Theor Appl Genet (2002) 104:661–668

J.C. Herrera · M.C. Combes · F. Anthony
A. Charrier · P. Lashermes

Introgression into the allotetraploid coffee (*Coffea arabica* L.): segregation and recombination of the *C. canephora* genome in the tetraploid interspecific hybrid (*C. arabica*×*C. canephora*)

Received: 26 March 2001 / Accepted: 29 June 2001

Abstract Transfer of desired characters from the diploid relative species such as *Coffea canephora* into the cultivated allotetraploid coffee species (*Coffea arabica* L.) is essential to the continued improvement of varieties. Behaviour of the *C. canephora* genome and its interaction with the *C. arabica* genome were investigated in tetraploid interspecific hybrids (*C. arabica*×*C. canephora* 4*x*) resulting from a cross between an accession of *C. arabica* and a tetraploid plant of *C. canephora* obtained following colchicine treatment. Segregation and co-segregation of restriction fragment length polymorphism (RFLP) and microsatellite loci-markers were studied in two BC_1 populations. These two populations of 28 and 45 individuals, respectively, resulted from the backcross of two tetraploid F_1 plants to *C. arabica*. The presence in BC_1 plants of specific *C. canephora* markers was scored for 24 loci (11 RFLP and 13 microsatellites) distributed on at least 7 of the 11 linkage groups identified in *C. canephora*. At almost all loci analysed, the segregation of *C. canephora* alleles transmitted by the (*C. arabica*×*C. canephora* 4*x*) hybrids conformed to the expected ratio assuming random chromosome segregation and the absence of selection. The recombination fractions of *C. canephora* chromosome segments were estimated for seven marker intervals, and compared with the recombination fractions previously observed in *C. canephora* for the equivalent marker intervals. The recombination frequencies estimated in both plant materials were rather similar, suggesting that recombination in the (*C. arabica*×*C. canephora* 4*x*) hybrid is not significantly restricted by the genetic differentiation between chromosomes belonging to the different genomes. The hybrid (*C. arabica*×*C. canephora* 4*x*) therefore appeared particularly favourable to intergenomic recombination events and gene introgressions.

Keywords Coffee · Wide cross · Introgression · Polyploidy · Evolution

Communicated by H.C. Becker

J.C. Herrera
CENICAFE, AA 2427 Manizales, Caldas, Colombia

M.C. Combes · P. Lashermes (✉)
IRD, GeneTrop, BP 5045, F-34032, Montpellier, France
e-mail: Philippe.Lashermes@mpl.ird.fr
Fax: +33-4-67-41-61-81

F. Anthony
CATIE, AP59, 7170 Turrialba, Costa Rica

A. Charrier
ENSA.M, Place Viala, F-34060, Montpellier, France

Introduction

Coffee is the most-important agricultural export commodity in the world. *Coffea arabica* (Arabica coffee) alone represents 70% of world coffee production and the remainder consists mainly of *Coffea canephora* (Robusta coffee). *C. arabica* is the only tetraploid species (2n=4*x*=44) in the genus and is self-fertile, while other species are diploid (2n=2*x*=22) and generally self-incompatible (Charrier and Berthaud 1985). Recent investigations established that *C. arabica* is an amphidiploid formed by natural hybridisation between *Coffea eugenioides* and *Coffea canephora* (Lashermes et al. 1999). *C. arabica* is characterised by low genetic diversity which has been attributed to the allotetraploid origin, reproductive biology and evolutionary process of this species. In contrast, considerable variability was reported among diploid coffee species, and some of these species form valuable gene reservoirs for different breeding purposes (Carvalho 1988).

Transfer of desirable genes from wild relative species to cultivated varieties through wide crosses is one of the proven breeding strategies for crop improvement. Gene exchange is possible due to the meiotic recombination process, which allows information from the parental chromosomes to be combined into new genetic entities that are passed to the next generation (Stebbins 1950; Puchta and Hohn 1996). However, inherent problems of interspecific crosses, such as hybrid instability, infer-

tility, non-Mendelian segregations, and low levels of intergenomic crossing-over, can constitute important limitations (Stebbins 1958).

In coffee, the transfer of desired characters, in particular disease resistances, from the diploid relative species such as *C. canephora* or *Coffea liberica* into cultivars of *C. arabica* without affecting quality traits, has been the main breeding objective world-wide. Spontaneous hybrids occurring when these species are cultivated together, as well as controlled interspecific hybrids called arabusta (Capot 1972) between *C. arabica* and *C. canephora*, have been extensively used (Cramer 1957; Bettencourt 1973; Sreenivasan et al. 1993). However, the ploidy difference between *C. arabica* and the donor species, and the lack of information regarding genome recombination in interspecific hybrids and DNA introgression into *C. arabica* constitute major bottlenecks.

