Theor Appl Genet (2004) 109:225–230
DOI 10.1007/s00122-004-1604-1

ORIGINAL PAPER

S. Noir · S. Patheyron · M.-C. Combes · P. Lashermes · B. Chalhoub

Construction and characterisation of a BAC library for genome analysis of the allotetraploid coffee species (*Coffea arabica* L.)

Received: 1 October 2003 / Accepted: 19 January 2004 / Published online: 2 March 2004

Abstract In order to promote genome research on coffee trees, one of the most important tropical crops, a bacterial artificial chromosome (BAC) library of the coffee allotetraploid species, *Coffea arabica*, was constructed. The variety IAPAR 59, which is widely distributed in Latin America and exhibits a fair level of resistance to several pathogens, was chosen. High-efficiency BAC cloning of the high molecular weight genomic DNA partially digested by *Hin*dIII was achieved. In total, the library contains 88,813 clones with an average insert size of 130 kb, and represents approximately eight *C. arabica* dihaploid genome equivalents. One original feature of this library is that it can be divided into four sublibraries with mean insert sizes of 96, 130, 183 and 210 kb. Characterisation of the library showed that less than 4.5% of the clones contained organelle DNA. Furthermore, this library is representative and shows good genome coverage, as established by hybridisation screening of high-density filters using a number of nuclear probes distributed across the allotetraploid genome. This Arabica BAC library, the first large-insert DNA library so far constructed for the genus *Coffea*, is well-suited for many applications in genome research, including physical mapping, map-based cloning, functional and comparative genomics as well as polyploid genome analyses.

Introduction

Coffee is one of the most important agricultural export commodities in the world. *Coffea arabica* (Arabica coffee) accounts for about 70% of the total coffee production. It is the only polyploid species ($2n = 4x = 44$) in the genus, while other species are diploid ($2n = 22$). Recent investigations have established that *C. arabica* is an amphidiploid formed by hybridisation of two closely related diploid species, *C. eugenioides* and *C. canephora* (Lashermes et al. 1999). In addition, the low divergence between the two diploid constitutive sub-genomes of *C. arabica*, namely C^a and E^a, and those of its modern-day progenitor species, C (*C. canephora*) and E (*C. eugenioides*) respectively, suggests that *C. arabica* results from a very recent speciation (Lashermes et al. 1999; Herrera et al. 2002).

Availability of a large-insert genomic DNA library is crucial for physical mapping, map-based gene cloning, and analysis of gene structure and function in most organisms, including plant species. In recent years, bacterial artificial chromosome (BAC) vectors have emerged as the cloning system of choice (Choi and Wing 2000). The easy handling and propagation of the clones, together with their relatively high stability and low degree of chimerism compared with yeast artificial chromosome (YAC) vectors, make BACs an invaluable tool (Osoegawa et al. 2000). In plants, numerous BAC libraries have been constructed and successfully used for a variety of applications including positional cloning and genome sequencing (Shizuya et al. 1992; Martin et al. 1993; Sasaki and Burr 2000). So far, no large-insert genomic DNA library has been constructed for coffee. Availability of a BAC library would therefore constitute an important breakthrough in coffee genome research.

In this paper, we report the construction and the characterisation of an eight-genome equivalent BAC library from the Arabica cultivar IAPAR 59. We demonstrate that this BAC library is a valuable resource for genome analysis in *C. arabica* by carrying out hybridisation screens using mitochondrial and chloroplast DNA probes as well as nuclear probes distributed across the allotetraploid genome.

Communicated by J.W. Snape

S. Noir · M.-C. Combes · P. Lashermes (✉)
IRD, GeneTrop, BP 64501, 34394 Montpellier, Cedex 5, France
e-mail: Philippe.Lashermes@mpl.ird.fr
Fax: +33-4-67416181

S. Patheyron · B. Chalhoub
INRA URGV, CP 5708, 91057 Evry, Cedex, France

Materials and methods

Plant material

The starting material, consisting of plants of the commercial inbred line (F_5) IAPAR 59 (Sera 2001), was grown under greenhouse conditions for 8 months. Before sampling, plants were placed in the dark for 4 days. Young leaves were collected and stored in liquid nitrogen until needed.

