aucuparia*, which contrasts with late-successional social species, might also be of importance. Frequent extinction and re-colonization events characteristic of pioneer species might have disrupted the strong spatial structure of cpDNA variation likely to have been caused by long-distance dispersal events during re-colonization after the last glaciation (Ibrahim *et al*. 1996; Le Corre *et al*. 1997), whereas such a structure is thought to have been maintained in late-successional, social tree species such as oaks (Le Corre *et al*. 1997; Petit *et al*. 1997). However, further comparative work on species with contrasting life history traits is needed to estimate the relative importance of these traits as a cause of the low observed differentiation.

Within species, frequent extinctions and re-colonizations might increase genetic differentiation among newly founded populations compared to old populations, particularly in the case of a propagule pool model of colonization (e.g. Whitlock & McCauley 1990). Such a phenomenon has been observed in *S. aucuparia* with isozyme markers (Raspé & Jacquemart 1998; O. Raspé, unpublished results) and could also apply to cpDNA markers. Indeed, the level of genetic differentiation among the Plateau des Tailles populations was nearly as high as that observed among regions. These populations were recently founded and have been shown to exhibit an unexpectedly high level of nuclear differentiation, given the small distance separating them. The sampling effect related to the founding of these populations may account for the prevalence of different haplotypes in neighbouring populations (haplotype B in Ta-1, haplotype C in Ta-2, and haplotypes B and D in Ta-3). This might also explain why population Ta-2 clusters with the population from Alsace in the UPGMA tree, whereas the other populations cluster according to their geographical location. In the previous isozyme study (Raspé & Jacquemart 1998), all populations also clustered according to their geographical location except two populations from the Pyrénées and the Plateau des Tailles. Therefore, genetic variation at both nuclear and chloroplast levels seems, to some extent, geographically structured.

Conclusion

Despite the limited number of populations considered, and the relatively restricted geographical range investigated, the *Sorbus aucuparia* chloroplast genome exhibited many polymorphisms, allowing the identification of 12 haplotypes. These polymorphisms were not randomly distributed among the cpDNA regions analysed. Two regions, CD and VL, were highly variable. One restriction fragment in particular (VL3) showed a high level of variation along with homoplasy, illustrating that PCR–RFLP cpDNA markers may differ in their evolutionary properties (i.e. mutation rate and occurrence of homoplasy), and that while such regions may be very useful for studies of population structure or gene flow on a micro-geographic scale, their usefulness in phylogeography is lowered by a high level of homoplasy.

The most interesting result of this study is the low genetic differentiation among *Sorbus aucuparia* populations compared to the other tree species studied so far. This discrepancy might be accounted for by the contrasting life history traits of *S. aucuparia*, including bird dispersal of seeds and pioneer habit. We also suspect the meta-population dynamics of the Plateau des Tailles populations to have led to an increase of genetic differentiation among them, as was observed with isozymes. A more detailed study of cpDNA variation in the Plateau des Tailles populations is currently being conducted to investigate in detail the meta-population processes occurring in this insect-pollinated and bird-dispersed pioneer species.

Acknowledgements

We would like to thank P. Coquillard, L. Ghelen, J. Gueugnot and G. Mahy for their help in sampling *S. aucuparia* populations. We are also grateful to R.J. Petit for providing his program HAPLODIV, which was used to compute gene diversity parameters, and to J.R. Kohn and B. Igic for their comments on the last draft of the manuscript. O. Raspé and Anne-Laure Jacquemart are presently 'Chargé de Recherches' and 'Chercheur Qualifié' of the Belgian National Fund for Scientific Research (FNRS).

References

Aagaard JE, Vollmer SS, Sorensen FC, Strauss SH (1995) Mitochondrial DNA products among RAPD profiles are frequent and strongly differentiated between races of Douglas fir. *Molecular Ecology*, **4**, 441–447.

Allan GJ, Clark C, Rieseberg LH (1997) Distribution of parental DNA markers in *Encelia virginensis* (Asteraceae, Heliantheae), a diploid species of putative hybrid origin. *Plant Systematics and Evolution*, **205**, 205–221.

Birky CW Jr (1995) Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proceedings of the National Academy of Sciences of the USA*, **92**, 11331–11338.

Birky CW Jr, Fuerst P, Maruyama T (1989) Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics*, **121**, 613–627.

Byrne M, Moran GF (1994) Population divergence in the chloroplast genome of *Eucalyptus nitens*. *Heredity*, **73**, 18–28.

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis models and estimation procedures. *Evolution*, **21**, 550–570.

Clapham AR, Tutin TG, Warburg EF (1962) *Flora of the British Isles*, 2nd edn. Cambridge University Press, London.