C. arabica displays a diploid-like meiotic behaviour (Krug and Mendes 1940; Lashermes et al. 2000a). Particular chromosomes of *C. arabica* therefore only pair homogenetically, in spite of the minor differentiation among the two constitutive genomes of *C. arabica* (Lashermes et al. 1999). While triploid interspecific hybrids are almost completely sterile, tetraploid arabusta hybrids, resulting from the hybridisation between *C. arabica* and auto-tetraploid *C. canephora* obtained following colchicine treatment, are reasonably fertile (Berthaud 1978; Owuor and Van der Wossen 1981). However, the meiotic behaviour of the tetraploid arabusta hybrid seems to differ markedly from that of *C. arabica*. Although bivalents have been reported to predominate greatly at meiosis in its hybrid (Grassias 1980; Owuor 1985), the nature of pairing as deduced from the mode of inheritance appears fundamentally different (Lashermes et al. 2000a). Segregation analyses of restriction fragment length polymorphism (RFLP) loci-markers showed tetrasomic inheritance, indicating that the four sets of chromosomes present in the arabusta hybrid might not display any preferential pairing.

Molecular markers represent an invaluable tool for plant genome analysis including polyploids (Da Silva and Sorrells 1996). Factors affecting genetic exchange between parental genomes in polyploid hybrids could be addressed by marker analysis in segregating populations or from patterns of introgression (Wang et al. 1995; Garcia et al. 1995; Parkin and Lydiate 1997). The purpose of the study described here was to analyse the behaviour of the *C. canephora* genome and its interaction with the *C. arabica* genome in the context of tetraploid arabusta hybrids. Using both restriction fragment length polymorphism (RFLP) and microsatellite markers, allele segregation and chromosome recombination were studied in two BC_1 populations derived from tetraploid arabusta hybrids. The investigations have allowed the meiotic patterns of chromosome pairing and recombination to be deduced. The results are discussed in relation to the mechanism of introgression into *C. arabica* and the efficient use of genetic resources in arabica breeding.

Materials and methods

Plant material

The plant material surveyed consisted of two BC_1 populations resulting from the backcross of two interspecific arabusta tetraploid F_1 plants (Et30×IF181T) to *C. arabica* (accession Et30). The populations included 28 (P1) and 45 (P2) individuals, respectively. Each population derived from a single hybrid plant, which resulted from a cross between a plant of *C. arabica* (accession Et30) used as female parent and a tetraploid plant of *C. canephora* (IF181T) previously obtained by colchiploidisation of the clone IF181.

Molecular-marker assay

Genomic DNA was isolated from lyophilised leaves through a nuclei isolation step as described by Agwanda et al. (1997). For restriction fragment length polymorphism (RFLP) analysis, DNA digestion, separation by gel electrophoresis, non-radioactive labelling and hybridisation were performed as previously reported (Lashermes et al. 1995). Single-copy nuclear genomic clones from *C. arabica* or arabusta libraries were used as probes. A preliminary screening was performed on blots containing DNAs of the accessions Et30, IF181, IF181 T, and arabusta hybrids digested by *Eco*RI, *Dra*I, or *Hind*III in order to identify the probe/restriction enzyme combinations revealing appropriate polymorphisms between Et30 and IF181. The suitable probes were then used on Southern blots containing DNAs from all individuals included in this study and digested by the selected restriction enzyme.

Thirteen microsatellites (M12, M20, M41, M42, M47, M157, M160, M161, M162, M166, M170, M180 and M184) that previously showed clear polymorphisms between Et30 and IF181 were pre-selected for use in this study. These microsatellites or simple sequence repeats (SSRs) were identified in DNA clones derived from genomic libraries of *C. arabica* cv Caturra enriched for $(ATC)_n$ and $(TG)_n$ (Rovelli et al. 2000). The specific primer pairs, amplification conditions, radioactive labelling and polyacrylamide gel-electrophoresis were as described by Combes et al. (2000).

Part of the analysed RFLP- and microsatellite-loci have been previously mapped in *C. canephora* using a doubled-haploid population derived from the clone IF200 (Lashermes et al. 2001). These loci, which are distributed on 7 of the 11 linkage groups of the *C. canephora* genetic map, are indicated in Table 1.

Data scoring and analysis

Restriction fragments (i.e. RFLP locus), as well as PCR-amplified products (i.e. microsatellite locus) of different sizes, were identified and easily interpreted as either *C. canephora* or *C. arabica* specific markers by comparing the parental accessions. When two different *C. canephora* specific markers were present in the same arabusta hybrid, the markers were interpreted as alleles of *C. canephora* at RFLP- or microsatellite-loci and designated arbitrarily by the letters C_1 and C_2. When only one *C. canephora* specific marker was identified in the arabusta hybrids at microsatellite loci, allelic interpretation was not undertaken since dosage could not be determined. In contrast for RFLP loci, variations in banding intensity within the same lane were considered to represent differences in allele copy number as previously reported (Lashermes et al. 2000a). The alleles of *C. canephora* were therefore designated by the letter C in single or double dose.