BAC library construction

Isolation of high molecular weight DNA, digestion and ligation were done as recently described by Chalhoub et al. (2004). Nuclei were isolated from 25 g of leaf tissue, embedded in low melting point agarose (InCert Agarose, Bio-Rad) plugs and the high molecular weight DNA was partially digested by the *Hin*dIII restriction enzyme. Restriction fragments were submitted to triple size selection by pulse field gel-electrophoresis (PFGE) using a CHEF-DRIII apparatus (Bio-Rad). Ligation reactions were carried out using the pIndigo BAC vector, prepared for *Hin*dIII cloning according to the 'single tube vector preparation' method (Chalhoub et al. 2004), and four different size-selected partially digested DNA samples. Ligated products from the size-selected ranges 100–150, 150–200, 200–250 and 250–350 kb were denoted ligation 1, ligation 2, ligation 3 and ligation 4, respectively. Competent *E. coli* DH10B cells (Invitrogen) were transformed with each ligation mix by electroporation and transformants were selected on LB-X-gal-IPTG plates containing 12.5 μg/ml chloramphenicol. White colonies were picked using a Genetix Q-Bot and stored in 384-well microtitre plates (Genetix) at −80°C.

BAC library characterisation

Estimation of the insert sizes of BAC clones

BAC clones were randomly selected from each ligation and grown for 48 h at 37°C in 1.5 ml 2YT medium containing chloramphenicol (12.5 μg/ml). The BAC DNA was extracted using an alkaline lysis procedure (Sambrook et al. 1989) and digested with *Not*I (Biolabs). Digested products were separated by PFGE on 1% agarose gels (Gold Seakem) in 0.5×TBE, using a CHEF-DRIII apparatus (Bio-Rad).

BAC filter and Southern blot hybridisations

High-density colony filters were prepared using the Genetix Q-Bot. BAC clones were gridded in double spots onto Hybond-N^+ filters (Amersham-Pharmacia) using a 4×4 pattern as described by Tomkins et al. (1999). In total, 27,648 clones (i.e. 44 plates of ligation 4, 27 plates of ligation 3 and 1 plate of ligation 2) of the BAC library were spotted onto a 22.5×22.5 cm filter. The inoculated filters were incubated for 14 h at 37°C. Colony DNA transfer onto the filter was performed according to standard procedures.

The filters were then hybridised with 15 low-copy-number Arabica probes mapped in *C. canephora*. These probes were selected to represent the 11 linkage groups of the genome of *C. canephora* (Lashermes et al. 2001). In addition, *Cox*III and *NAD*3/*rps*12 mitochondrial clones from *Oenothera* (Schuster et al. 1990) were used, while screening for chloroplast DNA was performed using Arabica PCR-generated probes corresponding to the *trn*D/*trn*T, *trn*K/*mat*K and *trn*Q/*Rps*16 fragments (Taberlet et al. 1991; Johnson and Soltis 1994). Probes were labelled with [^{32}P]-dCTP according to the manufacturer's recommendations (Megaprime DNA labelling kit, Amersham) and hybridisation was carried out as described by Sambrook et al. (1989).

RFLP analysis was performed from positive BAC clones and genomic DNAs of *C. arabica* accessions and *Coffea* diploid species accessions. For each probe, restriction enzymes were chosen to reveal duplicate loci in the Arabica accessions and polymorphism among a set of diploid species accessions of *C. canephora* and *C. eugenioides*, corresponding to modern-day progenitor species of *C. arabica* (Lashermes et al. 2000a). By comparison, these duplicate loci were interpreted as two homeologous loci associated with the two homeologous genomes C^a and E^a expected in the case of the allotetraploid Arabica genome. Positive BAC clones hybridising with a particular probe from *C. arabica* were grown for 48 h at 37°C in 3 ml 2YT medium containing chloramphenicol (12.5 μg/ml). The BAC DNA was extracted using an alkaline lysis procedure (Sambrook et al. 1989). Plasmid DNAs of all BAC clones related to a given probe and genomic DNAs extracted as described by Agwanda et al. (1997) were digested with the appropriate restriction enzyme (i. e. *Hin*dIII, *Eco*RI or *Dra*I) for Southern blotting. Digested fragments were separated on a 1% agarose gel. The separated DNA fragments were transferred to a Hybond-N^+ nylon membrane. Labelling and hybridisation were carried out as above for library screening.

