

Genetic relationships between diploid and allotetraploid cherry species (*Prunus avium*, *Prunus* × *gondouinii* and *Prunus cerasus*)

M Tavaud^{1,3}, A Zanetto¹, JL David², F Laigret¹ and E Dirlewanger¹

¹INRA, UREFV, BP81, 71, Avenue E Bourleaux, 33883 Villenave d'Ornon, France; ²INRA, UMR DGPC, Domaine de Melgueil, 34130 Mauguio, France

Prunus avium L. (diploid, AA, $2n=2x=16$), *Prunus cerasus* L. (allotetraploid, AAFF, $2n=4x=32$) species, and their hybrid *Prunus* × *gondouinii* Rehd., constitute the most widely cultivated cherry tree species. *P. cerasus* is supposed to be an hybrid species produced by the union of unreduced *P. avium* gametes and normal *P. fruticosa* gametes. A continuum of morphological traits between these three species makes their assignation difficult. The aim of this paper is to study the genetic relationships between tetraploid and diploid cherry species. In all, 114 genotypes belonging to these species were analyzed using 75 AFLP markers. The coordinates of these genotypes on the first axis of a

correspondence analysis allowed us to clearly distinguish each species, to identify misclassifications and to assign unknown genotypes to one species. We showed that there are specific alleles in *P. cerasus*, which are not present in the A genome of *P. avium* and which probably come from the F genome of *P. cerasus*. The frequencies of each marker in the A and the F genomes were estimated in order to identify A and F specific markers. We discuss the utility of these specific markers for finding the origin of the A and F genomes in the allopolyploid species.

Heredity (2004) 93, 631–638. doi:10.1038/sj.hdy.6800589
Published online 8 September 2004

Keywords: *Prunus avium*; *Prunus cerasus*; AFLP markers; species identification; genetic diversity

Introduction

The most widely cultivated cherry trees belong to *Prunus avium* L. and *P. cerasus* L. species. Along with *P. fruticosa* Pall., these species and their interspecific hybrids constitute the *Eucerasus* section of the *Cerasus* subgenus, based on morphological criteria (Rehder, 1947; Krussmann, 1978). This classification and the monophyletic origin of the *Eucerasus* clade have been confirmed by chloroplastDNA variation analysis (Badenes and Parfitt, 1995).

P. avium includes sweet cherries, cultivated for human consumption and wild cherry trees grown for their wood (Webster, 1996). This species is diploid (AA, $2n=2x=16$) and its natural range covers the temperate regions of Europe, from the Northern part of Spain to the Southeastern part of Russia (Hedrick, 1915). *P. fruticosa*, the ground cherry tree, is a tetraploid wild species ($2n=4x=32$) thought to be autotetraploid (FFFF), sometimes used as rootstocks for other *Prunus* species. This species is widespread over the major part of central Europe, Siberia and Northern Asia (Hedrick, 1915). *P. cerasus*, the sour cherry tree, is cultivated for fruit, used in jam or liquor. It is an allotetraploid species (AAFF, $2n=4x=32$), thought to result from natural hybridiza-

tion between *P. avium* (producing unreduced gametes) and *P. fruticosa* (Figure 1). This origin was first suggested by Olden and Nybom (1968), who observed that artificial hybrids between tetraploid *P. avium* and *P. fruticosa* were very similar to *P. cerasus*. Isozyme analyses, genomic *in situ* hybridization and karyotype analysis further confirmed the hybrid origin of *P. cerasus* (Hancock and Iezzoni, 1988; Santi and Lemoine, 1990a; Schuster and Schreiber, 2000). The patterns of inheritance of seven isozymes in different crosses of sour cherry indicated that *P. cerasus* may be a segmental allopolyploid (Beaver and Iezzoni, 1993). Recent studies based on cpDNA markers detected two distinct chlorotypes in *P. cerasus*, which strongly suggest that crosses between *P. avium* and *P. fruticosa* have occurred at least twice to produce sour cherry (Badenes and Parfitt, 1995; Iezzoni and Hancock, 1996; Brettin *et al.*, 2000). Moreover, these results suggest that, most of the time, *P. fruticosa* was the female progenitor of *P. cerasus*, but in few cases *P. avium* was the female parent, via the formation of unreduced ovules. Triploid hybrids through the fusion of normal gametes of *P. avium* and *P. fruticosa* occur naturally but remain sterile. Moreover, they are not clonally propagated by man since they exhibit many unfavorable *P. fruticosa* traits (Olden and Nybom, 1968).

The duke cherries, which result from crosses between *P. avium* and *P. cerasus*, are cultivated at a much smaller scale. Different names have been given to this species including *P. acida* Dum, *Cerasus regalis*, *P. avium* ssp *regalis*, but the name used currently is *P. × gondouinii* Rehd. (Faust and Suranyi, 1997; Saunier and Claverie, 2001). This species is allotetraploid (AAAF, $2n=4x=32$),

Correspondence: Dr M Tavaud, INRA, UREFV, BP81, 71, Avenue E Bourleaux, 33883 Villenave d'Ornon, France.

E-mail: tavaud@ensam.inra.fr

³Current address: INRA, UMR DGPC, Domaine de Melgueil, 34130 Mauguio, France.

Received 27 November 2003; accepted 16 July 2004; published online 8 September 2004

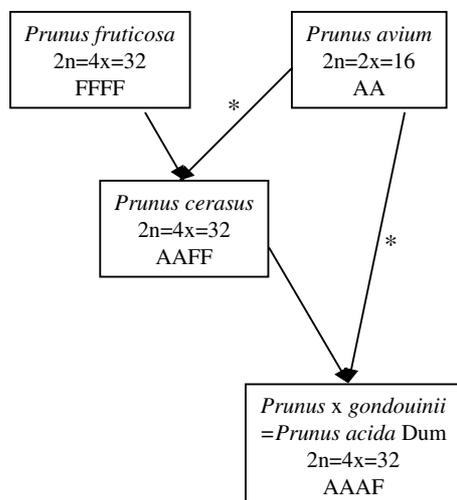


Figure 1 Hypothesis for the relationships between the four species belonging to the *Eucerasus* section. **P. avium* is thought to produce diploid gametes. A and F are haploid genomes coming from *P. avium* and *P. fruticosa* respectively.

stemming from the fertilization of sour cherry by unreduced gametes of sweet cherry (Iezzoni *et al*, 1990). These hybrids are often sterile, due to disturbances during meiosis, but they are clonally propagated by man. The duke cherry trees present tree and fruit characteristics intermediate between their progenitors.