Statistical analysis compared observed segregation patterns to the expected segregation frequencies assuming random chromosome segregation. In addition, linkage analysis was performed between different marker combinations. As expected in the arabusta hybrids used in the present study, four different marker-pair conditions were considered namely, (1) simplex coupling, (2) simplex repulsion, (3) asymmetrical coupling and (4) duplex coupling (Table 3). The recombination fraction for each linked marker pair

was determined by using maximum-likelihood equations according to the method described by Yu and Pauls (1993). The chi-square test was used to assess the goodness-of-fit to the appropriate expected ratios and to compare the recombination fractions.

Results

Identification of *C. canephora* markers in BC_1 individuals

In this study, different molecular markers were used to detect introgression events in the genome of the species *C. arabica*. Analysing segregation patterns of 24 polymorphic loci (11 RFLPs and 13 microsatellites), we scored for the presence of specific *C. canephora* markers in BC_1 plants resulting from backcrosses of two interspecific arabusta tetraploid F_1 plants to *C. arabica*. The two BC_1 populations (i.e. P1 and P2) derived from the two arabusta hybrids were considered separately for comparison (Table 1). Examples of allele segregation patterns for a microsatellite marker are presented in Fig. 1.

For all of the 24 loci analysed, comparison for the frequency of plants with *C. canephora* markers between the two BC_1 populations showed no significant difference ($P>0.05$). In overall analyses including both populations, for a large majority of loci studied (i.e. 83%), the proportions of plants with *C. canephora* markers were consistent with the expected proportion (i.e. 0.83) assuming random chromosome segregation (Table 1). The observed proportions deviated significantly from the expected value for only four loci (gA1, gA25, gA53 and M160). While three of those distorted loci showed abnormally low *C. canephora* marker frequencies, one locus (M160) exhibited a significantly higher frequency than expected. Furthermore, it is noticeable that the disturbed loci belong to at least three different linkage groups of the *C. canephora* map.

Frequency of *C. canephora* alleles in the BC_1 populations

The frequencies of *C. canephora* alleles were determined for 20 loci including 11 RFLP- and 9 microsatellite-loci (Table 2). These loci are distributed on at least 7 of the 11 linkage groups identified in *C. canephora*. For all loci considered, the observed frequencies of *C. canephora* alleles in the P1 population were in agreement with the expected proportion (i.e. 0.25) assuming random allele segregation. Regarding the P2 population, only two loci,

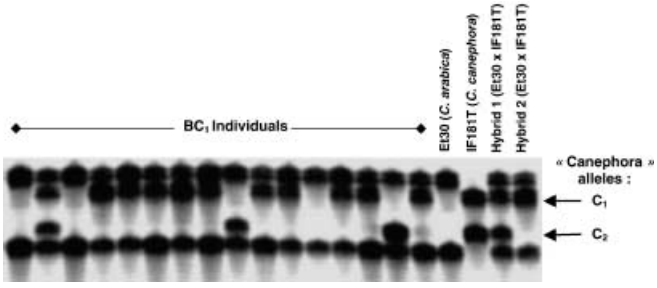

Fig. 1 DNA marker analysis of introgression. Example of microsatellite locus (M166) showing allele segregations among BC_1 individuals resulting from the backcross of arabusta hybrids (*C. arabica*×*C. canephora* 4*x*) to *C. arabica* (Acc. Et30)

Table 1 Proportion of individuals containing specific *C. canephora* markers at different loci in two BC_1 populations (i.e. P1 and P2) resulting from the backcross of arabusta hybrids (*C. arabica*×*C. canephora* 4*x*) to *C. arabica*. Chi-square values for expected Mendelian segregation ratios are also indicated

Locus		Number of plants			Proportion of plants containing *canephora* markers			χ^2 value[b]
Name	Linkage group[a]	P1	P2	Total	P1	P2	Overall	
cR167	4	26	–	26	0.85	–	0.85	0.04 (0.834)
gA1	9	28	45	73	0.86	0.60	0.70	8.95 (0.003)
gA19	9	17	28	45	0.76	0.71	0.73	2.90 (0.089)
gA25	3	12	26	38	0.83	0.54	0.63	10.43 (0.001)
gA53	–	22	40	62	0.59	0.60	0.60	24.11 (0.000)
gA60	–	13	18	31	0.77	0.89	0.84	0.02 (0.886)
gA61	7	28	45	73	0.82	0.93	0.89	1.88 (0.170)
gA71	3	27	43	70	0.81	0.86	0.84	0.08 (0.775)
gA72	7	17	36	53	0.88	0.92	0.90	2.14 (0.143)
gR13	2	27	42	69	0.70	0.83	0.78	1.12 (0.290)
gR41	–	11	27	38	0.91	0.85	0.87	0.42 (0.518)
M12	5	27	42	69	0.77	0.88	0.84	0.05 (0.822)
M20	9	18	36	54	0.94	0.92	0.84	0.03 (0.859)
M41	3	27	40	67	0.78	0.87	0.85	0.15 (0.694)
M42	3	26	41	67	0.92	0.73	0.80	0.27 (0.603)
M47	4	22	35	57	0.86	0.89	0.88	0.91 (0.341)
M157	3	26	42	68	0.92	0.74	0.81	0.20 (0.652)
M160	1	25	41	66	0.92	0.95	0.94	5.57 (0.018)
M161	–	25	42	67	0.84	0.83	0.83	0.02 (0.897)
M162	–	23	35	58	0.83	0.89	0.86	0.44 (0.507)
M166	–	27	42	69	0.67	0.86	0.78	1.12 (0.290)
M170	–	26	40	66	0.81	0.90	0.86	0.52 (0.471)
M180	–	21	35	56	0.80	0.74	0.77	1.55 (0.213)
M184	–	25	42	67	0.92	0.88	0.89	2.05 (0.153)