Results

Insert size estimation and organellar DNA contamination control

The Arabica line BAC library we constructed consists of a total of 88,320 clones arranged in four sublibraries corresponding to the different ligation reactions. For each ligation, the estimation of mean insert size was based on PFGE gel analysis of a total of 120 randomly chosen clones digested with the *Not*I restriction enzyme. The estimated mean insert sizes for each ligation reaction (sublibrary) are indicated in Table 1 and examples of insert distribution (*Not*I digestions of minipreps) are illustrated in Fig. 1. Although the mean insert size appeared to be lower than that anticipated based on the gel-size selection, BAC sublibraries with satisfactory mean insert sizes were constructed. Sublibrary 1 shows a mean insert size of 96 kb whereas this was 210 kb for sublibrary 4 (Table 1). The insert size distribution of BAC clones derived from ligation reactions 3 and 4 (i.e. sublibraries 3 and 4) is shown in Fig. 2. Taking into

Table 1 Composition of the Arabica BAC library. The mean insert size, together with the standard deviation, and the insert size range, are given in kilobase pairs

Ligation reaction	Size selection range	Number of clones picked	Mean insert size ± SD	Insert size range	Proportion (%) of the whole library size
1	100–150	34,944	96+/-13	70–120	26%
2	150–200	23,040	130+/-9	120–150	25%
3	200–250	13,824	183+/-39	130–240	20%
4	250–350	16,896	210+/-65	80–310	29%

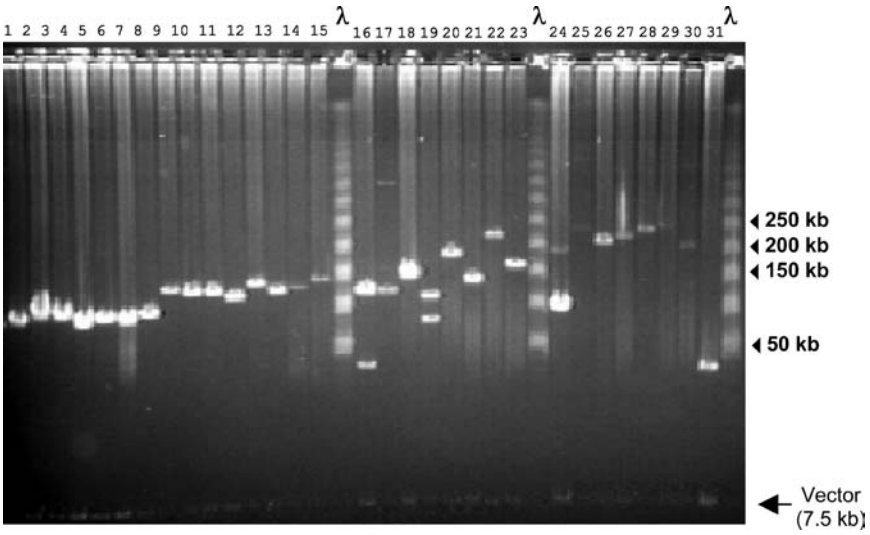

Fig. 1 Pulsed-field gel electrophoresis (PFGE) analysis of BAC clones. Ethidium bromide stained gel of BAC DNA digested with *Not*I and separated in a 1% agarose gel. *Lanes 1–8* BAC DNA from ligation 1, *lanes 9–15* BAC DNA from ligation 2, *lanes 16–23* BAC DNA from ligation 3, *lanes 24–31* BAC DNA from ligation 4. λ is a standard ladder (PFG marker, Biolabs)