Great morphological variation exists among *P. avium*, *P. fruticosa* and *P. cerasus* species. Multivariate analysis on sour cherry revealed continuous variation between the *P. avium*, *P. cerasus* and *P. fruticosa* traits throughout the geographic distribution of the species. In Western Europe, *P. cerasus* trees look like *P. avium*, whereas in Eastern Europe *P. cerasus* are closer to *P. fruticosa* (Hillig and Iezzoni, 1988; Krahl *et al*, 1991). This continuum of morphological characteristics makes species assignment difficult when considering only phenotypic traits.

Genetic resource conservation, essential for future breeding programs, requires a good characterization of germplasm genetic diversity and a proper assignment of individual genotypes to species. Molecular markers supply tools to analyze genetic diversity in *Prunus* species (Baird *et al*, 1996). They could be helpful by giving an accurate and unambiguous assignment of each genotype to a particular species. Isozymes have been already used to characterize genetic diversity in *P. avium* and *P. cerasus*, but the number of available isozymes is limited and their polymorphisms are low (Santi and Lemoine, 1990b). AFLP markers could be useful because they are reproducible, they need no previous information about the genome and they allow the detection of many markers in few analysis (Vos *et al*, 1995; Jones *et al*, 1997). They are well suited for studying unexplored species and have already been used successfully to discriminate different species with varying ploidy levels in *Gossypium* and *Oryza* (Aggarwal *et al*, 1999; Abdalla *et al*, 2001).

The aim of this work was first to determine whether AFLP markers could be useful in assigning each cherry genotype accurately to one species (*P. avium*, *P. cerasus* or *P. x gondouinii*). We also wanted to find specific markers of the different genomes (A and F) of these cherry species. These markers would be useful for comparing the diversity of the A genome between *P. avium* and *P.*

cerasus and to identify the origin of the A genome in both tetraploid species (*P. cerasus* and *P. x gondouinii*).

Materials and methods

Plant material

In all, 114 genotypes were studied: 68 *P. avium* genotypes (33 cultivated and 35 wild) sampled throughout the natural range of the species, 25 *P. cerasus* genotypes and 9 *P. x gondouinii* genotypes according to morphological characters (Table 1). In total, 12 French genotypes which had not been previously studied were not assigned to a species. The 33 cultivated sweet cherry trees were old cultivars, from 13 different countries. The 35 wild cherry trees were from several European forests. Sour and duke cherry trees and unassigned genotypes were from the French germplasm, mainly maintained in INRA orchards, at the 'Unité Expérimentale d'Arboriculture de Bordeaux'.

AFLP analyses

Genomic DNA was extracted from young leaves, following the method described by Viruel *et al* (1995). We essentially followed the AFLP protocol developed by Vos *et al* (1995) with minor modifications. DNA was digested with *EcoRI* and *MseI* restriction enzymes. Digestion was performed for 4 h at 37°C in a final volume of 17.5 µl containing 10 mM Tris-Ac, 10 mM MgAc, 50 mM KAc, 5 mM DTT, 25 U *EcoRI* (Pharmacia), 4 U *MseI* (New England Biolabs) and 250 ng of genomic DNA. Two linkers, one for the *EcoRI* sticky ends and the other for the *MseI* sticky ends, were ligated for 3 h at 37°C to the digested DNA by adding 2.5 µl of a mix containing 2.5 pmol *EcoRI* linker, 25 pmol *MseI* linker, 8 mM ATP, 10 mM Tris-HAc, 10 mM MgAc, 50 mM KAc, 5 mM DTT and 0.85 U T4 DNA ligase (Pharmacia). This ligation product was diluted five-fold. A first preselective PCR amplification was performed using *EcoRI* + A and *MseI* + C primers in a 50 µl mix of 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 1.5 mM MgCl₂, 0.2 mM of each dNTP, 30 ng of each primer, 1 U *Taq* DNA polymerase (GibcoBRL) and 5 µl of diluted ligation product. The reaction was carried out in a Perkin-Elmer 9600 thermocycler and the samples were subjected to 28 amplification cycles with three steps of 30 s at 94°C, 60 s at 56°C and 60 s at 72°C. The preamplification products were diluted 10-fold and used as starting material for the selective radioactive amplification.

For selective amplification, *EcoRI* and *MseI* primers with three and two additional nucleotides, respectively, were used (Table 2), the first one being ³³P-labelled using T₄ polynucleotide kinase. The PCR reaction was performed in a 20 µl volume of 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 1.5 mM MgCl₂, 5 ng [³³P] *EcoRI* primer, 30 ng *MseI* primer, 1 U *Taq* DNA polymerase (GibcoBRL) and 5 µl of diluted preamplified DNA. The selective amplification was carried out using the following cycling parameters: 11 cycles of 30 s at 94°C, 30 s at 65°C, 60 s at 72°C, in which the annealing temperature was lowered by 0.7°C per cycle, followed by 24 cycles of 30 s at 94°C, 30 s at 56°C and 60 s at 72°C.

