Tree Genetics & Genomes (2011) 7:781–798
DOI 10.1007/s11295-011-0374-6

ORIGINAL PAPER

# Improving the quality of African robustas: QTLs for yield- and quality-related traits in *Coffea canephora*

**Thierry Leroy · Fabien De Bellis · Hyacinthe Legnate · Edmund Kananura · Gustavo Gonzales · Luiz Felipe Pereira · Alan Carvalho Andrade · Pierre Charmetant · Christophe Montagnon · Philippe Cubry · Pierre Marraccini · David Pot · Alexandre de Kochko**

Received: 29 October 2010 / Revised: 27 January 2011 / Accepted: 7 February 2011 / Published online: 1 March 2011


**Abstract** *Coffea canephora* breeding requires combining sustainable productivity with improved technological and cup quality characteristics. Beverage quality is a complex and subjective trait, and breeding for this trait is time consuming and depends on knowledge of the genetics of its components. A highly variable *C. canephora* progeny resulting from an intraspecific cross was assessed for 63 traits over 5 years. To identify quantitative trait loci (QTLs) controlling agronomic, technological, and quality-related traits, a genetic map comprising 236 molecular markers was constructed, and composite interval mapping was performed. Beverage quality was evaluated in relation to biochemical and cup tasting traits. QTLs were identified for almost half of the traits evaluated, with effects ranging from 6% to 80% of phenotypic variation. Most of them present a consistent detection over years. The strongest QTLs explained a high percentage of the variation for yield in 2006 (34% to 57%), bean size (25% to 35%), content of chlorogenic acids (22% to 35%), sucrose and trigonelline content (29% to 81%), and acidity and bitterness of coffee beverages (30% to 55%). Regions of the *C. canephora* genome influencing beverage quality were identified. Five QTL zones were co-localized with candidate genes related to the biosynthesis of the analyzed traits: two genes coding for caffeine biosynthesis, one gene implicated in the biosynthesis of chlorogenic acids, and two genes implicated in sugar metabolism. This is one of the first studies on the identification of QTLs combining agronomic and quality

Communicated by D. Grattapaglia

**Electronic supplementary material** The online version of this article (doi:10.1007/s11295-011-0374-6) contains supplementary material, which is available to authorized users.

T. Leroy (✉) · F. De Bellis · P. Charmetant · P. Cubry · P. Marraccini · D. Pot
UMR AGAP, CIRAD,
34398 Montpellier, France
e-mail: Thierry.leroy@cirad.fr

C. Montagnon
UMR RPB, CIRAD,
34398 Montpellier, France

H. Legnate
Centre National de Recherche Agronomique (CNRA),
BP 808,
Divo, Ivory Coast

E. Kananura
National Agricultural Research Organisation (NARO),
P.O. Box 185, Kituza, Mukono, Uganda

G. Gonzales
Department of Analytical chemistry, University of Seville,
41012 Seville, Spain

L. F. Pereira
Instituto Agronomico do Paraná (IAPAR), LBI-AMG,
CP 481,
86001–970 Londrina, Paraná, Brazil

A. C. Andrade
Laboratory of Molecular Genetics (LGM-NTBio),
Embrapa Genetic Resources and Biotechnology,
CP 02372,
70770–900 Brazilian Federal District, Brazil

A. de Kochko
UMR DIADE, IRD,
34394 Montpellier, France

traits in coffee. The high variability of quality traits within *C. canephora* and the presence of consistent QTLs offer breeders a promising tool to improve coffee cup quality.

**Keywords** Coffee · QTL · Beverage quality · Yield · Cup tasting · Biochemical traits

**Abbreviations**

| | |
|---|---|
| QTL | Quantitative trait loci |
| MAS | Marker-assisted selection |
| Y_200*X* | Yield for the year 200*X* |
| CY_200*X*/*Y* | Cumulated yield from year 200*X* to 200*Y* |
| BS | Bean size |
| PB | Rate of pea berries |
| CA | Caffeine content |
| TR | Trigonelline content |
| SU | Sucrose content |
| 3C | 3-Caffeoylquinic acid content (3-CQA) |
| 4C | 4-Caffeoylquinic acid content (4-CQA) |
| 5C | 5-Caffeoylquinic acid content (5-CQA) |
| 5F | 5-Feruloylquinic acid content (5-FQA) |
| 34dC | 3,4-Dicaffeoylquinic acid (3,4di-CQA) |
| 35dC | 3,5-Dicaffeoylquinic acid (3,5di-CQA) |
| 45dC | 4,5-Dicaffeoylquinic acid (4,5di-CQA) |
| FR | Fragrance |
| AR | Aroma |
| BO | Body |
| FL | Flavor |
| AC | Acidity |
| BI | Bitterness |
| AF | Aftertaste |
| GL | Global note |
| LG | Linkage group |

## Introduction

Two coffee species have worldwide economic importance, namely *Coffea arabica* L. and *Coffea canephora* Pierre known commercially as Arabica and Robusta, respectively (Wintgens 2004). Large variations within and between cultivated species are observed (Anthony et al. 1993, 2001; Gomez et al. 2009; Cubry et al. submitted for publication). Arabica represents approximately 65% of worldwide coffee production and presents higher quality than Robusta due to its lower bitterness and caffeine content and its more appreciated flavor (for a review, see Leroy et al. 2006). To date, quality has not often been considered before the last steps of the breeding process for *C. canephora* (Charrier and Berthaud 1988). Two main genetic and geographic *C. canephora* groups were identified: the Guinean group from western Africa and the Congolese group from central Africa (Berthaud 1986). Further studies also divided the Congolese group into four subgroups (Montagnon et al. 1992; Dussert et al. 1999). A recurrent selection program has been built on *C. canephora* breeding in Côte d'Ivoire since 1984. This recurrent selection breeding is based on the improvement of complementary Guinean and Congolese populations (Leroy et al. 1993).