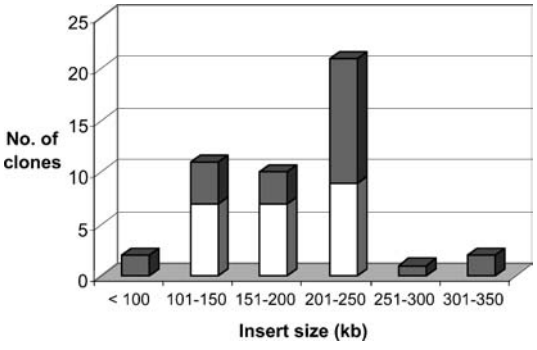

Fig. 2 Insert size distribution of Arabica BAC clones. The data are based on 23 and 24 randomly selected clones derived from ligation reactions 3 (*white*) and 4 (*grey*), respectively

account the relative number of clones picked and the mean insert size of each sublibrary, the average insert size of the library was estimated at 130 kb, and more than 89% of BAC clones had inserts larger than 100 kb. Of the 120 BAC clones analysed, 5 (4%) had no insert, suggesting that 84,787 (96%) of the clones in the library have inserts. Furthermore, only one insert band per clone was generally observed after *Not*I digestion of BAC DNA (Fig. 1).

The percentage of clones in the library containing organelle DNA was determined by colony hybridisation of BAC clone spotted filters. Filters of 27,648 clones were screened with three probes derived from different regions of the chloroplast DNA (*trn*D/*trn*T, *trn*K/*mat*K and the *trn*Q/*Rps*16). Depending on the probe used, the number of positive clones ranged from 1,160 to 1,190, with 1,150 being detected by the three probes. Altogether, in the fraction of the library screened (i.e. 27,648 clones), 1,237 clones contained chloroplast DNA indicating that roughly 4.5% of the BACs in the library had inserts that originated from the chloroplast genome.

Screening of the same BAC filters with two mitochondrial DNA probes, *Cox*III and *NAD*3/*rps*12, detected eight positive clones for each probe, four being in common. Thus, 12 different positive clones were observed suggesting that about 0.0004% of the BAC clones are derived from the mitochondrial genome. In total, less than 4.5% of the BAC clones appeared to be derived from the organellar DNA.

BAC library screening

To test the distribution and genome coverage of the BAC library, 15 RFLP probes distributed on the 11 gametic chromosomes of the *C. canephora* genome (C genome) were used to screen high-density filters containing 27,648 BAC clones and representing approximately 44% of the overall BAC library (Table 2). Twelve of these probes revealed two loci in *C. arabica*, one on each of the two diploid sub-genomes (C^a and E^a) associated in *C. arabica* (Lashermes et al. 2000a). The three remaining probes exhibited only one constant allele in *C. arabica* (Table 2). Thus, the exact number of loci in *C. arabica* remains ambiguous since this figure could reflect the presence of either one locus or two loci (one on each sub-genome) exhibiting the same allele (Lashermes et al. 1999).Positive clones were identified for all 15 probes (Fig. 3). Depending on the probe used, and the number of loci that it revealed, the number of hits per probe ranged from 2 to 16 (7.8 on average). There was a good correlation between the number of positive BAC clones that were detected by each probe and the number of loci revealed (Table 2). An average of four BAC clones were revealed relative to each locus number, which suggests that the fraction of clones analysed represents about four genome equivalents. Taking into account that this fraction is approximately half (44%) of the overall BAC library size, the estimated number of dihaploid genome equivalents would be around eight. The small number of positive BAC clones for the gR19 and the gA11 probes suggests that these probes correspond to single loci in the *C. arabica* genome.