The PCR products, in which an equal volume of load buffer (98% formamide, 10 mM EDTA, 0.05% bromophenol blue and 0.05% xylene cyanol) had been added, were denatured 5 min at 94°C and immediately placed on ice. In all, 2 µl of each sample was loaded on a

Table 1 (a) *A priori* species assignment, geographical origin and sample number of the cultivated and wild genotypes studied. (b) Names of cherry varieties analysed and their species assignment according to their coordinates on the first axis of the factorial analysis

(a) ^a						
<i>P. avium</i> priori assignments	Geographical origin	Sample number		<i>P. avium</i> posteriori assignments		
		Cultivated cherry	Wild cherry	<i>P. av</i>	<i>P. gon</i>	<i>P. cer</i>
<i>P. avium</i>	Czech republic	3		3		
	Canada	1		1		
	France	10	8	14	4	
	Germany	3	5	8		
	Georgia		4	4		
	Great Britain	2	2	4		
	Greece	2		2		
	Italy	3	5	5	3	
	Japan	1		1		
	Poland	1		1		
	Romania	2	3	5		
	Slovakia		1	1		
	Spain		5	3		2
	Sweden		2	2		
	Switzerland	2		2		
	USA	1		1		
	Yugoslavia	2		1		1
	Total	33	35	58	7	3
<i>P. cerasus</i>	France	25		2	2	21
<i>P. × gondouinii</i>	France	9		2	3	4
Undetermined genotypes	France	12		2	4	6
Total			114			

(b)			
Variety names	Origin	a priori assignment	a posteriori assignment
Bigarreau Pelissier	France	<i>P. avium</i>	<i>P. avium</i>
Guigne noire Luisante	France	<i>P. avium</i>	<i>P. avium</i>
Cœur de Bœuf	France	<i>P. avium</i>	<i>P. avium</i>
Koromilokeraso Vitalou	Greece	<i>P. avium</i>	<i>P. avium</i>
Petrokeraso Tragano Achaïas	Greece	<i>P. avium</i>	<i>P. avium</i>
Pivovka	Czech republic	<i>P. avium</i>	<i>P. avium</i>
Chlumecka cerna	Czech republic	<i>P. avium</i>	<i>P. avium</i>
Sakvicka	Czech republic	<i>P. avium</i>	<i>P. avium</i>
Bigarreau Cœur de Noyon	France	<i>P. avium</i>	<i>P. avium</i>
Knauffs Schwarze	Germany	<i>P. avium</i>	<i>P. avium</i>
Karina	Germany	<i>P. avium</i>	<i>P. avium</i>
Büttner's Rote Knorpelkirsche	Germany	<i>P. avium</i>	<i>P. avium</i>
Bedford Prolific <i>P. avium</i>	Great Britain	<i>P. avium</i>	<i>P. avium</i>
Nutberry Black	Great Britain	<i>P. avium</i>	<i>P. avium</i>
Camus	France	<i>P. avium</i>	<i>P. avium</i>
Vittoria	Italy	<i>P. avium</i>	<i>P. avium</i>
Maru Ruby [®]	USA	<i>P. avium</i>	<i>P. avium</i>
Van 2D 9–11	Canada	<i>P. avium</i>	<i>P. avium</i>
Amar 153	Yugoslavia	<i>P. avium</i>	<i>P. avium</i>
Bianca di Verona	Italy	<i>P. avium</i>	<i>P. avium</i>
Sato Nishiki	Japan	<i>P. avium</i>	<i>P. avium</i>
Arodel	France	<i>P. avium</i>	<i>P. avium</i>
Kunzegeo	Poland	<i>P. avium</i>	<i>P. avium</i>
Silva	Romania	<i>P. avium</i>	<i>P. avium</i>
Roz Amar de Mărculesti (RAM)	Romania	<i>P. avium</i>	<i>P. avium</i>
Ravenna	Italia	<i>P. avium</i>	<i>P. avium</i>
Frühe Luxburger	Switzerland	<i>P. avium</i>	<i>P. avium</i>
Webers Sämpling	Switzerland	<i>P. avium</i>	<i>P. avium</i>
Marasca	Yugoslavia	<i>P. avium</i>	<i>P. cerasus</i>
Gros Guin de Cœur	France	<i>P. avium</i>	<i>P. × gondouinii</i>
Guigne Boissière	France	<i>P. avium</i>	<i>P. × gondouinii</i>
Guin des Charentes	France	<i>P. avium</i>	<i>P. × gondouinii</i>
Meynard	France	<i>P. avium</i>	<i>P. × gondouinii</i>
Frumi	Canada	<i>P. cerasus</i>	<i>P. avium</i>
Griotte jaune d'Ollins	France	<i>P. cerasus</i>	<i>P. avium</i>
Ferracida	France	<i>P. cerasus</i>	<i>P. cerasus</i>
Montmorency1	France	<i>P. cerasus</i>	<i>P. cerasus</i>
Montmorency2	France	<i>P. cerasus</i>	<i>P. cerasus</i>
Griotte de Moissac	France	<i>P. cerasus</i>	<i>P. cerasus</i>
Griotte Layat	France	<i>P. cerasus</i>	<i>P. cerasus</i>

Table 1 (continued)

Variety names	Origin	a priori assignation	a posteriori assignation
Griotte du Nord	France	<i>P. cerasus</i>	<i>P. cerasus</i>
Montmorency3	France	<i>P. cerasus</i>	<i>P. cerasus</i>
Acide Haut Rhin1	France	<i>P. cerasus</i>	<i>P. cerasus</i>
Acide Haut Rhin2	France	<i>P. cerasus</i>	<i>P. cerasus</i>
Acide Haut Rhin3	France	<i>P. cerasus</i>	<i>P. cerasus</i>
Cerise de Toussaint	France	<i>P. cerasus</i>	<i>P. cerasus</i>
Grinque Montmorency	France	<i>P. cerasus</i>	<i>P. cerasus</i>
Griotte précoce	France	<i>P. cerasus</i>	<i>P. cerasus</i>
Montmorency pleureur de Sauvigny	France	<i>P. cerasus</i>	<i>P. cerasus</i>
Iwa	Poland	<i>P. cerasus</i>	<i>P. cerasus</i>
Sabina	Poland	<i>P. cerasus</i>	<i>P. cerasus</i>
Lucyna	Poland	<i>P. cerasus</i>	<i>P. cerasus</i>
Cerise acide Mimbielle	France	<i>P. cerasus</i>	<i>P. cerasus</i>
Griotte de Moissac	France	<i>P. cerasus</i>	<i>P. cerasus</i>
Montmorency4	France	<i>P. cerasus</i>	<i>P. cerasus</i>
Griotte du Lyonnais	France	<i>P. cerasus</i>	<i>P. cerasus</i>
Griotte de Provence	France	<i>P. cerasus</i>	<i>P. × gondouinii</i>
Guigne Boissière	France	<i>P. cerasus</i>	<i>P. × gondouinii</i>
Anglaise Hative1	France	<i>P. × gondouinii</i>	<i>P. avium</i>
Anglaise Hative2	France	<i>P. × gondouinii</i>	<i>P. avium</i>
Cerise Olivet Hative	France	<i>P. × gondouinii</i>	<i>P. cerasus</i>
Cerise Olivet Tardive	France	<i>P. × gondouinii</i>	<i>P. cerasus</i>
Cerise Reine Hortense	France	<i>P. × gondouinii</i>	<i>P. cerasus</i>
Belle Magnifique	France	<i>P. × gondouinii</i>	<i>P. cerasus</i>
Cerise Impératrice Eugénie	France	<i>P. × gondouinii</i>	<i>P. × gondouinii</i>
Gros guin noir de Gironde	France	<i>P. × gondouinii</i>	<i>P. × gondouinii</i>
Maynard	France	<i>P. × gondouinii</i>	<i>P. × gondouinii</i>
La Carrée	France	unknown	<i>P. avium</i>
Petite joue vermeille	France	unknown	<i>P. avium</i>
Belle de Varennes	France	unknown	<i>P. cerasus</i>
Grosse cerise tardive	France	unknown	<i>P. cerasus</i>
Grosse griotte	France	unknown	<i>P. cerasus</i>
Triaux de Fondettes	France	unknown	<i>P. cerasus</i>
Belle Magnifique	France	unknown	<i>P. cerasus</i>
Cerise Royale Tardive	France	unknown	<i>P. cerasus</i>
Cerise Cure1	France	unknown	<i>P. × gondouinii</i>
Cerise Cure2	France	unknown	<i>P. × gondouinii</i>
Guin des Charentes	France	unknown	<i>P. × gondouinii</i>