The main traits commonly taken into account for *C. canephora* coffee quality improvement are as follows: bean size and extractable soluble solids with respect to technological qualities; aroma precursors, such as sugars, caffeine, trigonelline, lipids, and chlorogenic acids, as biochemical traits; and organoleptic traits assessed by cup tasting. Ky et al. (2001a) described the diversity observed in a number of quality traits, including caffeine, trigonelline, chlorogenic acids, and sucrose in Robusta and Arabica coffees. For most compounds, they indicated that the geographical origin of the plants within the genetic groups was the main factor contributing to variability. Regarding cup quality, Moschetto et al. (1996) reported differences between genetic groups related to preference, aroma, acidity, body, and bitterness. Additionally, the same study concluded that Guinean genotypes were on average inferior to Congolese genotypes for preference and aroma. They also indicated good linear correlation coefficients between preference and certain factors, such as acidity and aroma.

Montagnon et al. (1998) observed that the variations in yield and quality traits were independent within *C. canephora*, meaning that quality could be improved without lowering yield. In the same report, the authors showed that narrow sense heritability was high for caffeine content ($h^2$=0.80) and bean weight (0.73), intermediate for chlorogenic acids (0.36), and low for sucrose content (0.11). For cumulative yield over four harvests, narrow sense heritability values were estimated as intermediate with values ranging from 0.3 to 0.4 (Leroy et al. 1994; Montagnon et al. 2003).

Other results for interspecific hybrids suggested a high value (0.71) for the heritability of trigonelline content (Ky et al. 2001b), with a maternal mode of inheritance. For sucrose content, while Montagnon et al. (1998) indicated that this trait could be difficult to improve because $h^2$ was low, Ky et al. (2000a) found additive transmission among their interspecific hybrids, presenting the possibility of choosing parents for its improvement. Relationships between different chlorogenic acid monomers or dimers were also studied by Ky et al. (1999) using an interspecific cross between *Coffea liberica* and *Coffea pseudozanguebariae*, indicating a linear relationship between caffeoylquinic and dicaffeoylquinic contents. In the same interspecific population, Barre et al. (1998) indicated that caffeine was under polygenic control, with strong genetic effects.

Unraveling the genetic basis of complex traits, such as yield potential and stability and coffee quality, can be

undertaken through the construction of a genetic linkage map followed by quantitative trait loci (QTL) identification.

The first *C. canephora* map was constructed with doubled haploids using restriction fragment length polymorphism (RFLP) markers (Paillard et al. 1996). Another map has been developed using doubled haploids for the analysis of segregation distortion (Lashermes et al. 2001). Recently, a map has been developed (Lefebvre-Pautigny et al. 2010) from a segregating population of 93 individuals resulting from a cross between heterozygous genotypes, using mainly RFLP, simple sequence repeat (SSR), and expressed sequence tag (EST)–SSR markers. Maps were also developed for interspecific crosses (Akaffou et al. 2003; Coulibaly et al. 2003; Ky et al. 2000b; N'Diaye et al. 2007) to identify QTLs involved in contrasting traits between wild species using mainly AFLP markers.

To date, few QTL studies have been performed on coffee, and no QTL report is available related to quality traits for coffee. QTLs have been identified for the incompatibility S locus (Lashermes et al. 1996), pollen viability restoration (Coulibaly et al. 2003), fructification time (Akaffou et al. 2003), morphological traits (N'Diaye et al. 2007), and for somatic embryogenesis capacity (Priyono et al. 2010). However, the genetic diversity from the Guinean pool has not been explored in these QTL studies, and the main *C. canephora* breeding program is based on reciprocal recurrent selection between the Congolese and Guinean pool (Leroy et al. 1993).

The aims of the present study were: (1) to construct an intraspecific genetic linkage map for *C. canephora* using a pseudo-backcross progeny between Guinean and Congolese genotypes, (2) to identify QTLs for yield and quality-related traits, and (3) to draw conclusions related to yield and quality breeding of *C. canephora*.

## Materials and methods

### Plant material

Genetic mapping was performed on an intraspecific population of 273 pseudo-backcross individuals of *C. canephora*. This progeny resulted from a cross between an intergroup progeny genotype (Guinean *410*×Congolese *A03*) used as the female and a Guinean genotype used as the male (*02183*). A trial was set up in 2000 on Divo station at the Centre National de Recherche Agronomique (CNRA) in Republic of Côte d'Ivoire (RCI) (5°46′04.07″N, 5°13′22.09″W, altitude 200 masl). The experimental design was a fully randomized single tree plot. Trees were planted at a high density (2.5×1.5 m, approximately 3,000 plants per hectare). The grandparents *410* and *A03* and the parent *02183* were planted in the same plot as the progeny. Data were collected between 2002 and 2006.

During the trial (2000–2006), the pattern of rainfall was quite normal, with two rainy seasons (May to June and September to November). Eighteen trees from our progeny died during the first dry season; the trees were not irrigated and they were fertilized twice a year with 20 g of 23-10-5 (NPK)+4 Mg. Treatment against weeds was performed six times a year. Coffee leaf rust appeared on the trees, but no treatment was performed for this because this disease had no significant effect on yield.

Cherries from the entire trial were sun-dried. Samples for technological, chemical, and organoleptic analyses were wet processed. After pulping, beans were fermented and then dried without direct sun exposure.

### Evaluation of phenotypic traits

We analyzed each trait separately within each year to take into account environmental annual effects. A total of 21 phenotypic traits were recorded for variable number of years, totalling 57 yearly phenotypic traits. Cumulative yields were analyzed for the 2002 to 2003, 2002 to 2004, 2002 to 2005, 2002 to 2006, 2003 to 2006, and 2004 to 2006 harvests. A total of 63 traits were thus evaluated for this trial. These traits can be split into four classes: yield, technological, biochemical, and organoleptic.

#### *Yield traits*

Yield was calculated by summing the weights of fresh berries (in grams) harvested during the productive period (from the second fortnight of September until the end of December). Fresh red fruits were harvested by hand each month and weighed immediately for each tree. Observations were recorded from the second until the sixth year after planting (i.e., 2002 to 2006) for 248 trees. It was observed that 25 of the 273 trees never produced any fruit, including the 18 trees that died during the first dry season.

#### *Technological traits*

Samples of 250 g with 12% moisture were used to estimate bean size (BS) as the weight of 100 kernels (in grams) and the percentage of round beans, or pea berries (PB, in percentage). Both traits were recorded during 4 years (2003–2006) in samples of 155, 227, 216, and 191 trees, respectively, depending on the availability of fruits.