[a] Genetic linkage groups as determined in *C. canephora* (Lashermes et al. 2001)
[b] Comparison between the overall observed proportion and the expected value (i.e. 0.83) assuming random chromosome segregation. Probability values are indicated in parenthesis

Table 2 Frequency of *C. canephora* alleles at marker-defined loci in BC_1 populations resulting from the backcross of arabusta hybrids (*C. arabica*× *C. canephora* 4*x*) to *C. arabica*. Chi-square values for expected Mendelian segregation ratios are also indicated

Locus		Frequency of *canephora* alleles		Overall frequency	χ^2 value[c]
Name	Linkage group[a]	P1[b]	P2[b]		
cR167	4	0.26	–	0.26	0.05 (0.821)
gA1	9	0.23	0.18*	0.20	3.58 (0.059)
gA19	9	0.22	0.25	0.24	0.12 (0.731)
gA25	3	0.25	0.13**	0.17*	5.05 (0.025)
gA53	–	0.17	0.19	0.19*	5.51 (0.019)
gA60	–	0.29	0.29	0.29	1.08 (0.299)
gA61	7	0.27	0.28	0.28	1.17 (0.280)
gA71	3	0.25	0.28	0.27	0.69 (0.408)
gA72	7	0.24	0.31	0.28	1.23 (0.267)
gR13	2	0.21	0.26	0.24	0.17 (0.677)
gR41	–	0.27	0.24	0.25	0.00 (1.000)
M41	3	0.23	–	0.23	0.20 (0.657)
M42	3	0.28	0.21	0.24	0.18 (0.672)
M157	3	0.28	0.23	0.25	0.00 (1.000)
M160	1	0.30	0.31	0.30*	3.96 (0.047)
M161	–	0.27	–	0.28	0.21 (0.651)
M162	–	0.28	–	0.28	0.52 (0.470)
M166	–	0.20	–	0.20	1.23 (0.267)
M170	–	0.25	–	0.25	0.00 (1.000)
M184	-	0.29	0.31	0.30*	3.90 (0.048)

[a] Genetic linkage groups as determined in *C. canephora* (Lashermes et al. 2001)
[b] For each locus, observed frequency was compared with expected value (i.e. 0.25) assuming random chromosome segregation. *, ** indicate significance at $P<0.05$ and $P<0.01$, respectively
[c] Comparison between the overall observed proportion and the expected value (i.e. 0.25) assuming random chromosome segregation. Probability values are indicated in parenthesis

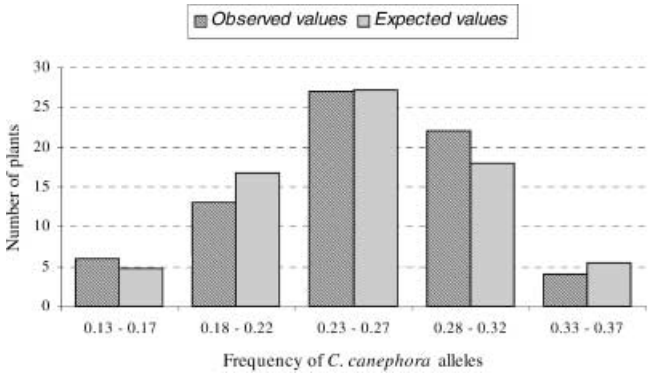

Fig. 2 Histograms of the numbers of BC_1 individuals in which the particular frequency of *C. canephora* alleles was detected. For comparison, the expected values for a theoretical binomial distribution, assuming random segregation at all loci, are also illustrated

gA1 and gA25 (with $P=0.039$ and $P=0.007$, respectively) showed significantly lower values than the expectations. When the overall observed frequencies were analysed, four loci exhibited ratios significantly different from those expected. With the exception of one locus, gA25, all distorted loci showed a similar tendency either in favour (M160, M184), or disfavour (gA53), of *C. canephora* alleles in both populations.

For each BC_1 plant, the frequencies of *C. canephora* alleles were determined. The distribution of the 73 individuals for the frequency of *C. canephora* alleles is presented in Fig. 2. The observed distributions were not significantly different from the expected values for a theoretical binomial distribution assuming random chromosome segregation at all loci. Likewise, the overall mean (i.e. 0.245) of *C. canephora* alleles in BC_1 plants was equivalent to the predicted value (i.e. 0.250).