*A posteriori species repartition of genotypes in the three species was determined according to their coordinates on the first axis of the factorial analysis. For each species, total numbers are indicated in bold.

polyacrylamide gel (4.5% acrylamide/bisacrylamide 20:1, 7.5 M Urea and 0.5 × TBE) and were run at 95 W for 2 h. After electrophoresis, gels were dried on a standard slab gel drier for 2 h and exposed for 5 days to an X-ray film. Each analysis was performed twice for each individual.

Only intense, polymorphic and reproducible bands were taken into account to generate a binary matrix 0–1, where 1 and 0 denote, respectively, the presence and the absence of a band.

Statistical analyses

Species assignation: To properly assign based genotypes to their species, a correspondence analysis based on the AFLP markers was first performed on the whole sample using the procedure CORRESP of the SAS software (SAS Institute, 1985). An *a posteriori* species group was assigned to each genotype according to its coordinates on the first axis of this analysis.

Variation of individual band number (IBN): As AFLP markers are dominant, the phenotype [1] corresponds either to the genotype (11) or to the genotypes (01) or (10), whereas phenotype [0] is only due to the (00) genotype. Let q_i be the frequency of allele 0 at locus i in one species. If we assume that this species can be

Table 2 Sequences of the primers used for AFLP selective amplification and number of reducible polymorphic bands analysed

EcoRI/MseI primers	Primer sequences	Number of markers
<i>EcoRI</i> +AGG	5'-GACTGCGTACCAATTCAGG-3'	20
/MseI+CA	5'-GATGAGTCCTGAGTAACA-3'	
<i>EcoRI</i> +AGG	5'-GACTGCGTACCAATTCAGG-3'	17
/MseI+CG	5'-GATGAGTCCTGAGTAACG-3'	
<i>EcoRI</i> +AAC	5'-GACTGCGTACCAATTC AAC-3'	22
/MseI+CG	5'-GATGAGTCCTGAGTAACG-3'	
<i>EcoRI</i> +AGA	5'-GACTGCGTACCAATTCAGA-3'	16
/MseI+CA	5'-GATGAGTCCTGAGTAACA-3'	
Total		75

considered as being at Hardy–Weinberg equilibrium (random mating, nonoverlapping generations, large population size, mutation ignored and migration overlooked), this allelic frequency q_i can be estimated by $\hat{q}_i^2 = [0]$ in a diploid species and by $\hat{q}_i^4 = [0]$ in an autotetraploid species, where [0] is the frequency of the phenotype [0] for the locus i . Allele frequencies were calculated following this method for each locus in diploid *P. avium* species.

We wanted to know whether an increase of IBN in tetraploid *P. cerasus* species (AAFF), in comparison to *P. avium*, a diploid species (AA), was simply due to polyploidization or could be attributed to specific markers belonging to the F genome. Indeed, autopolyploidization increases IBN since the frequency of the phenotype [1] is $1 - q_i^4$ in a putative autotetraploid species, whereas this frequency is $1 - q_i^2$ in diploid species for each locus. IBN was calculated as the sum of all amplified bands [1] in each genotype considering all the AFLP markers. To determine the effect of autopolyploidization on the IBN, we simulated 500 autotetraploid genotypes (AAAA), formed by the autotetraploidization of *P. avium* with identical allelic frequencies q_i to *P. avium*. For each simulated genotype, the phenotype at each locus i was randomly assigned to [0] or [1] with probabilities \hat{q}_i^4 and $1 - \hat{q}_i^4$, respectively, where \hat{q}_i is the estimated allele 0 frequency in *P. avium*. The mean value of the IBN, in each species, is computed as $\sum_{i=1}^{i=75} (1 - \hat{q}_i^2)$ in diploid species, and $\sum_{i=1}^{i=75} (1 - \hat{q}_i^4)$ in tetraploid species, where i denotes the locus, \hat{q}_i the allele 0 frequency in the species. We compared the IBN distribution in *P. avium*, the 'simulated' autotetraploid species, *P. cerasus* and *P. × gondouinii*.