#### *Biochemical traits*

Biochemical traits were analyzed for the 2003 (103 samples) and 2005 (204 samples) harvests. Caffeine,

trigonelline, and chlorogenic acid content were analyzed by reversed-phase high-performance liquid chromatography (HPLC) with spectrophotometric detection. A 5-μm particle size C18 150×4.6 mm column was used. Mobile phase consisted of a mixture of methanol and an aqueous solution of 5 mM triethylamine and acetic acid (pH 3.0, 30:70 *v*/*v*). Elution was performed in isocratic mode at 1 ml $min^{-1}$. Detection was conducted with a diode array detector. Caffeine and trigonelline were quantified using the peak area (measuring the absorbance at 273 nm for caffeine and 265 nm for trigonelline) by external calibration using standards. Prior to HPLC determination, the powdered assay portion was extracted using an aqueous suspension of MgO at 105°C for 20 min. Quantification of chlorogenic acids was achieved by peak area measurement. Seven chlorogenic acids were measured: three caffeoylquinic acids, 3-CQA (3C), 4-CQA (4C), and 5-CQA (5C); the feruloylquinic acid 5-FQA (5F) and three dicaffeoylquinic acids, 3,4di-CQA (34dC), 3,5di-CQA (35dC), and 4,5di-CQA (45dC). These are the main chlorogenic acids present in *C. canephora* (Ky et al. 2001a).

For chlorogenic acid determinations, the methodology used was reversed-phase HPLC using mobile phases A (2 mM $H_3PO_4$ in methanol 5% (*v*/*v*) pH 2.7) and B (2 mM $H_3PO_4$ in methanol 5% (*v*/*v*) pH 3.9). A gradient program was used for 45 min with different volumes of phases A and B. Chromatograms were recorded at 325 nm. The flow rate was 0.8 ml $min^{-1}$. Sucrose (SU) was quantified by enzymatic spectrophotometric determination on green coffee, as described by Alcazar et al. (2004).

### *Organoleptic traits*

The liquoring method used was adapted from CIRAD protocol (Ribeyre, personal communication) and performed at NARO facilities (Uganda). Samples were received from the experimental field (CNRA, RCI) and stored until testing. A maximum of four samples per day were roasted to medium roast using a PROBERT roasting machine. Each sample was tested in triplicate using 10 g of roasted beans ground to medium size and 200 ml of controlled nearly boiled water (95°C, pH ~7.0 and dry residues <100 mg) for each replicate. After mixing and prior to analysis, time was allowed for coffee powder to settle and for the temperature to cool down to 50°C. The same panel of five well-trained tasters conducted the whole experiment. A sample score using a six-class scale (from 0=not detected to 5=fully detected) was performed on eight coffee quality descriptors: fragrance, aroma, body, flavor, acidity, bitterness, and aftertaste. A global note was attributed by each taster based on a global preference of the beverage. Organoleptic analyses were carried out on the three last harvests (from 2004 to 2006), with an unequal number of samples (trees) analyzed each year: 62 trees in 2004, 112 trees in 2005, and 180 trees in 2006.

### *Data analysis*

The mean, standard deviation and variance were calculated for all traits. Histograms were built, and normality was checked using a Shapiro and Wilk test; an appropriate transformation was performed to reach a normal distribution whenever needed.

## Genotypic analyses and map construction

### *DNA extraction*

Genomic DNA was extracted from ground leaves following an extraction procedure using MATAB buffer adapted from Risterucci et al. (2000). Purification of the extracts was then performed using an anion exchange resin column (Nucleo-Bond AX 20 from Macherey-Nagel, Düren, Germany).

### *Molecular markers*

A set of 147 genomic microsatellites (SSRs) was mapped. They were identified in enriched libraries (Poncet et al. 2004, 2007; Combes et al. 2000; Baruah et al. 2003; Moncada and Mac Couch 2004). We also mapped 69 SSRs derived from EST sequences, genes or BAC end sequences (Leroy et al. 2005; Lin et al. 2005; Geromel et al. 2006).

Twenty fragments of genes were mapped for sugar, lipid, caffeine, and chlorogenic acid metabolism (Bouchet et al. 2005; Geromel et al. 2006). Fragments of genes related to drought tolerance in coffee were also mapped (Freire et al. 2010). The genes mapped are presented in Table 1.

### *Genotyping*

For the genes to be mapped, primers were defined at different locations within genes in promoters, exons, or 3' or 5' UTR regions. Then, PCR amplification was performed for parents, indicating whether a size polymorphism (insertion/deletion) was present. If a size polymorphism was present, the genes were mapped using the same genotyping method as was used for SSRs.

For each PCR reaction, 5 ng of DNA in 5 μl of water was used as a template and mixed in a final reaction volume of 10 μl with 1× buffer (10 mM Tris–HCl, 50 mM KCl, 2 mM $MgCl_2$, 0.001% glycerol), 200 μM dNTPs, 0.10 μM of reverse primer, 0.08 μM of forward primer tailed with M13 sequence, 0.10 μM of fluorescently labeled M13 primer, and 0.1 U of Taq polymerase. PCR amplifications were performed in an Eppendorf Mastercycler ep 384 (Eppendorf, Westbury, NY, USA). The amplification