Clegg MT, Gaut BS, Learn GH, Morton BR (1994) Rates and patterns of chloroplast DNA evolution. *Proceedings of the National Academy of Sciences of the USA*, **91**, 6795–6801.

Comps B, Thiébaut B, Paule L, Merzeau D, Letouzey J (1990) Allozymic variability in beechwoods (*Fagus sylvatica* L.) over central Europe: spatial differentiation among and within populations. *Heredity*, **65**, 407–417.

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. *Molecular Ecology*, **4**, 129–131.

Demesure B, Comps B, Petit RJ (1996) Chloroplast DNA phylogeography of the common beech (*Fagus sylvatica* L.) in Europe. *Evolution*, **50**, 2515–2520.

Desplanque B, Viard F, Bernard J *et al.* (2000) The linkage disequilibrium between chloroplast DNA and mitochondrial DNA haplotypes in *Beta vulgaris* ssp. *maritima* (L.): the usefulness of both genomes for population genetics studies. *Molecular Ecology*, **9**, 141–154.

Dong J, Wagner DB (1993) Taxonomic and population differentiation of mitochondrial diversity in *Pinus banksiana* and *Pinus contorta*. *Theoretical and Applied Genetics*, **86**, 573–578.

Dow BD, Ashley MV (1996) Microsatellite analysis of seed dispersal and parentage of saplings in bur oaks, *Quercus macrocarpa*. *Molecular Ecology*, **5**, 615–627.

Dumolin-Lapègue S, Pemonge MH, Petit RJ (1997) An enlarged set of consensus primers for the study of organelle DNA in plants. *Molecular Ecology*, **6**, 393–397.

Edwards K, Johnstone C, Thompson C (1991) A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Research*, **19**, 1349.

El Mousadik M, Petit RJ (1996a) Chloroplast DNA phylogeography of the argan tree of Morocco. *Molecular Ecology*, **5**, 547–555.

El Mousadik M, Petit RJ (1996b) High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. *Theoretical and Applied Genetics*, **92**, 832–839.

Ennos R (1994) Estimating the relative rates of pollen and seed migration among plant populations. *Heredity*, **72**, 250–259.

Gielly L, Taberlet P (1994) The use of chloroplast DNA to resolve plant phylogenies—noncoding versus rbcL sequences. *Molecular Biology and Evolution*, **11**, 769–777.

Hamrick JL, Godt MJW, Sherman-Broyles SL (1992) Factors influencing genetic diversity in woody plant species. *New Forests*, **6**, 95–124.

Hipkins VD, Marshall KA, Neale DB, Rottman WH, Strauss SH (1995) A mutation hotspot in the chloroplast genome of a conifer (Douglas-fir: Pseudotsuga) is caused by variability in the number of direct repeats derived from a partially duplicated tRNA gene. *Current Genetics*, **27**, 572–579.

Hu XS, Ennos R (1997) On estimation of the ratio of pollen to seed flow among plant populations. *Heredity*, **79**, 541–552.

Ibrahim KM, Nichols RA, Hewitt GM (1996) Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity*, **77**, 282–291.

King RA, Ferris C (1998) Chloroplast DNA phylogeography of *Alnus glutinosa* (L.) Gaertn. *Molecular Ecology*, **7**, 1151–1161.

Kullman L (1986) Temporal and spatial aspects of subalpine populations of *Sorbus aucuparia* in Sweden. *Annales Botanici Fennici*, **23**, 267–275.

Latta RG, Mitton JB (1997) A comparison of population differentiation across four classes of gene marker in limber pine (*Pinus flexilis* James). *Genetics*, **146**, 1153–1163.

Le Corre V, Machon N, Petit RJ, Kremer A (1997) Colonization with long-distance seed dispersal and genetic structure of maternally inherited genes in forest trees: a simulation study. *Genetical Research*, **69**, 117–125.

Li P, Adams WT (1988) Range-wide patterns of allozyme variation in Douglas fir (*Pseudotsuga menziesii*). *Canadian Journal of Forest Research*, **19**, 149–161.

McCauley DE (1995) The use of chloroplast DNA polymorphism in studies of gene flow in plants. *Trends in Ecology and Evolution*, **10**, 198–202.

Moran GF (1992) Patterns of genetic diversity in Australian tree species. *New Forests*, **6**, 49–66.

Ogihara Y, Terachi T, Sasakuma T (1992) Structural analysis of length mutations in a hot-spot region of wheat chloroplast DNAs. *Current Genetics*, **22**, 251–258.

Palmer JD, Jansen RK, Michaels HJ, Chase MW, Manhart JR (1988) Chloroplast DNA and plant phylogeny. *Annals of the Missouri Botanical Garden*, **75**, 1180–1206.