Determination of genome-specific markers: To compare genetic diversity in each genome, among species with different ploidy levels, the determination of genome-specific markers is needed. In an allotetraploid species like *P. cerasus* (AAFF), markers could be present specifically in one genome or in both genomes. We cannot distinguish if the phenotype [1] in *P. cerasus* is due to marker amplification in A or/and F genome. We estimated allele frequencies of each marker in the A and F genomes in order to detect A and F genome-specific markers. Let q_{A_i} be the frequency of the allele 0 in the *P. avium* genome at the locus i and let q_{F_i} be the frequency of the allele 0 in the F genome of *P. cerasus* brought by *P. fruticosa* through allopolyploidization. Thus, markers for which $q_{F_i} = 1$ are specific to the A genome (no amplification in F genome because the allele 0 is fixed), whereas markers for which $q_{A_i} = 1$ are specific to the F genome markers. Assuming that *P. avium* is at Hardy-Weinberg equilibrium, the allelic frequency q_{A_i} was estimated by $\hat{q}_{A_i} = \sqrt{[0]_{avium}}$, where $[0]_{avium}$ is the frequency of the phenotype [0] in *P. avium*. Several assumptions have to be made to carry on the estimation of q_{F_i} , that is, to identify specific markers. We first assumed that both *P. avium* and *P. cerasus* are weakly spatially structured and that their allelic frequencies are still close to their initial state (ie before the allopolyploidization event). Under these hypotheses, q_{A_i} in *P. avium* and in *P. cerasus* should be closely related. Thus, in *P. cerasus*, $[0]_{cerasus} = \hat{q}_{F_i}^2 \hat{q}_{A_i}^2$, where $[0]_{cerasus}$ is the frequency of the phenotype [0] in this species. For each locus, q_{F_i} was estimated as $\hat{q}_{F_i} = \sqrt{[0]_{cerasus} / \hat{q}_{A_i}^2}$. Many \hat{q}_{F_i} were higher than 1, indicating that the real frequency q_{A_i} in *P. cerasus* is lower than in *P. avium* for these loci; these markers are clearly A genome specific and we considered that $\hat{q}_{F_i} = 1$. For each marker, we plotted the frequencies \hat{q}_{F_i} versus \hat{q}_{A_i} .

Genetic diversity of the A genome in *P. avium* and *P. cerasus*: *P. cerasus* comes from crosses between *P. avium* and *P. fruticosa*. We wanted to know whether the genetic

diversity of the A genome is lower in *P. cerasus* than in *P. avium*. To estimate genetic diversity in the two species, we used average heterozygosity (He), also known as gene diversity (Nei, 1978). Using only A genome-specific markers, we computed He as $He_i = 2q_i(1 - q_i)$, where q_i is the frequency of allele 0 at locus i in the species studied. In *P. avium*, such frequency is q_{A_i} , whereas in *P. cerasus*, this frequency is calculated by $q_{A_i, cerasus} = \sqrt{[0]_{cerasus}}$ using only A-specific markers. For each species, we calculated He as the mean of He_i . We computed $He_{i, avium} - He_{i, cerasus}$ for each locus and tested the mean difference using a sign test.

Results

Species assignment

A total of 75 polymorphic AFLP markers were detected using the four primer pairs, as indicated in Table 2. On the two first axes of the correspondence analysis based on the 75 AFLP markers, three clear-cut groups appeared (Figure 2). The first axis corresponds to 31.5% of total inertia. The *P. cerasus* and *P. avium* genotypes mainly compose the two extreme groups with some misclassified individuals. In the *P. avium* group, one Canadian variety and one old French variety were *a priori* declared as belonging to *P. cerasus*; two old French varieties were previously classified as *P. × gondouinii* (Table 1). The *P. cerasus* group contains three and four genotypes previously classified as *P. avium* and *P. × gondouinii*, respectively. The intermediate group contains nine misidentifications (seven presumed *P. avium* and two *P. cerasus*), four undefined genotypes and only three genotypes identified as *P. × gondouinii*. The coordinates on the first axis of the factorial analysis allow to clearly assign *a posteriori* each genotype to one species, that is, to

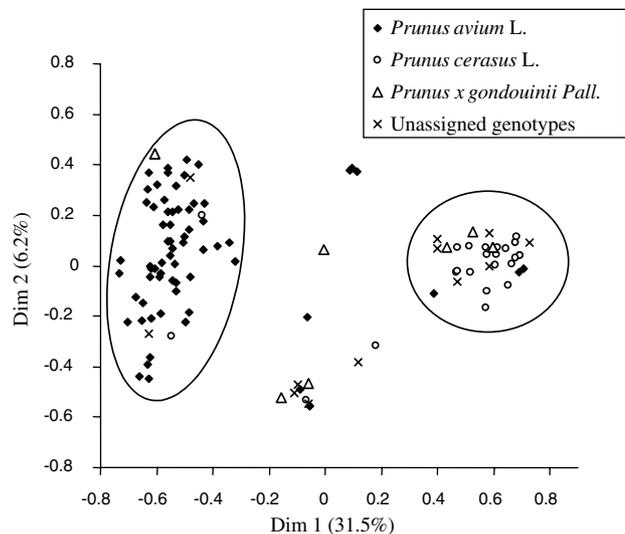


Figure 2 Coordinates of correspondence analysis on 75 AFLP markers of 114 genotypes on the two first axes. Symbols indicate *a priori* assignment genotypes, based on morphological descriptors. The two first axes represent 36.7% of total variation. Each genotype with coordinates on the first axis lower than -0.3 is *a posteriori* classified in *P. avium* species; each genotype with coordinates up to 0.3 is *P. cerasus* genotype, and each genotype with coordinates between -0.3 and 0.3 is *P. × gondouinii* genotype.

classify unknown genotypes in one species and to correct misidentifications as summarized in Table 1. Each sample having a coordinate on the first axis lower than -0.3 was classified as *P. avium* species; each sample with a coordinate higher than 0.3 was classified as *P. cerasus* genotype. Between these two groups, intermediate genotypes are classified in *P. × gondouinii* species. With the 75 markers used, we detected five pairs of identical genotypes: one pair of *P. cerasus*, two pairs previously classified in *P. avium* (a posteriori *P. × gondouinii*) and two pairs previously classified in *P. × gondouinii* (a posteriori *P. avium* and *P. cerasus*). These five pairs have thus the same coordinates on the two first axes of the factorial analysis and are consequently merged in Figure 2.