**Table 1** List of genes mapped

| Marker | Function | GenBank *Coffea*[a] | BLASTx[b] | Species | *E* value | Primer F | Primer R |
|---|---|---|---|---|---|---|---|
| G_4CL_C2 | 4-Coumarate CoA ligase | – | XP_002267459 | *Vitis vinifera* | 9.00E-06 | AAATCCAAAGGCGAATTGTG | ACAAGGTCGGGCATGATTAC |
| G_4CL_C5 | 4-Coumarate CoA ligase | – | XP_002307770 | *Populus trichocarpa* | 1.00E-26 | CCGTACAAGCTCGCTCTATG | AGACACGTGGAGACGGATTC |
| G_4CL_C6 | 4-Coumarate CoA ligase | – | XP_002514904 | *Ricinus communis* | 1.00E-42 | TTGCCCCAGAAATTCTCAAC | GGCGGTTTCATGTTCATTCT |
| G_C18332cenar1 | No hit | – | No hit | – | – | ATGGTGGACATCCTGGTGAG | GCCAGCAAGTACATGGAGTG |
| GC03_CCoAMT_intron1 | Caffeoyl-CoA-O-methyltransferase | **ABO77959** | | | | GCCATAAAAGCCTTCTGCAA | GGCTCTGGCTCTCTTGGATA |
| G_CFS_1D | SAM dependent carboxyl methyltransferase | **BAC43760** | | | | CATATGAATGGAGGCGAAGG | CAATGTCCCGAACTGTTGAA |
| G_CWI_SSR01 | Cell wall invertase | *AM231577* | | | | CAATACGGCATGCATTTGAC | TGTTGAACACGCAATTGACC |
| G_CWI_SSR05 | Cell wall invertase | **ABI17893** | | | | ATGTGGTGCTGATGTGCAGT | GTCACGTGGGATGATGAGAA |
| GC17_CP26_intron1 | Chlorophyll a/b binding protein | – | Q9XF89 | *Arabidopsis thaliana* | 6.00E-23 | ACGACGAGCTCGCCAAGT | GCAGGAAAATTCTTCTGTCAGG |
| G_KO_C1_2 | Ent-kaurene oxydase | *FJ409844* | | | | GCCTCGACCACATCTTTGTT | GGCAGGAGAACAATTCAAGC |
| G_KO_sing2_1 | Ent-kaurene oxydase | **ACQ99375** | | | | CTATGTTGATCGCGTGCATT | ATGGAGCTCAAGAAGCTGGA |
| GC19_MYB61_5′UTR | R2R3 Myb TF | – | P81393 | *Antirrhinum majus* | 1.00E-74 | TCAGCCTGTCCTGCATATTG | TGAGCTTTCTCACAGCAAGG |
| G_NMT_A | SAM dependent carboxyl methyltransferase | **BAC43756** | | | | ACCGCAAACTCGAGAAAGAA | ATCCCCAATTCAATCACCAA |
| G_promSUSY_SSR08 | Sucrose synthase | *AM231581* | | | | CGATTTTACACAAGCGTGACA | TCTTTTCTTTTCTTCCGGATTG |
| G_promSUSY_SSR09 | Sucrose synthase | *AM231580* | | | | CAAACAAAACAGTACAATTCAATCC | ATCCCTGCGAGACCTGACTA |
| GC32_RBCS_intron1 | Ribulose-l,5-bisphosphatecarboxylase | **CAD11990** | | | | TTACTTCCCTTGCCACCAAC | AAGTCTCGTTCTTCAACTTTCCA |
| GC25_SDD_3′UTR_1 | Serine-type endopeptidase | – | O65351 | *Arabidopsis thaliana* | 2.00E-65 | CAGCGATACACGGTCACATT | AGAGCCCCGATTGATCTTCT |
| G_SUSY_SSR12 | Sucrose synthase | *AM231583* | | | | CAAACAAACAGTACAATTCAATCC | ACCCCTGTTTTGCTGTTCAC |
| G_SUSY_SSR14 | Sucrose synthase | **CAI56307** | | | | GGATCTTATCGCAATGAACCA | CCAACAGTGTCCTTGCTGAA |
| G_SUSY2_c6_d | Sucrose synthase | **CAJ32597** | | | | TGGCTGGAGTTTATGGCTTC | CCTTATTATTATTGAGGAGCAACC |

The name of the markers on the map, their function, the mnemonic number in databases, and the primers used for these markers are indicated

[a] *Coffea* spp. sequences. Nucleotide sequence mnemonics are shown in italics and protein sequence mnemonics in bold

[b] Best match when performing a BLASTx search in the NCBI nr database (BLASTx 2.2.24)

program consisted of an initial denaturation cycle of 5 min at 94°C, followed by ten cycles of "touch-down" PCR consisting of 45 s at 94°C, 1 min at 60°C to 55°C, decreasing by 0.5°C each cycle, and 1 min at 72°C. The next 20 cycles consisted of 94°C for 45 s, 55°C for 1 min, and 72°C for 1 min, prior to a final elongation step at 72°C for 5 min.

Fluorescently labeled PCR products were analyzed by electrophoresis on a 6.5% polyacrylamide gel using a LI-COR 4300 automated sequencer (LI-COR Biosciences, Lincoln, NE, USA). Gel images were retrieved, and gels were scored manually. Individuals were scored according to the parental segregation types.

### *Statistical analysis*

Statistical analysis was performed to determine variation in the data and to calculate correlation coefficients. Means over years and coefficients of variation were computed for each trait. These coefficients of variation were calculated considering data throughout harvest years.

Pearson phenotypic correlation coefficients between traits were evaluated, and the significance was tested by a *P* value test. These correlations were calculated for individual annual data for all traits, as well as for mean phenotypic data over years.

### *Map construction*

Due to the high heterozygosity of both parents, a pseudo-testcross strategy was used (Grattapaglia and Sederoff 1994). After building independent datasets with segregating markers for each parent, parental maps were constructed using JoinMap 4 software (Van Ooijen 2006) with Kosambi's mapping function and default parameters. A consensus map was then built for a log-of-odds (LOD) score higher than 4.4 and visualized using Spidermap software (Rami, unpublished).

### QTL analysis

A MapQTL 5 software (Van Ooijen 2004) was used for the detection of QTLs on the consensus genetic map for the 63 traits observed, with between 62 and 246 individuals scored for each trait. The data for all 63 traits were used for QTL detection.

Interval mapping was used to detect QTLs. An LOD threshold was experimentally determined for a given trait using the permutation test of MapQTL 5 with 500 iterations. Specific thresholds were determined for the 1% confidence level for each trait in each linkage group (LG). At the genome-wide level, thresholds were determined for 10%, 5%, and 1% confidence levels. In a second step, composite interval mapping referred as MQM in MapQTL (Multiple QTL models, Jansen and Stam 1994) was used after an automatic cofactor selection allowing the removal of some loci flanking the most important QTLs and then localizing QTLs with smaller effects more precisely. The confidence interval was determined by the LOD-1 method for each QTL.

Allelic effects of the QTLs were estimated as $A_f = [(\mu_{ac} + \mu_{ad}) - (\mu_{bc} + \mu_{bd})]/4$ for female additivity, $A_m = [(\mu_{ac} + \mu_{bc}) - (\mu_{ad} + \mu_{bd})]/4$ for male additivity, and $D = [(\mu_{ac} + \mu_{bd}) - (\mu_{ad} + \mu_{bc})]/4$ for dominance, where $\mu_{ac}$, $\mu_{ad}$, $\mu_{bc}$, and $\mu_{bd}$ are the estimated phenotypic means associated with each of the four possible genotypic classes, ac, bc, ad, and bd, derived from a <ab × cd> cross (Segura et al. 2009). The genetic LGs carrying QTLs were presented using MapChart software (Voorrips 2002).