Table 2 BAC library filter-hybridisation results using 15 low-copy number Arabica RFLP probes anchored to the 11 linkage groups of the genome of *Coffea canephora* The 27,268 BACs analysed represent approximately 44% of the overall BAC library

Probe	Linkage group in *C. canephora*	No. of loci in *C. arabica*	No. of hits among the 27,648 BACs tested
gA67	1	2	8
gR19	2	1 or 2	2
gR13	2	2	10
gA14	3	1 or 2	8
gA71	3	2	10
cR167	4	2	2
gA12	5	2	8
gA29	5	2	7
gA11	6	1 or 2	3
gA59	6	2	16
gA61	7	2	12
gA6	8	2	9
gA1	9	2	8
gA13	10	2	6
gR1210	11	2	3

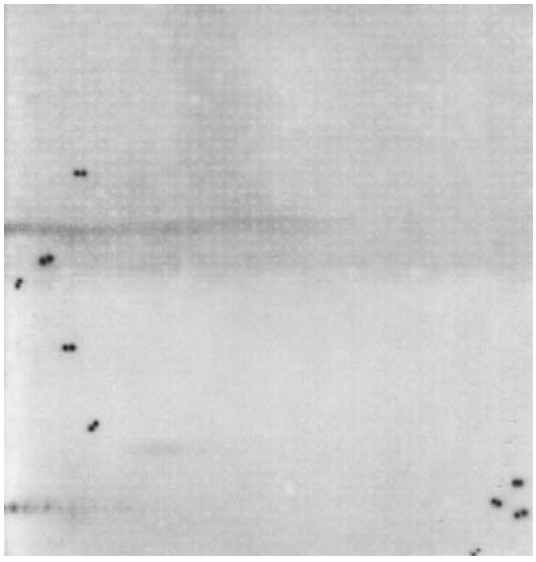

Fig. 3 Colony hybridisation of BAC filters. A total of 27,648 clones were double-spotted onto nylon membrane and hybridised with the nuclear single-copy probe gA6

Particular probe/restriction enzyme combinations were tested by Southern blotting to estimate the contribution of the two diploid constitutive sub-genomes of *C. arabica* in the BAC library derived from the homozygous inbred line IAPAR 59. Positive clones identified from BAC filter screening were analysed for nine selected probes (i.e. gA67, gR13, gA71, cR167, gA29, gA6, gA1, gA13 and gR1210). Four of the 91 BAC clones tested (i.e. 4.4%) did not hybridise with the probe used for their initial identification, and are thus false-positive clones. From BAC clones analysed, we were able to distinguish two alleles for eight of the nine tested loci (Table 3). RFLP pattern comparison between the positive BAC clones and the genomic DNAs of *C. arabica* and *Coffea* diploid species allowed us to assign BAC clones to one or the other sub-genome of *C. arabica* (Table 3). An illustration of this analysis with the gR13 probe is given in Fig. 4. Based on the results, it was estimated that 45% of the BAC clones analysed were derived from the E^a sub-genome and 55% from the C^a sub-genome. No allelic distinction was revealed between the BAC clones identified with the cR167 probe. This could reflect either the low number of positive BAC clones hybridising with this probe or the fact that the positive BAC clones possibly corresponded to a single locus.

Discussion

An Arabica BAC library was successfully constructed from the cultivar IAPAR 59. This introgressed variety, derived from the Timor Hybrid (Lashermes et al. 2000b),

Table 3 Respective number of BAC clones putatively associated with the constitutive sub-genomes of *C. arabica* (C^a or E^a)

Linkage group in *C. canephora*	Probe/restriction enzyme combinations	Number of BAC clones compared	Putative attribution to a sub-genome of *C. arabica*	
			E^a	C^a
1	gA67/*Eco*RI	7	2	5
2	gR13/*Hin*dIII	8	4	4
3	gA71/*Dra*I	8	5	3
4	cR167/*Dra*I	2	–	2
5	gA29/*Dra*I	5	3	2
8	gA6/*Dra*I	9	4	5
9	gA1/*Dra*I	6	1	5
11	gR1210/*Dra*I	2	1	1