The second axis, corresponding to 6.2% of total inertia, gives information on intraspecific variation. An analysis of the three species separately did not show clear genetic structure in *P. avium* or in *P. cerasus*. However, among the *P. × gondouinii* genotypes, we noticed three distinct groups also distinguishable in the global factorial analysis, one composed of Italian genotypes previously classified as *P. avium* and the two others were French varieties (data not shown). Using the a posteriori classification, 61 markers (81.33%) of the 75 used are polymorphic in *P. avium*, 55 markers (73.33%) are polymorphic in *P. cerasus* and 52 (69.33%) are polymorphic in *P. × gondouinii*. Thus, the allelic richness is highest in *P. avium*.

Variation of IBN among the different species

When using this a posteriori species classification, the species means of IBN are very different between species: 28.8 and 50.2 for *P. avium* and *P. cerasus*, respectively, and 36.6 for the simulated autotetraploid species (AAAA), far less than values from *P. cerasus* (data not shown). The difference in IBN between this 'simulated autotetraploid' and *P. cerasus* means that the number of markers observed in *P. cerasus* may not be explained by doubling chromosome number alone but may be also due to additional alleles, specific to the second genome of *P. cerasus*, that is, the genome coming from *P. fruticososa*. This supports the hypothesis of the allopolyploid origin of *P. cerasus* and suggests that indeed specific alleles come from the F genome. The *P. × gondouinii* genotypes have a mean IBN value of 38.6, which is intermediate between *P. avium* and *P. cerasus* IBNs.

Identification of genome-specific markers

To detect genome-specific markers, we plotted \hat{q}_F versus \hat{q}_A , the estimated allele frequencies of the 75 markers in the F and the A genomes, respectively. It was then possible to distinguish five classes of markers (Figure 3):

- **Class 1:** When $\hat{q}_F \geq 1$ and $0 < \hat{q}_A < 1$, these markers are always absent in the F genome and polymorphic in the A genome; these 25 markers are called A-specific markers.
- **Class 2:** When $\hat{q}_A = 1$ and $0 < \hat{q}_F < 1$, these markers are always absent in the A genome and are polymorphic in the F genome. These 14 markers are F-specific.
- **Class 3:** When $\hat{q}_F = 0$ and $0 < \hat{q}_A < 1$, these 10 markers are fixed in the F genome and they are polymorphic in the A genome.
- **Class 4:** There are five markers for which $\hat{q}_F = 0$ and $\hat{q}_A = 1$. They are F genome diagnostic markers in our

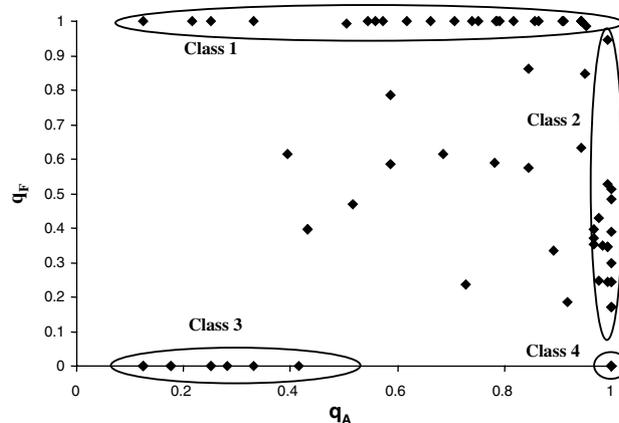


Figure 3 Frequencies of the absent allele (0) for the 75 AFLP markers in *P. avium* genome (q_A) versus in F genome in *P. cerasus* (q_F). Class 1 contains AFLP markers specific to the A genome, class 2 contains AFLP markers specific to the F genome. Class 3 contains AFLP markers fixed in the F genome of *P. cerasus* and polymorphic in *P. avium* and class 4 is composed of five markers always found in the F genome of *P. cerasus* species but never in *P. avium* species.

sample, that is, they are always present in the F genome and always absent in the A genome.

- **Class 5:** When $0 < \hat{q}_F < 1$ and $0 < \hat{q}_A < 1$, these 21 markers show polymorphism on both genomes.

With the A- and F-specific markers, the *P. × gondouinii* genotypes analyzed are separated in three and two groups, respectively, indicating that there could have been at least three origins of the A genome and two origins of the F genome in their hybrid species. This suggests that crosses between *P. avium* and *P. cerasus* may occur recurrently.

Comparison of genetic diversity in A genome between *P. avium* and *P. cerasus*

Gene diversity (H_e) is 0.319 in *P. avium* and 0.173 in *P. cerasus* using only the 25 genome A-specific markers. The difference $H_{e_{P. avium}} - H_{e_{P. cerasus}}$ is positive for 21 markers among the 25 used, and is therefore highly significant using a sign test ($P < 0.001$).

Discussion

Species assignment

Using AFLP markers, the three species *P. avium*, *P. cerasus* and *P. × gondouinii* are clearly distinguished according to their coordinates on the first axis of the correspondence analysis. This separation was confirmed using 725 genotypes of *P. avium*, representing the whole natural range of this species (data not shown). Beaver et al (1995) analyzed seven isozymes markers on 36 *P. cerasus*, *P. avium* and *P. fruticososa* and their hybrids. They identified F-specific markers and showed that the first principal coordinate of the PCO clearly separates the diploid from the tetraploid species, but we were not able to distinguish each species in the tetraploid species group.

AFLP markers are good tools for cherry species identification, especially to clearly distinguish tetraploids. Microsatellite markers have been recently developed in *P. avium* species (Schueler et al, 2003) or

developed in other *Prunus* species and used in *P. avium* (Wunsch and Hormaza, 2002). These microsatellite markers are codominant and should be useful for the study of genetic diversity in specific components of the genome, particularly if primers which only amplify a product from one particular genome can be found.

Several phenomena could explain the number of misidentifications in our sample. Many cherry genotypes in this study were old cherry trees, maintained in national orchards and only few morphological traits were observed for species assignment. Moreover, the distinction between *P. avium* and *P. cerasus* species may not be easy, especially during winter (Hillig and Iezzoni, 1988). The morphological continuum between *P. avium* and *P. cerasus* could be due to the range of cold selection among geographical areas (Iezzoni and Mulnix, 1992) and to complex back-crosses occurring with the local progenitor species (Beaver *et al.*, 1995). The identification of *P. × gondouinii* is still more difficult because of its phenotypic variability, which ranges from *P. cerasus* to *P. avium* morphologic traits.