Recombination of *C. canephora* chromosome segments in arabusta hybrids

Co-segregation studies of specific *C. canephora* markers detected in the tetraploid arabusta hybrids were carried out based on the analysis of BC_1 populations. The expected gamete types and their frequencies with respect to pairs of *C. canephora* markers in different locus conditions are reported in Table 3. Eleven marker-loci, previously mapped in *C. canephora* and distributed on four different linkage groups, were included in this analysis (Table 4). For each pair of adjacent markers, the number of BC_1 individuals for marker-pair phenotypes was determined in the two BC_1 populations.

For the seven marker pairs considered, gamete types unexpected in the absence of recombination were observed (Table 4). This pointed to the presence of crossing-over in the *C. canephora* chromosome segments of the two arabusta hybrids. The observed co-segregation patterns were compared with the expected values for unlinked markers (Table 4). For three of the six marker pairs analysed, the observed co-segregations in at least one BC_1 population were not significantly different ($P>0.05$) from those expected for unlinked markers. The recombination fraction (p) for each marker pair was also estimated in both BC_1 populations. For three of the five marker pairs analysed, the p values estimated in the two BC_1 populations were not significantly different ($P>0.05$). The estimated recombination fractions diverged significantly between the two populations for only two marker pairs, gA71-M157 and M157-M42. Both marker pairs are located on linkage group 3, which showed segregation distortion in population P2 (Table 2).

Table 3 Co-segregation analysis of *C. canephora* markers in BC_1 populations resulting from the backcross of arabusta hybrids (*C. arabica*×*C. canephora* 4*x*) to *C. arabica*. Expected gamete types and their frequencies with respect to pairs of canephora markers in different locus conditions were adapted from Yu and Pauls (1993)

[a] The gamete ratios are equal to the phenotypic ratios observed in the BC_1 population because the recurrent parent (i.e. *C. arabica*) is double recessive for all of the useful markers

[b] The recombination fraction for each marker pair is designated by "p"

Canephora-marker pair conditions in the arabusta hybrid	Types of gametes from the arabusta	General gamete ratios[a, b]	Gamete ratios if C1 and C2 linked[a]	Gamete ratios if C1 and C2 not linked[a]
Simplex coupling (SC)	C1 C2	1–p	1	1
C1 C2.(c1 c2)$_3$	C1 c2	p	0	1
	c1 C2	p	0	1
	c1 c2	1–p	1	1
Simplex repulsion (SR)	C1 C2	1+p	1	1
C1 c2.c1 C2.(c1 c2)$_2$	C1 c2	2–p	2	1
	c1 C2	2–p	2	1
	c1 c2	1+p	1	1
Asymmetrical coupling (AC)	C1 C2	3–p	6	5
C1 C2.C1 c2.(c1 c2)$_2$	C1 c2	2+p	4	5
	c1 C2	p	0	1
	c1 c2	1–p	2	1
Duplex coupling (DC)	C1 C2	5–2p+p^2	35	25
(C1 C2)$_2$. (c1 c2)$_2$	C1 c2	2p–p^2	0	5
	c1 C2	2p–p^2	0	5
	c1 c2	1–2p+p^2	1	1

Table 4 Co-segregation analysis of *C. canephora* markers in the two BC_1 populations. Observed co-segregation patterns and estimated recombination fractions for each pair of marker-defined loci location on the same linkage group of the *C. canephora* map

Linkage groups	Marker intervals	BC_1 pop. (P1/P2)	Marker-pair condition[a]	No. individuals for marker-pair phenotypes in BC_1				χ^2 value[b]	Recombination fraction (p)[c]	Comparison P1/P2[d] χ^2
				C1 C2	C1 c2	c1 C2	c1 c2			
3	M41–gA71	P1	AC	23	19	2	8	4.82 (0.185)	0.206±0.129	
		P2	DC	29	2	3	2	4.24 (0.237)	0.284±0.046	0.85 (0.356)
	gA71–M157	P1	AC	25	17	5	3	2.04 (0.565)	0.521±0.158	
		P2	AC	32	34	2	8	3.56 (0.313)	0.265±0.115	14.88 (0.000)
	M157–M42	P1	SC	13	16	17	5	6.96 (0.073)	0.647±0.067	
		P2	SC	31	6	4	41	49.40 (0.000)	0.122±0.036	40.10 (0.000)
	M42–gA25	P2	AC	12	12	3	15	41.31 (0.000)	0.177±0.136	–
4	M47–cR167	P1	DC	18	1	0	1	–	0.107±0.037	–
7	gA72–gA61	P1	AC	14	14	2	4	–	0.358±0.185	
		P2	SC	29	13	10	16	12.35 (0.006)	0.338±0.057	0.06 (0.806)
9	gA1–gA19	P1	DC	12	1	2	2	–	0.272±0.066	
		P2	AC	17	21	5	11	10.9 (0.012)	0.348±0.146	0.61 (0.435)

[a] SC, simplex coupling; SR, simplex repulsion; AC, asymmetrical coupling and DC, duplex coupling

[b] To fit the expected ratio (see Table 3) for pairs of unlinked markers in SC, SR (1:1:1:1), AC (5:5:1:1), or DC (25:5:5:1) conditions. Probability values are indicated in parenthesis. Comparisons were made only when the population size was sufficient (individuals >35)