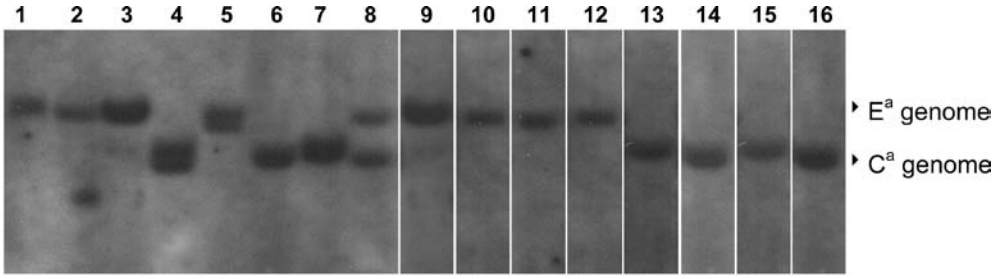

Fig. 4 Southern hybridisation of *Hin*dIII-digested DNA from BAC clones identified by the probe gR13 (*lanes 9–16*) and genomic DNA from *Coffea* species (*lane 8 C. arabica* IAPAR 59 accession, *lanes 1–3 C. eugenioides* accessions, *lanes 4–7 C. canephora* accessions), using the gR13 probe. The putative association between the alleles revealed and the diploid genome is precise

is widely distributed in Latin America, and the cultivar IAPAR 59 was selected for its resistance to leaf rust and root-knot nematodes (Sera 2001). It also presents the introgressed DNA fragment carrying the *Mex-1* locus that confers the resistance to *Meloidogyne exigua* (Noir et al. 2003). The protocol (Chalhoub et al. 2004) adopted in the present study appeared highly satisfactory, generating a high proportion of BACs with large insert sizes. The Arabica library consists of 88,320 clones distributed in four sublibraries with 55% of all BAC clones presenting an insert size above 200 kb.

Although nuclear isolation and purification steps were performed, results of screening with mitochondrial and chloroplast probes showed that approximately 4.5% of the library sequences were derived from organellar DNA. Contamination with chloroplast DNA appeared much more serious than with mitochondrial DNA, suggesting that coffee chloroplast DNA is easier to recover along with the nuclei than that of the mitochondria. Unintentionally, some of these BACs could correspond to the entire chloroplast genome and could be particularly useful when studying the coffee chloroplast genome (Cros et al. 1998). Alternative hypotheses to explain the presence of the chloroplast or mitochondrial DNA-containing BAC clones can be suggested. The exchange of DNA sequences between mitochondria, chloroplasts, and the nucleus is not a rare event. For instance, large insertions of mitochondrial DNA have been observed in the genome of *Arabidopsis thaliana* (Stupar et al. 2001). Studies have shown that nuclear genomes of *Beta vulgaris* and related *Chenopodiaceae* contain large amounts of chloroplast DNA (Ayliffe et al. 1998). Similarly, it has been estimated that the nuclear genome of spinach contains 5–10 copies of the chloroplast genome (Scott and Timmis 1984). Therefore, it remains open whether the organelle DNA sequences found in the Arabica library all originate from chloroplasts and mitochondria, or from integrated nuclear copies.

Analysis of the BAC clones by *Not*I digestion followed by PFGE showed that the majority of Arabica DNA inserts were typically present as single *Not*I fragments. Thus, the Arabica genome apparently contains few *Not*I sites, a feature commonly observed with the genomes of other dicot species (Danesh et al. 1998; Nam et al. 1999; Deng et al. 2001) and contrary to the results obtained with monocot species (Wang et al. 1995; Yang et al. 1997).