Specific markers of the F genome have been identified using the frequency of the different markers in both genomes of *P. cerasus*. The diagnostic markers (designated 4 in Figure 3), always present in F genome of *P. cerasus* and always absent in *P. avium* genome in our sample, will now be helpful to detect if a genotype contains the F genome. These could be used for species discrimination, especially to sort *P. avium* from *P. × gondouinii*. It would therefore be beneficial to transform the AFLPs into SCAR markers, to allow efficient analysis of a large number of genotypes (Xu *et al.*, 2001).

Genetic resources characterization, management and conservation are essential for future breeding programs. As species assignment based only on morphological traits could lead to misidentifications, we demonstrated that this assignment could efficiently be carried out with AFLP markers. Moreover, the same AFLP analysis could be used to fingerprint varieties and then to assess the genetic diversity present in germplasm, to detect redundant genotypes in order to constitute and conserve a core collection.

Origin of *P. cerasus*

The difference in distribution of the IBN in *P. cerasus* compared to a simulated theoretical autotetraploid (AAAA) indicates that *P. cerasus* is an allotetraploid containing another genome that is differentiated from the A genome. This fact is in agreement with previous studies indicating that the *P. cerasus* origin was a cross between *P. avium* and *P. fruticosa* (Olden and Nybom, 1968). The presence of 46 genome-specific markers and only 29 markers detected in both the A and F genomes demonstrates that these two genomes are quite divergent but also that some regions of these genomes are conserved, which is in accordance with the segmental allopolyploidy proposed by Beaver and Iezzoni (1993).

The reduction of genetic diversity observed in the genome A of *P. cerasus* in comparison with *P. avium* is probably due to the constitution of *P. cerasus* involving only a part of *P. avium* genetic diversity. However, in this work, mainly French sour cherry trees were analyzed. Therefore, these results have to be confirmed

using a representative sample of whole *P. cerasus* genetic diversity.

The original habitat of wild *P. cerasus* is currently unknown and is difficult to determine because many feral forms have escaped from orchards. It is not possible, therefore, to determine *a priori* which populations of *P. avium* contributed to the A genome of *P. cerasus*. The national ranges of *P. avium* and *P. fruticosa* overlap around the Caspian Sea and Asia Minor; wild *P. cerasus* could be assumed to have appeared in this sympatric area (De Candolle, 1883; Hedrick, 1915). The genome-specific markers detected in our study could be very useful for studying the evolution of one genome among different species with different ploidy levels. We tried to use A genome-specific markers to study the origin of the A genome of *P. cerasus*, that is, to determine which population(s) of *P. avium* has contributed to the constitution of *P. cerasus*. Our inability to unambiguously define the origin of the A genome of *P. cerasus* is probably due to the low number of *P. avium* genotypes in our sample, which is not large enough to represent the whole genetic diversity of this species. Moreover, the low geographic differentiation in *P. avium* makes this determination difficult (Mariette *et al.*, 1997; Mohanty *et al.*, 2001).

The formation of *P. cerasus* and *P. × gondouinii* species both involved unreduced *P. avium* gametes. These abnormalities during meiosis occur at low frequencies and could be affected by genetic and environmental factors (Iezzoni and Hancock, 1984, 1996; Ramsey and Schemske, 1998). Nevertheless, we found two different groups of *P. × gondouinii* using A genome-specific markers. This suggests that the frequency of unreduced gamete production is sufficient to have led to two distinct hybridization events.

Detection of genome-specific markers

There are some limitations in our definition of genome-specific markers. Our sets of specific markers are stable with many *P. avium* samples (unpublished data), but the classification of markers was based on some unproven assumptions. The genetic diversity of *P. avium* has been shown to be weakly spatially structured (Mariette *et al.*, 1997; Mohanty *et al.*, 2001), but we have no data about impacts on the genetic diversity of any population bottleneck which could have accompanied the constitution of *P. cerasus*, nor the genetic drift since this constitution. Moreover, Beaver *et al.* (1995) suggested that *P. fruticosa*, *P. cerasus* and *P. avium* share a common gene pool and/or are continually sharing alleles through introgression.

In each species, the genotypes studied in our sample did not represent the whole intraspecific genetic diversity. For example, *P. cerasus* analyzed in our sample were collected in French orchards and could originate from the same pool of progenitors. That is probably why the correspondence analysis, with only A and F genome-specific markers, respectively, did not reveal the clear genetic structure in *P. cerasus*, which has previously been reported (Iezzoni and Hancock, 1996).

To confirm that F genome-specific markers come from *P. fruticosa*, samples of this species should be genotyped. It would then be easier to classify each set of markers according to their presence or absence in each diploid

species and in tetraploid species carrying the same genome: a procedure which has been successful in *Gossypium* (Abdalla et al, 2001).

A large number of crosses between species could occur within the genus *Prunus* and particularly within the subgenus *Cerasus* (Iezzoni et al, 1990; Webster, 1996). Backcrosses between *P. cerasus* and its progenitors exist. The four species *P. avium*, *P. fruticosa*, *P. cerasus* and *P. × gondounii* are often involved in many crosses with other species, even those from other botanical sections. For example, *P. fruticosa* could cross with *Prunus canescens* (Rehder, 1947). It would be very interesting to study all the *Prunus* subgenus *Cerasus* species, in order to analyse the genetic relationships between them and to have a better knowledge of the crosses involved.

Acknowledgements

We thank those who supplied plant samples, especially R Saunier, F Santi, M Leroux, M Dechaume, B Heinze, J Kleinschmidt, members of the Cytofor European program and members of the European and French networks for *Prunus* genetic resources. We also thank J Ronfort and T Bataillon for their comments on the manuscript. This work was supported by a grant from the 'Association Danone pour les Fruits', by a CEDRE project no. 99 Fb F12/L14.