[c] The p values were determined by using maximum-likelihood equations according to the method described by Yu and Pauls (1993). Standard errors were estimated according to the method described by Fisher (1937)

[d] Comparison between the recombination fraction values for P1 and P2. Probability values are indicated in parenthesis

Table 5 Comparison for different *C. canephora* chromosome segments of the recombination frequencies estimated in the arabusta hybrid (*C. arabica*× *C. canephora* 4*x*) and in *C. canephora*. The contingency χ^2 tests whether the proportion of parental and recombinant gametes is the same in the two situations

[a] Probability values are indicated in parenthesis

Linkage groups	Marker intervals	Recombination fractions		χ^2 value[a]
		Arabusta hybrid	*C. canephora*	
3	M41–gA71	0.24	0.31	0.86 (0.355)
	gA71–M157	0.37	0.32	0.36 (0.548)
	M157–M42	0.13	0.04	3.94 (0.047)
	M41–M42	0.27	0.40	1.66 (0.198)
4	M47–cR167	0.10	0.07	0.15 (0.700)
7	gA72–gA61	0.37	0.20	6.65 (0.010)
9	gA1–gA19	0.35	0.49	2.29 (0.129)

Comparison of recombination rates in arabusta vs *C. canephora*

The recombination fractions of *C. canephora* chromosome segments estimated in the arabusta hybrids were compared with the recombination fractions previously observed in *C. canephora* for equivalent marker intervals (Table 5). The two BC_1 populations were considered as a whole, and only marker intervals involving non-significantly distorted loci were included in the analysis. Seven marker intervals, distributed on four different linkage groups, were studied. The recombination frequencies estimated in both plant materials were rather similar. Only two of the seven intervals analysed exhibited a significant difference in recombination frequency between the arabusta hybrids and *C. canephora*. Slightly higher recombination fractions were observed in the arabusta hybrids.

Discussion

Efficient use of the genetic resources available in wild relative species is essential to the continued improvement of coffee varieties (Charrier and Eskes 1997). However, breeding programmes face considerable difficulties in doing so. In particular, strong limitations are due to the long generation time of coffee-trees, the high cost of field trials, and the lack of accuracy of the current strategy. New insights into the introgression of *C. arabica* are required to develop improved breeding methodology. In a recent study (Lashermes et al. 2000b), based on the analysis of arabica introgression lines derived from a spontaneous interspecific hybrid, the amount of alien genetic materiel was reported to be significant. In addition, it was suggested that introgression in *C. arabica* is not restricted to chromosome substitutions. The results from this study represent the first attempt to characterise mechanisms of introgression in an interspecific hybrid between *C. arabica* and a diploid relative species. Behaviour of the *C. canephora* genome in tetraploid interspecific hybrids (*C. arabica*×*C. canephora* 4*x*) was successfully investigated by analysing both segregation and co-segregation of *C. canephora* specific markers in two BC_1 progenies. Although *C. canephora* is one of the progenitor species of *C. arabica*, and there is low divergence between the two constituent genomes of *C. arabica* and those of its progenitor species (Lashermes et al 1999), we were able to identify markers specific to the *C. canephora* parent (i.e. IF181 T) of studied hybrids. Moreover, the 24 detected loci presenting *C. canephora* specific markers are well distributed on the genome, representing at least seven different linkage groups.

Interspecific hybrids are often characterised by distorted marker-segregations (Zamir and Tadmor 1986; Ky et al. 2000). Segregation distortion can arise either from competition among gametes or from abortion at the gamete or zygote stage (Lytte 1991; Kreike and Stiekema 1997). In the present study, a remarkably high frequency of *C. canephora* specific markers was observed in both BC_1 populations. The segregation of *C. canephora* alleles transmitted by the arabusta hybrids conformed to the expected ratio in the absence of selection at almost all loci analysed. In relation to its outcrossing nature due to a self-incompatibility system, *C. canephora* is considered to carry a high level of deleterious recessive alleles (Lashermes et al 1994). Furthermore, a zygotic selection due to the expression of lethal or sub-lethal genes in the homozygous condition was indicated to be the principal cause of the pronounced degree of segregation distortion noted in doubled-haploid populations of *C. canephora* (Lashermes et al. 2001). The different behaviour observed in the progenies of arabusta hybrids is likely to be a consequence of the polyploid condition. The buffering capacity of polyploidy against the effects of deleterious recessive alleles has been already demonstrated in several plant species, such as alfalfa and oilseed rape (Sharpe et al. 1995; Brouwer and Osborn 1999).

The absence of important segregation distortion in the progenies of arabusta hybrids indicates also an equal representation of alleles among the functional gametes of the arabusta hybrids. Although characterised by predominantly bivalent pairing and tetrasomic inheritance, the meiotic behaviour of the tetraploid arabusta hybrid is to some extent irregular (Grassias 1980). For instance, significant formations of univalents and multivalents, as well as unbalanced gametes, have been reported. Our results suggest an overall random association of chromosomes in the functional gametes. However, involvement with a low frequency of gametes resulting from irregular meiosis cannot be discounted. In particular, double-reduction events (Burnham 1962) were detected in an especially high proportion in the arabusta hybrid (Lashermes et al. 2000a). The limited segregation deviations observed at a few loci are likely to be due to such meiotic abnormality. Similarly, the slight differences observed in the behaviour of the two analysed arabusta hybrids could be the consequence of irregular meiosis.