Taking into consideration the BACs without insert and the organelle clones, the library is composed of 80,813 clones of *C. arabica* nuclear DNA. The dihaploid genome size of *C. arabica* has been estimated at 1,300 Mb (Cros et al. 1995). Thus, considering the average BAC insert size of 130 kb, the entire library corresponds to 10,506 Mb of cloned nuclear DNA and would represent approximately eight *C. arabica* genome equivalents. The probability (*P*) of having a particular nuclear sequence represented at least once in this library would be $P=1-e^{N[\ln(1-I/GS)]}=0.9997$ (Clarke and Carbon 1976; *I* being the average insert size, *N* the total of clones in the library and GS the genome size).

Furthermore, hybridisation screening of the library has been performed with 15 distinct probes associated with the 11 chromosomes of the *C. canephora* genome. All these probes, distributed throughout the genome, hybridised with positive BAC clones. The number of positive clones in each case was roughly consistent with the allotetraploid structure and the estimated eight-fold genome coverage of the library. For three probes, the slight difference between the estimated and the expected number of copies per dihaploid genome equivalents (i.e. eight) could result from either a small under-representation of particular sequences, or difficulties in interpreting the data associated with the allotetraploid structure (e.g. estimation of RFLP locus number).

Based on the previous analysis of Lashermes et al. (2000a), appropriate probe/restriction enzyme combinations allowed us to show that two distinct alleles can be observed from BAC clones identified for a given locus. Since the IAPAR 59 variety is an inbred line, the observation of distinct alleles at a same locus supports the allotetraploid structure of *C. arabica*. These observations indicate that both of the two sub-genomes of *C. arabica* are represented in the IAPAR 59 BAC library. In addition, the association between the BAC clones analysed and the C^a or E^a sub-genomes attributed suggested that the constitutive genomes of Arabica are represented in the roughly same proportion in the library. Altogether, these evaluations indicated that the Arabica allotetraploid genome is well represented by the BAC library.

In conclusion, the Arabica BAC library described here represents the first large-insert DNA library for the genus *Coffea*. It contains large inserts, has sufficient genome coverage (8×), combines the two sub-genomes C^a and E^a, and contains a low fraction of clones with mitochondrial or chloroplast DNA. These features indicate that the BAC library is suitable for many applications in classical genome research (i.e. physical mapping, map-based cloning, functional and comparative genomics). In addition, in term of polyploid genome analysis, this BAC library offers the opportunity to reconstruct the phylogeny of homeologous loci in this allotetraploid genome in order to better investigate the mechanisms of speciation of the *C. arabica* species.

Acknowledgements This work was supported in part by the European Community through the International Scientific Co-operation Programme (INCO Contract ICA4-CT-2001-10070). We thank Dr. B. Bertrand CIRAD/Promecafe for the seed material.

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

Ayliffe MA, Scott NS, Timmis JN (1998) Analysis of plastid DNA-like sequences within the nuclear genomes of higher plants. Mol Biol Evol 15:738–745

Chalhoub B, Belcram H, Caboche MN (2004) Efficient cloning of plant genomes into bacterial artificial chromosome (BAC) libraries with larger and more uniform insert size. Plant Biotechnol J (in press)

Choi S, Wing RA (2000) The construction of bacterial artificial chromosome (BAC) libraries. In: Gelvin S, Schilperoort (eds) Plant molecular biology manuals, 2nd edn, suppl IV. Kluwer, Dordrecht

Clarke L, Carbon J (1976) A colony bank containing synthetic Col El hybrid plasmids representative of the entire *E. coli* genome. Cell 9:91–99

Cros J, Combes MC, Chabrillange N, Duperray C, Monnot des Angles A, Hamon S (1995) Nuclear DNA content in the subgenus *Coffea* (Rubiaceae): inter- and intra-specific variation in African species. Can J Bot 73:14–20

Cros J, Combes MC, Trouslot P, Anthony F, Hamon S, Charrier A, Lashermes P (1998) Phylogenetic analysis of chloroplast DNA variation in *Coffea* L. Mol Phylogenet Evol 9:109–117

Danesh D, Peñuela S, Mudge J, Denny RL, Nordstrom H, Martinez JP, Young ND (1999) A bacterial artificial chromosome library for soybean and identification of clones near a major cyst nematode resistance gene. Theor Appl Genet 96:196–202