References

- Abdalla AM, Reddy OUK, ElZik KM, Pepper AE (2001). Genetic diversity and relationships of diploid and tetraploid cottons revealed using AFLP. *Theor Appl Genet* **102**: 222–229.
- Aggarwal RK, Brar DS, Nandi S, Huang N, Khush GS (1999). Phylogenetic relationships among *Oryza* species revealed by AFLP markers. *Theor Appl Genet* **98**: 1320–1328.
- Badenes ML, Parfitt DE (1995). Phylogenetic relationships of cultivated *Prunus* species from an analysis of chloroplast DNA variation. *Theor Appl Genet* **90**: 1035–1041.
- Baird WV, Ballard RE, Rajapakse S, Abbott AG (1996). Progress in *Prunus* mapping and application of molecular markers to germplasm improvement. *Hortscience* **31**: 1099–1106.
- Beaver JA, Iezzoni AF (1993). Allozyme inheritance in tetraploid sour cherry (*Prunus cerasus* L.). *J Am Soc Hortic Sci* **118**: 873–877.
- Beaver JA, Iezzoni AF, Ramm CW (1995). Isozyme diversity in sour, sweet, and ground cherry. *Theor Appl Genet* **90**: 847–852.
- Brettin TS, Karle R, Crowe EL, Iezzoni AF (2000). Chloroplast inheritance and DNA variation in sweet, sour, and ground cherry. *J Hered* **91**: 75–79.
- De Candolle A (1883). *L'origine des plantes cultivées*. Diderot Multimédia: Paris, pp 214–221 (edn 1998).
- Faust M, Suranyi D (1997). Origin and dissemination of cherry. *Hortic Rev* **19**: 263–317.
- Hancock AM, Iezzoni AF (1988). Malate dehydrogenase isozyme patterns in seven *Prunus* species. *Hortscience* **23**: 381–383.
- Hedrick UP (1915). The history of cultivated cherries. In: Hedrick UP, Howe GH, Taylor OM, Tubergen CB, Wellington R (eds) *The Cherries of New York*. JB Lyon company: Albany, NY, pp 39–64.
- Hillig KW, Iezzoni AF (1988). Multivariate analysis of a sour cherry germplasm collection. *J Am Soc Hortic Sci* **113**: 928–934.
- Iezzoni AF, Hancock AM (1996). Chloroplast DNA variation in sour cherry. *Acta Hort* **410**: 115–120.
- Iezzoni AF, Mulinix CA (1992). Variation in bloom time in a sour cherry germplasm collection. *Hortscience* **27**: 1113–1114.
- Iezzoni AF, Schmidt H, Albertini A (1990). Cherries (*Prunus*). In: Moore JN, Bellington Jr JR (eds) *Genetic Resources of Temperate Fruit and Nut Crops 1*. International Society of Horticultural Science: Wageningen, pp 109–173.
- Iezzoni AF, Hancock AM (1984). A comparison of pollen size in sweet and sour cherry. *Hortscience* **19**: 560–562.
- Jones CJ, Edwards KJ, Castaglione S, Winfield MO, Sala F, van de Wiel C et al (1997). Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Mol Breeding* **3**: 381–390.
- Krahl KH, Lansari A, Iezzoni AF (1991). Morphological variation within a sour cherry collection. *Euphytica* **52**: 47–55.
- Krussmann G (1978). *Manual of Cultivated Broadleaved Trees and Shrubs*, Vol. III. PRU-Z.B.T. Batsford Ltd: London, pp 18–58.
- Mariette S, Lefranc M, Legrand P, Taneyhill 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, Martin JP, Aguinalgalde I (2001). A population genetic analysis of chloroplast DNA in wild populations of *Prunus avium* L. in Europe. *Heredity* **87**: 421–427.
- Nei M (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**: 583–590.
- Olden EJ, Nybom N (1968). On the origin of *Prunus cerasus* L. *Hereditas* **70**: 3321–3323.
- Ramsey J, Schemske DW (1998). Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Ann Rev Ecol Syst* **29**: 467–501.
- Rehder A (1947). *Manual of Cultivated Trees and Shrubs*, 2nd edn. Macmillan Company: New York, pp 452–481.
- Santi F, Lemoine M (1990a). Genetic markers for *Prunus avium* L. 2: clonal identifications and discrimination from *P. cerasus* and *P. cerasus* × *P. avium*. *Ann For Sci* **47**: 219–227.
- Santi F, Lemoine M (1990b). Genetic markers for *Prunus avium* L. 1: inheritance and linkage of isozyme loci. *Ann For Sci* **47**: 131–139.
- SAS Institute I (1985). *SAS User's Guide: Statistics. Version 8* (Available at <http://www.zi.unizh.ch/software/unix/statmath/sas/sasdoc/stat/>).
- Saunier R, Claverie J (2001). Le cerisier: évolution de la culture en France et dans le monde. Point sur les variétés, les portegreffe. *Le fruit belge* **490**: 50–62.
- Schueler S, Tusch A, Schuster M, Ziegenhagen B (2003). Characterization of microsatellites in wild and sweet cherry (*Prunus avium* L.). Markers for individual identification and reproductive processes. *Genome* **46**: 95–102.
- Schuster M, Schreiber H (2000). Genome investigation in sour cherry, *P. cerasus* L. *Acta Horticulture* **538**: 375–379.
- Viruel MA, Messeguer R, Vicente MCd, Garcia-Mas J, Pingdomenech P, Vargas F et al (1995). A linkage map with RFLP and isozyme markers for almond. *Theor Appl Genet* **91**: 964–971.
- Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M et al (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res* **23**: 4407–4414.
- Webster AD (1996). The taxonomic classification of sweet and sour cherries and a brief history of their cultivation. In: Webster AD, Looney NE (eds) *Cherries: Crop Physiology, Production and Uses*. Cab International: Wallingford, pp 3–23.
- Wunsch A, Hormaza JI (2002). Molecular characterization of sweet cherry (*P. avium* L.) genotypes using peach (*Prunus persica* L. Batsch) SSR sequences. *Heredity* **89**: 56–63.
- Xu M, Huaracha E, Korban SS (2001). Development of sequence-characterized amplified regions (SCARs) from amplified fragment length polymorphism (AFLP) markers tightly linked to the Vf gene in apple. *Genome* **44**: 63–70.