The recombination rate of *C. canephora* chromosome segments estimated in the arabusta hybrids was found to be very similar to the recombination frequencies reported in *C. canephora*. The compared equivalent-marker intervals, although widely distributed, do not cover the whole genome, and local differences in recombination frequency may exist. However, our results clearly demonstrate that recombination in the tetraploid arabusta hybrid is not significantly affected by the genetic differentiation between chromosomes belonging to the different genomes. The four sets of chromosomes present in the arabusta hybrid might not only display preferential pairing but also have no difficulty in recombining. Indirectly, this suggests that the genome of *C. canephora* has remained essentially unaltered since the original interspecific hybridisation that formed *C. arabica*, and that these two constitutive diploid genomes are extremely

similar to the genome of modern-day *C. canephora*. This conclusion could be anticipated since *C. arabica* is assumed to result from a very recent speciation (Lashermes et al. 1999). Similar results have been reported in another allotetraploid species, *Brassica napus* (Parkin and Lydiate 1997).

In conclusion, the present study revealed that the arabusta hybrid (*C. arabica*×*C. canephora* 4*x*) is particularly favourable to intergenomic recombination events, and that genes may be more readily introgressed into *C. arabica* than originally believed. In the arabusta context, the introgression of desirable genes into *C. arabica* from *C. canephora*, and most probably from other diploid related species, should not be limited by differences either in sequence homology or in chromosomal structure. Although requiring further investigations, efficient recombination in subsequent generations could be reasonably expected and should reduce any problems associated with linkage drag. Furthermore, *C. arabica* and the arabusta hybrids could constitute an interesting model for examining genome evolution in polyploids (Comai 2000; Wendel 2000).

Acknowledgements This work was supported in part by the European Community through the International Scientific Cooperation Programme (INCO-DC Contract ERBIC18CT970181). The authors are indebted to Dr. Mathias Lorieux (IRD) for skilful recommendations in statistical data analysis, and Prof. Giorgio Graziosi (Universita di Trieste, Italy) for his fruitful co-operation in the development of microsatellite DNA markers in coffee.

References

Agwanda C, Lashermes P, Trouslot P, Combes MC, Charrier A (1997) Identification of RAPD markers for resistance to Coffee Berry Disease, *Colletotrichum kahawae*, in Arabica coffee. Euphytica 97:241–248

Berthaud J (1978) L'hybridation interspécifique entre *Coffea arabica* L. et *Coffea canephora* Pierre. Obtention et comparaison des hybrides triplo, arabusta et hexaplo. Café Cacao Thé 22:87–109

Bettencourt A (1973) Cosideraçoes gerais sobre o "Hibrido de Timor". Instituto Agronomico de campinas, Circular No 31, Campinas, Brasil, p 256

Burnham CR (1962) Discussions on cytogenetics. Burgess Publishing, Minneapolis, Minnesota, USA

Brouwer DJ, Osborn TC (1999) A molecular marker linkage map of tetraploid alfalfa (*Medicago sativa* L.). Theor Appl Genet 99:1194–1200

Capot J (1972) L'amélioration du caféier en Côte d'Ivoire. Les hybrides Arabusta. Café Cacao Thé 16:3–18

Carvalho A (1988) Principles and practices of coffee plant breeding for productivity and quality factors: *Coffea arabica*. In: Clarke RJ, Macrae R (eds) Coffee, vol V/4: agronomy. Elsevier Applied Science, London, pp 129–165

Charrier A, Berthaud J (1985) Botanical classification of coffee. In: Clifford MN, Wilson KC (eds) Coffee: botany, biochemistry and production of beans and beverage. Croom Helm, London, pp 13–47

Charrier A, Eskes AB (1997) Les caféiers. In: Charrier A, Jacquot M, Hamon S, Nicolas D (eds) L'amélioration des plantes tropicales. Collection Repères, CIRAD et ORSTOM, Paris, pp 171–196

Comai L (2000) Genetic and epigenetic interactions in allopolyploid plants. Plant Mol Biol 43:387–399

Combes MC, Andrzejewski S, Anthony F, Bertrand B, Rovelli P, Graziosi G, Lashermes P (2000) Characterisation of microsatellite loci in *Coffea arabica* and related coffee species. Mol Ecol 9:1178–1180

Cramer PJS (1957) A review of literature of coffee research in Indonesia from about 1602 to 1945. Interamerican Institute of Agriculture Sciences, Turrialba, Costa Rica

Da Silva JAG, Sorrells ME (1996) Linkage analysis in polyploids using molecular markers. In: Jahuar PP (ed) Methods of genome analysis in plants. CRC Press, Boca Raton, USA, pp 211–228