## Results

### Quantitative trait analysis

A total of 63 quantitative traits were analyzed separately. The mean, SD, min, and max values were calculated for the progeny. Mean values were calculated for three parental genotypes: *410* and *A03* as grandparents, and *02183* as the male parent of the progeny.

### *Yield traits*

Individual yields for each of the 248 trees of the progeny were recorded for five consecutive years (2002 to 2006) along with the calculation of cumulative yields. These five harvests constitute the first cycle of production for young trees. After this fifth harvest, trees were cut for a new production cycle. The 2004 to 2006 harvests were the most productive. For further analyses, all yield data were log transformed (log (yield in grams+1)) to fit data normality.

Summary data for parents and progeny are presented in Table S1. The trees were planted in 2000; yields increased yearly until 2004 and remained stable from 2004 to 2006, with a slight decrease occurring in 2005 due to alternation in production. Yield presented a high variability, with a coefficient of variation reaching 62.2% (Table 2). This high variability is always observed for yield in coffee, as strong environmental and ontogenic effects are observed for young plants.

Correlations were estimated for yearly and cumulative yields after log transformation (see supplementary Table S2). All of the yearly and cumulative data were moderately to highly correlated (0.21 to 0.55), except for yield in 2002 and 2006. The cumulative 5-year yield was highly correlated with yearly yields from 2004 to 2006, with correlation coefficients higher than 0.6.

**Table 2** Mean and coefficient of variation for all traits, calculated from data throughout the 5 years of the experiment

| Type of trait | Trait | Mean | Coefficient of variation in percent |
|---|---|---|---|
| Yield | Y | 1.82 | 62.23 |
| Technological | PB | 36.28 | 34.12 |
| | BS | 9.91 | 19.33 |
| Biochemical | SU | 5.27 | 22.63 |
| | CA | 2.56 | 19.56 |
| | TR | 0.84 | 22.94 |
| | 3C | 9.62 | 20.36 |
| | 5C | 55.06 | 11.16 |
| | 4C | 13.62 | 17.38 |
| | 5F | 11.48 | 26.55 |
| | 34dC | 3.26 | 36.78 |
| | 35dC | 3.27 | 58.19 |
| | 45dC | 4.39 | 53.47 |
| Organoleptic | FR | 2.03 | 13.29 |
| | AR | 1.97 | 14.63 |
| | BO | 1.98 | 16.48 |
| | FL | 1.85 | 20.16 |
| | AC | 0.13 | 99.24 |
| | BI | 1.33 | 39.79 |
| | AF | 1.72 | 22.09 |
| | GL | 1.86 | 23.60 |

*CA* caffeine, *FR* fragrance, *AR* aroma, *BO* body, *FL* flavor, *AC* acidity, *BI* bitterness, *AF* aftertaste, *GL* global note

### *Technological traits*

BS and the rate of PB were measured for harvests from 2003 to 2006. The results for the progeny and for parental genotypes are presented in Table S1. Bean size remained low for the progeny and the rate of pea berries was high, always being higher than 30%. The values observed for the progeny were within the values observed for the parents and grandparents (Table S1). Technological traits exhibited medium coefficients of variation (19% for BS and 34% for PB, Table 2). The phenotypic correlations were highly significant for each trait considering successive years, with values above 0.5 in most cases (Table S2). Significant positive correlations were also observed between PB and BS, except in 2005.

### *Biochemical traits*

Biochemical traits were estimated for the 2003 and 2005 harvests. A high variability was observed every year within the progeny, and values were quite stable from 1 year to another. For all compounds, the values observed for the progeny were included between the values measured for the parents *410* and *A03* (Table S1). Some data were modified to fit normality; for example, data on chlorogenic acids were transformed using the ArcSIN $((x/100)^{0.5})$ formula. The concentration of biochemical compounds in the green beans presented a coefficient of variation ranging from 11.16% (5-CQA) to 58.19% (3,5di-CQA); see Table 2. Phenotypic variability was moderate for caffeine, sucrose, and trigonelline content (19% to 22%). Variability for the different chlorogenic acids was highly variable, with a positive correlation with the compound's mean value. The lowest variability was observed for 5-CQA (11.16%, for a mean of 55.06), and the highest was observed for 3,5di-CQA (58.19%, for a mean of 3.27).

Correlations for biochemical traits are presented in supplementary Table S2. For sucrose, caffeine and 3-CQA, 4-CQA and 5-FQA chlorogenic acid content, a high positive correlation was observed between values in 2003 and 2005. A highly significant positive correlation was observed between 3-CQA and 4-CQA content (0.99 in 2003 and 0.96 in 2005), whereas highly significant negative correlations were observed between these two chlorogenic acids and the other chlorogenic acids, including 5-CQA and 5-FQA.

For the dicaffeoylquinic acids, highly significant positive correlations were observed between the three compounds analyzed in both years of the analysis. Correlations between the contents of all caffeoylquinic acids (except 5-FQA) and the contents of dicaffeoylquinic acids were negative in 2003 (−0.12 to −0.66) and were significantly negative with the three di-CQAs in 2005 (−0.34 to −0.43).

Additionally, no significant correlation was observed between sucrose, caffeine or trigonelline content, and the content of other compounds, except for caffeine content in 2003 and dicaffeoylquinic acid contents in 2005.

### *Organoleptic traits*

Each coffee sample produced from each plant was individually tested for its organoleptic characteristics for three successive years from 2004 to 2006. Small differences were observed for all traits between genotypes and years. The Guinean parents of the progeny (*410* and *02183*) presented much higher bitterness than the Congolese grandparent *A03*, but the differences were quite small for acidity. Within the progeny, the bitterness was quite high but decreased over the years. Meanwhile, acidity remained low for all years, with values also decreasing from year to year. The global preference notation given by the judges decreased in 2006, along with decreasing aroma and body (Table S1). The variability of organoleptic traits (Table 2) was generally low to medium (13% to 23%), except for acidity and bitterness, which showed high levels of variability (99% and 40%, respectively) with very low mean values. The high level of variability found for acidity

was due to its very low content for the beans tested, and thus, this trait was very difficult for the tasters to estimate.