Petit RJ, Kremer A, Wagner DB (1993a) Finite island model for organelle and nuclear genes in plants. *Heredity*, **71**, 630–641.

Petit RJ, Kremer A, Wagner DB (1993b) Geographic structure of chloroplast DNA polymorphisms in European oaks. *Theoretical and Applied Genetics*, **87**, 122–128.

Petit RJ, Pineau E, Demesure B *et al.* (1997) Chloroplast DNA footprints of postglacial recolonization by oaks. *Proceedings of the National Academy of Sciences of the USA*, **94**, 9996–10001.

Pons O, Petit RJ (1995) Estimation, variance and optimal sampling of gene diversity I. Haploid locus. *Theoretical and Applied Genetics*, **90**, 462–470.

Powell W, Morgante M, McDevitt R, Vendramin GG, Rafalski JA (1995) Polymorphic simple sequence repeat regions in chloroplast genomes—applications to the population genetics of pines. *Proceedings of the National Academy of Sciences of the USA*, **92**, 7759–7763.

Prat D, Leger C, Bojovic S (1992) Genetic diversity among *Alnus glutinosa* (L.) Gaertn. populations. *Acta Oecologica*, **13**, 469–477.

Rameau JC, Mansion D, Dumé G (1989) *Flore Forestière Française*. Institut pour le Développement Forestier, Paris.

Raspé O (1998) *Biologie de la reproduction et variation génétique d'un arbre entomophile*: Sorbus aucuparia *L. (Rosaceae: Maloideae)*. PhD Thesis, Catholic University of Louvain, Louvain-la-Neuve, Belgium.

Raspé O, Jacquemart AL (1998) Allozyme diversity and genetic structure of European populations of *Sorbus aucuparia* L. (Rosaceae: Maloideae). *Heredity*, **81**, 537–545.

Reboud X, Zeyl C (1994) Organelle inheritance in plants. *Heredity*, **72**, 132–140.

Snow B, Snow D (1988) *Birds and Berries. A Study of an Ecological Interaction*. T & AD Poyser, London.

Soltis DE, Gitzendanner MA, Strenge DD, Soltis PS (1997) Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. *Plant Systematics and Evolution*, **206**, 353–373.

Swofford DL (1993) PAUP: *Phylogenetic Analysis Using Parsimony*, Version 3.1. Illinois Natural History Survey, Champaign, Illinois.

Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453–464.

Tarayre M, Saumitou-Laprade P, Cuguen J, Couvet D, Thompson JD (1997) The spatial genetic structure of cytoplasmic (cpDNA) and nuclear (allozyme) markers within and among populations of the gynodioecious *Thymus vulgaris* (Labiatae) in southern France. *American Journal of Botany*, **84**, 1675–1684.

Wagstaff SJ, Hickerson L, Spangler R, Reeves PA, Olmstead RG (1998) Phylogeny in Labiatae s.l., inferred from cpDNA sequences. *Plant Systematics and Evolution*, **209**, 265–274.

Weising K, Gardner RC (1999) A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. *Genome*, **42**, 9–19.

Wheeler NC, Gurries RP (1982) Population structure, genetic diversity, and morphological variation in *Pinus contorta* Dougl. *Canadian Journal of Forest Research*, **12**, 595–606.

Whitlock MC, McCauley DE (1990) Some population genetic consequences of colony formation and extinction: genetic correlations within founding groups. *Evolution*, **44**, 1717–1724.

Wolfe KH, Li WH, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast and nuclear DNAs. *Proceedings of the National Academy of Sciences of the USA*, **84**, 9054–9058.

Zanetto A, Roussel G, Kremer A (1994) Geographic variation of inter-specific differentiation between *Quercus robur* L. & *Quercus patraea* (Matt.) Liebl. I. Monolocus patterns of variation. *Forest Genetics*, **1**, 111–123.

This work is part of Olivier Raspé's thesis, which dealt with the reproductive biology and population genetics of *Sorbus aucuparia*. He is working as part of the plant ecology and population biology team of the Ecology and Biogeography Laboratory in Louvain-la-Neuve (Belgium). The team, headed by Anne-Laure Jacquemart, is mainly interested in the population biology and mating system evolution of Ericaceae, as well as in the ecology and conservation of peat bogs and heathlands. The molecular analyses presented in this paper were realized in the Laboratory of Genetics and Evolution of Plant Populations (CNRS-University of Lille 1, France), where Pierre Saumitou-Laprade co-ordinates the development of molecular markers in many different plant species. Joël Cuguen is mainly interested in population genetics of spatially structured populations, investigating nuclear and cytoplasmic components of gene flow using molecular markers.