Deng Z, Tao Q, Chang YL, Huang S, Ling P, Yu C, Chen C, Gmitter FG Jr, Zhang HB (2001) Construction of a bacterial artificial chromosome (BAC) library for citrus and identification of BAC contigs containing resistance gene candidates. Theor Appl Genet 102:1177–1184

Herrera JC, Combes MC, Anthony F, Charrier A, Lashermes P (2002) 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*). Theor Appl Genet 104:661–668

Johnson LA, Soltis E (1994) *matK* DNA sequences and phylogenetic reconstruction in *Saxifragaceae* s. str. Syst Bot 19:143–156

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 interspecific hybrid *C. arabica* x *C. canephora*. J Hered 91:81–85

Lashermes P, Andrzejewski S, Bertrand B, Combes MC, Dussert S, Graziosi G, Trouslot P, Charrier A (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

Martin GB, Brommonschenkel SH, Chunwongse J, Frary A, Ganal MW, Spivey R, Wu TE, Earle ED, Tanksley SD (1993) Map-based-cloning of a protein kinase gene conferring disease resistance in tomato. Science 262:1432–1436

Nam Y-W, Penmesta RV, Endre G, Uribe P, Kim D, Cook DR (1999) Construction of a bacterial artificial chromosome library of *Medicago truncatula* and identification of clones containing ethylene-response genes. Theor Appl Genet 98:638–646

Noir S, Anthony F, Combes MC, Lashermes P (2003) Identification of a major gene (*Mex*-1) from *Coffea canephora* conferring resistance to *Meloidogyne exigua* in *Coffea arabica*. Plant Pathol 52:97–103

Osoegawa K, Tateno M, Woon PY, Frengen E, Mammoser AG, Catanese JJ, Hayashizaki Y, de Jong PJ (2000) Bacterial artificial chromosome libraries for mouse sequencing and functional analysis. Genome Res 10:116–128

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Habour Laboratory Press, New York

Sasaki T, Burr B (2000) International rice genome sequencing project: the effort to completely sequence the rice genome. Curr Opin Plant Biol 3:138–141

Schuster W, Wissinger B, Unseld M, Brennicke A (1990) Transcripts of the NADH-dehydrogenase subunit 3 gene are differentially edited in Oenothera mitochondria. EMBO J 9:263–269

Scott NS, Timmis JN (1984) Homologies between nuclear and plastid DNA in spinach. Theor Appl Genet 67:279–288

Sera T (2001) Coffee genetic breeding at IAPAR. Crop Breed Appl Biotechnol 1:179–199

Shizuya H, Birren B, Kim U-J, Mancino V, Slepak T, Tachiiri Y, Simon M (1992) Cloning and stable maintenance of 300 kb fragments of human DNA in *Escherichia coli* using an F-factor-based vector. Proc Natl Acad Sci USA 89:8794–9797

Stupar RM, Lilly JW, Town CD, Cheng Z, Kaul S, Buell CR, Jiang J (2001) Complex mtDNA constitutes an approximate 620 kb insertion on *Arabidopsis thaliana* chromosome 2: implication of potential sequencing errors caused by large-unit repeats. Proc Natl Acad Sci USA 98:5099–5103

Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol Biol 17:1105–1109

Tomkins JP, Yu Y, Miller-Smith H, Frisch DA, Woo SS, Wing RA (1999) A bacterial artificial chromosome library for sugarcane. Theor Appl Genet 99:419–424

Wang GL, Holsten TE, Song WY, Wang HP, Ronald PC (1995) Construction of a rice bacterial artificial chromosome library and identification of clones linked to the *Xa-21* disease resistance locus. Plant J 7:525–533

Yang D, Parco A, Nandi S, Subudhi P, Zhu Y, Wang G, Huang N (1997) Construction of a bacterial artificial chromosome (BAC) library and identification of overlapping BAC clones with chromosome 4-specific RFLP markers in rice. Theor Appl Genet 95:1147–1154