Significant correlations between organoleptic traits were observed each year, but not between years, indicating the independence of successive years (Table S2). As expected, acidity was highly positively correlated with the global preference note, with values ranging from 0.49 to 0.60, but it was negatively correlated with bitterness (−0.33 to −0.37).

Phenotypic correlations among all mean trait values over the years were evaluated (Table 3), including cumulative yield over five years. Significant but low positive correlations were observed between cumulative yield (CY_2002/6), bean size (0.21), and sucrose content (0.15). Cumulative yield was also significantly correlated with favorable organoleptic traits, such as acidity, flavor, and the global note (0.19 to 0.29). Technological traits (bean size and rate of pea berries) presented highly significant correlations with 3-CQA (0.15 to 0.17) and with 5-FQA (negative correlations from −0.17 to −0.22). Finally, the correlations between flavor, acidity, and the global note, with respect to organoleptic traits, and sucrose and 5-CQA content, among biochemical traits, were highly and positively significant. A negative correlation was observed between flavor and acidity with trigonelline and 5-FQA content. Bitterness was highly negatively correlated with sucrose (−0.25) and was highly positively correlated with caffeine content (0.20).

## LG analysis

A total of 248 markers were initially used for genotyping the population of 248 individuals. Markers or genotypes with more than 10% of missing data were eliminated. Finally, 238 markers were mapped using 184 individual trees. Eleven LGs were constructed, corresponding to the 11 gametic chromosomes of *C. canephora*. The total length of our consensus map was 1,290 cM, with an average distance of 5.5 cM between markers and a maximum distance of 37 cM between markers. The length of the LGs was variable, ranging from 57.8 cM for LG I to 243.3 cM for LG B. Segregation distortion was observed for all markers in LG I, with $\chi^2$ test values found to be highly significant ($p<0.05$ to 0.0001). The consensus map is presented in Fig. 1. We built both parental maps, but as the map for the male Guinean parent *02183* was sparse, we only used the female map to confirm the position of markers on the consensus map.

Sixteen genes were mapped to nine LGs. These genes were identified in EST libraries in the framework of our collaboration with Brazilian teams from IAPAR and EMBRAPA/CENARGEN (Geromel et al. 2006; Freire et al. 2010). Two genes involved in caffeine metabolism were mapped to LGs A and I. One gene (two markers) related to lipid metabolism has been mapped to LG J. Four genes involved in the biosynthesis of chlorogenic acids were mapped to LGs B, F, G, and H. For CGA biosynthesis, we mapped one CCoAOMT gene that was previously mapped by Campa et al. (2003) and three 4CL genes. This last enzyme acts at different levels in the CGA metabolic chain pathway (Lepelley et al. 2007; Joët et al. 2009). Five putative candidate genes implicated in drought tolerance were mapped on to the B, C, F, G, and J LGs. Finally, four genes for sucrose metabolism were mapped to LGs A, D, and F (five markers from two sucrose synthase genes, *SUS1* and *SUS2*, and two *CWI* genes encoding cell wall invertases). With respect to previous work on the characterization of a *C. canephora* BAC library using sucrose synthase genes (Leroy et al. 2005), the copy number of these genes were confirmed, with one copy found for *SUS2* and three copies found for *SUS1*.

## QTL detection

QTL detection was performed from a synthetic map with 209 markers. Markers that were very close to each other were eliminated for QTL detection. Phenotypic data were available for all of the trees, and these trees were integrated in the QTL study independently of their elimination in map construction. The QTL detection was, thus, performed using all of the 248 trees of the progeny. Figure 2 presents the main QTLs identified per LG for all data and Table 4 presents all of the QTLs and their characteristics.

Seven QTLs were identified for yield (Table 4). One main QTL zone was located in LG K for the 2006 yield and the cumulative yield for 3 to 5 years of harvest. Two QTLs that explained 34% to 42% of the phenotypic variations for the 2006 yield were identified in LG I, and one QTL was identified in the D LG for the 2003 yield. Dominance effects were predominant for cumulative yield.

For the rate of pea berries, six QTLs were identified in LGs F, J, and K. A female additive effect was found to be the most important effect for this trait. Eleven QTLs for bean size were found in LGs A, B, D, F, and J, with little consistency throughout the years, and quite surprisingly, dominance effects were the most important for this heritable trait. One QTL in LG A explained 35% of the variation in this trait in 2005, while two QTL zones in LG B explained up to 25% of the trait variation.

Co-localization between agronomic and technological traits was limited to LG F, where a QTL for bean size and the rate of pea berries co-localized. With respect to biochemical traits, 27 QTL zones were implicated in this study (Table 4). Female additive effects were quite important for chlorogenic acids and caffeine. For sucrose and trigonelline content, male additive and dominant effects were also found. QTLs were identified for all traits analyzed. For these