Garcia GM, Stalker HT, Kochert G (1995) Introgression analysis of an interspecific hybrid population in peanuts (*Arachis hypogea* L.) using RFLP and RAPD markers. Genome 38:166–176

Grassias M (1980) Etude de la fertilité et du comportement méiotique des hybrides interspécifiques Arabusta *Coffea arabica*×*C. canephora*. PhD thesis, University of Paris XI-Orsay, France

Fisher RA (1937) The design of experiments, 2nd edition. Oliver and Boyd, Edinburgh, London

Ky CL, Barre P, Lorieux M, Trouslot P, Akaffou S, Louarn J, Charrier A, Hamon S, Noirot M (2000) Interspecific genetic linkage map, segregation distortion and genetic conversion in coffee (*Coffea* sp.). Theor Appl Genet 101:669–676

Kreike CM, Stiekema WJ (1997) Reduced recombination and distorted segregation in a *Solanum tuberosum* (2*x*)×*S. spegazzinii* (2*x*) hybrid. Genome 40:180–187

Krug CA, Mendes AJT (1940) Cytological observations in *Coffea.* IV. J Genet 39:189–203

Lashermes P, Couturon E, Charrier A (1994) Doubled haploids of *Coffea canephora*: development, fertility and agronomic characteristics. Euphytica 74:149–157

Lashermes P, Combes MC, Cros J (1995) Use of non-radioactive digoxigenin-labelled DNA probes for RFLP analysis in coffee. In: Berville A, Tressac M (eds) Techniques et utilisations des marqueurs moléculaires. Les Colloques no. 72, INRA, Paris pp 21–25

Lashermes P, Combes MC, Robert J, Trouslot P, D'hont A, Anthony F, Charrier A (1999) Molecular characterisation and origin of the *Coffea arabica* L. genome. Mol Gen Genet 261:259–266

Lashermes P, Paczek V, Trouslot P, Combes MC, Couturon E, Charrier A (2000a) Single-locus inheritance in the allotetraploid *Coffea arabica* L. and interespecific hybrid *C. arabica*× *C. canephora*. J Heredity 91:81–85

Lashermes P, Andrzejewski S, Bertrand B, Combes MC, Dussert S, Graziosi G, Trouslot P, Anthony F (2000b) Molecular analysis of introgressive breeding in coffee (*Coffea arabica* L.). Theor Appl Genet 100:139–146

Lashermes P, Combes MC, Prakash NS, Trouslot P, Lorieux M, Charrier A (2001) Genetic linkage map of *Coffea canephora*: effect of segregation distortion and analysis of recombination rate in male and female meioses. Genome 44:589–596

Lytte TW (1991) Segregation distorters. Annu Rev Genet 25:511–557

Owuor JBO (1985) Interspecific hybridization between *Coffea arabica* L. and tetraploid *C. canephora* P. ex Fr. II. Meiosis in F_1 hybrids and backcrosses to *C. arabica*. Euphytica 34:355–360

Owuor JBO, Van der Vossen HAM (1981) Interspecific hybridization between *Coffea arabica* L. and tetraploid *C. canephora* P. ex. Fr. I. Fertility in F_1 hybrids and backcrosses to *C. arabica*. Euphytica 30:861–866

Parkin IAP, Lydiate DJ (1997) Conserved patterns of chromosome pairing and recombination in *Brassica napus* crosses. Genome 40:496–504

Puchta H, Hohn B (1996) From centiMorgans to base pairs: homologous recombination in plants. Trends in Plant Sci 1:340–348

Rovelli P, Mettulio R, Anthony F, Anzueto F, Lashermes P, Graziosi G (2000) Microsatellites in *Coffea arabica* L. In: Sera T, Soccol CR, Pandey A, Roussos S (eds) Coffee biotechnology and quality. Kluwer Academic Publishers, Dordrecht, pp 123–133

Sharpe AG, Parkin IAP, Keith DJ, Lydiate DJ (1995) Frequent nonreciprocal translocations in the amphidiploid genome of oilseed rape (*Brassica napus*). Genome 38:1112–1121

Sreenivasan MS, Ram AS, Prakash NS (1993) Tetraploid interspecific hybrids in coffee breeding in India. 15th Conf ASIC, Montpellier, France, pp 226–233

Stebbins GL (1950) Variation and evolution in plants. Colombia University Press, New York

Stebbins GL (1958) The inviability, weakness and sterility in interspecific hybrids. Adv Genet 9:147–215

Yu KF, Pauls KP (1993) Segregation of random amplified polymorphic DNA markers and strategies for molecular mapping in tetraploid alfalfa. Genome 36:844–851

Wendel JF (2000) Genome evolution in polyploids. Plant Mol Biol 42:225–249

Wang GL, Dong JM, Paterson AH (1995) The distribution of *Gossypium hirsutum* in *G. barbadense* germplasm: molecular analysis of introgressive plant breeding. Theor Appl Genet 91:1153–1161

Zamir D, Tadmor Y (1986) Unequal segregation of nuclear genes in plants. Bot Gaz 147:355–358