**Table 3** Phenotypic correlations among mean values of the traits over the years (Pearson's correlation test)

| | BS | CY_2002/6 | FR | AR | BO | FL | AC | BI | AF | GL | SU | CA | TR | 3C | 5C | 4C | 5F | 34dC | 35dC | 45dC |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PB | 0.19** | ns | 0.22** | ns | ns | ns | ns | ns | ns | ns | ns | ns | ns | 0.17** | ns | 0.16* | −0.17** | ns | ns | ns |
| BS | | 0.21** | ns | ns | 0.14 * | ns | 0.15* | ns | ns | ns | ns | 0.16* | 0.17* | 0.15* | ns | ns | −0.22** | ns | 0.14* | ns |
| CY_2002/6 | | | ns | ns | 0.19** | 0.19** | 0.18** | 0.18* | 0.29** | 0.20** | 0.15* | ns | ns | ns | ns | ns | ns | ns | ns | ns |
| FR | | | | 0.29** | 0.18** | ns | ns | ns | 0.18* | ns | ns | ns | ns | 0.16* | ns | ns | ns | ns | ns | ns |
| AR | | | | | 0.16* | 0.19** | ns | ns | 0.15* | ns | ns | ns | ns | ns | ns | ns | ns | ns | ns | ns |
| BO | | | | | | 0.42** | 0.29** | ns | 0.45** | 0.44** | ns | ns | ns | ns | ns | ns | ns | ns | ns | 0.17* |
| FL | | | | | | | 0.59** | −0.26** | 0.53** | 0.74** | 0.23** | ns | −0.18* | ns | 0.26** | ns | −0.15* | ns | ns | ns |
| AC | | | | | | | | −0.19** | 0.50** | 0.66** | 0.30** | ns | −0.24** | ns | 0.18* | ns | −0.18** | ns | ns | ns |
| BI | | | | | | | | | ns | −0.23** | −0.25** | 0.20** | ns | ns | ns | ns | ns | ns | ns | ns |
| AF | | | | | | | | | | 0.59** | 0.14* | ns | ns | ns | 0.18* | ns | −0.16* | ns | ns | ns |
| GL | | | | | | | | | | | 0.25** | ns | ns | ns | 0.29** | ns | ns | ns | ns | ns |
| SU | | | | | | | | | | | | ns | ns | −0.15* | 0.25** | −0.15* | −0.16* | ns | ns | ns |
| CA | | | | | | | | | | | | | 0.21** | −0.15* | ns | −0.13* | ns | 0.15* | 0.22** | 0.20** |
| TR | | | | | | | | | | | | | | ns | ns | ns | 0.26** | ns | ns | ns |
| 3°C | | | | | | | | | | | | | | | −0.27** | 0.96** | −0.39** | −0.36** | −0.43** | −0.43** |
| 5°C | | | | | | | | | | | | | | | | −0.22** | −0.19** | −0.44** | −0.35** | −0.44** |
| 4°C | | | | | | | | | | | | | | | | | −0.41** | −0.41** | −0.47** | −0.47** |
| 5°F | | | | | | | | | | | | | | | | | | −0.17* | ns | ns |
| 34dC | | | | | | | | | | | | | | | | | | | 0.82** | 0.88** |
| 35dC | | | | | | | | | | | | | | | | | | | | 0.78** |

*ns* not significant, *FR* fragrance, *AR* aroma, *BO* body, *FL* flavor, *AC* acidity, *BI* bitterness, *AF* aftertaste, *GL* global note

*$P<0.05$; **$P<0.01$; only significant correlations are shown

three traits, no consistency was observed throughout the years, but some QTLs in LGs I and K for trigonelline content explained a large part of the variation of the trait: 81.2% and 41.8% in 2003 and 2005, respectively.

Eight QTLs were identified for 3-CQA and 4-CQA contents in LGs A, B, and I, with a good consistency throughout years. Both QTLs for 4-CQA content in 2003 and 2005 in LG B represented 28.8% and 40.6% of the variation in this trait, respectively (26.7% and 16% for 3-CQA). For 5-FQA, six QTLs were identified in LGs A, B, D, F, and I, with a common QTL found in LG I in 2003 and 2005. The QTLs for this trait in LG F explained up to 35% of its variation in 2003.

In LG A, a unique QTL zone was related to both 4-CQA and 5-FQA content, which are negatively correlated in 2005. In LG B, a large QTL zone included QTLs for 3-CQA, 4-CQA, 5-CQA, and 5-FQA content. In LG D, a small zone included QTLs for 5-FQA and 3,5di-CQA. Another zone in the same LG includes QTLs for trigonelline, 5-FQA, and 4-CQA. In LG I, the QTL zone included QTLs for caffeine (2005), CGA monomers and sucrose in 2005. In LG J, a small zone included QTLs for 3,4 and 3,5di-CQA. In LG K, one QTL zone was identified for caffeine and trigonelline in 2005, and another was found for 5-CQA in 2003.

Six QTLs were identified for organoleptic traits (Table 4). One QTL was identified for the global note in LG H (not presented in Fig. 2). For bitterness (2004), one QTL was identified in LG D. In LG I, we observed a QTL zone for acidity in 2006 and bitterness in 2005 and 2006. This result is consistent with the negative phenotypic correlations observed between these traits. For bitterness, female additive effects were predominant, while for acidity, only a male additive effect was identified.

All of the QTLs found for organoleptic properties explained more than 15% of the trait variability and up to 54.8% for acidity in LG I. The major co-localizations of QTL for traits from different sources of data were as follows:

- In LG B, bean size in 2004 and 2005 with caffeoylquinic acids in 2005
- In LG I, co-localization between bitterness, acidity (in 2006), sucrose content (2005), content of 3-CQA, 4-CQA and 5-FQA (2003 and 2005), caffeine content (2005), and yield (2006) was found.
- In LG J, bean size (2006) with dicaffeoylquinic acids (2005)

Some QTL zones co-localized with genes implicated in different metabolic pathways related to coffee quality. In LG A, a QTL for caffeine content co-localized with a caffeine synthase gene implicated in the last steps of caffeine biosynthesis (Misako and Kouichi 2004). In LG B, a composite QTL zone, including all caffeoylquinic acids and bean size was localized near a 4CL gene implicated in the biosynthesis of chlorogenic acids and near genes encoding two invertase genes (mapped from BAC ends sequences). In LG D, QTLs for bean size (2006) were located in the same zone as a gene encoding a cell wall invertase *CWI* and a copy from a *SUS1* gene mapped from BAC end sequences (Leroy et al. 2005). In LG F, QTLs for bean size and the rate of pea berries were located near a *SUS1* sucrose synthase gene. In LG I, a gene implied in caffeine metabolism, one N-methyltransferase, was found close to QTLs for caffeine (2005), monomer CGAs and sucrose (2005), and acidity and bitterness (2006).

**Fig. 1** Genetic linkage map of *C. canephora* based on SSR markers and genes, constructed with 238 markers and 184 plants with a minimum LOD score of 4.4. Distortion of markers in linkage groups is indicated by *stars* corresponding to the significance of the $\chi^2$ test (*double asterisk*, $p$=0.05 to *septuple asterisk*, $p$=0.0001). Candidate genes are indicated in *bold*

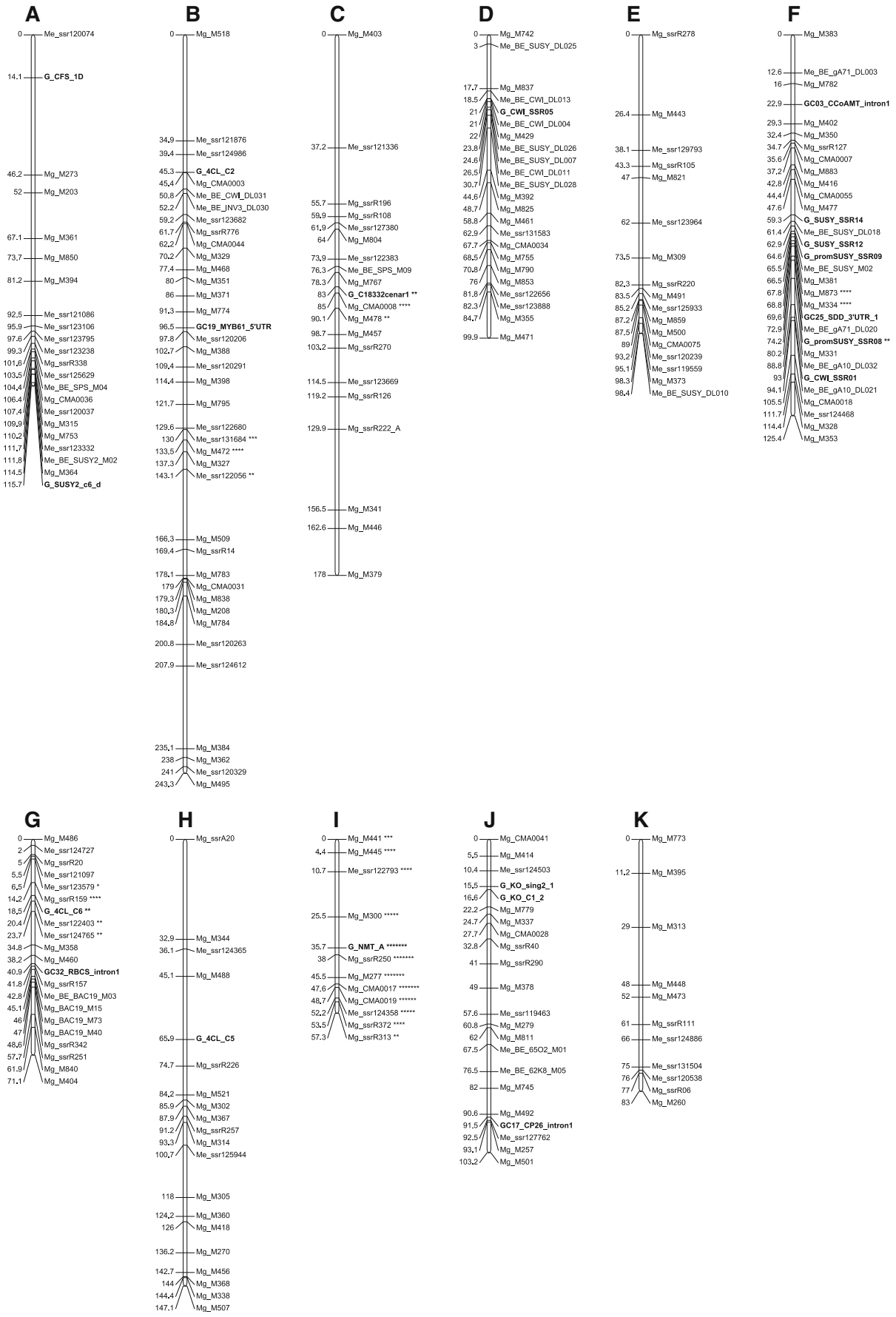
## Discussion

Our population of *C. canephora* possesses a large amount of phenotypic variability for yield and for traits related to quality. This outcome was expected for yield. For sucrose and caffeine content, it presents a higher variability than that observed in *C. canephora* accessions by Ky et al. (2001a). As it is known that sucrose content has additive transmission, selection for this trait will be possible. For biochemical traits, a good correlation was observed between the years, which is consistent with the results obtained by Ky et al. (2000a).

CGA biosynthesis in coffee seeds has been analyzed by Joët et al. (2010). These authors defined "control boxes" for CGA biosynthesis and showed relationships between mono- and di-CQA chlorogenic acids and the influence of abiotic factors such as low temperatures. Linear relationships between di-CQA have also been established (Bertrand et al. 2003b). Our results confirm linear relationships between CQA and di-CQA, all di-CQA being derived from 5-CQA (Ky et al. 1999).

### The genetic map

As expected for *C. canephora*, a genetic map of 11 linkage groups was produced using 238 markers, with an average of 5.5 cM between markers. As previously reported for intra- and interspecific maps in coffee (Paillard et al. 1996; Lashermes et al. 2001; Ky et al. 2000b), segregation distortion was observed for LG I. The distorted markers have not been discarded during map construction as they have for other species (Venkateswarlu et al. 2006) because


A
0 Me_ssr120074
14.1 G_CFS_1D
46.2 Mg_M273
52 Mg_M203
67.1 Mg_M361
73.7 Mg_M850
81.2 Mg_M394
92.5 Me_ssr121086
95.9 Me_ssr123106
97.6 Me_ssr123795
99.3 Me_ssr123238
101.6 Mg_ssrR338
103.5 Me_ssr125629
104.4 Me_BE_SPS_M04
106.4 Mg_CMA0036
107.4 Me_ssr120037
109.9 Mg_M315
110.2 Mg_M753
111.7 Me_ssr123332
111.8 Me_BE_SUSY2_M02
114.5 Mg_M364
115.7 G_SUSY2_c6_d
B
0 Mg_M518
34.9 Me_ssr121876
39.4 Me_ssr124986
45.3 G_4CL_C2
45.4 Mg_CMA0003
50.8 Me_BE_CWI_DL031
52.2 Me_BE_INV3_DL030
59.2 Me_ssr123682
61.7 Mg_ssrR776
62.2 Mg_CMA0044
70.2 Mg_M329
77.4 Mg_M468
80 Mg_M351
86 Mg_M371
91.3 Mg_M774
96.5 GC19_MYB61_5'UTR
97.8 Me_ssr120206
102.7 Mg_M388
109.4 Me_ssr120291
114.4 Mg_M398
121.7 Mg_M795
129.6 Me_ssr122680
130 Me_ssr131684 ***
133.5 Mg_M472 ****
137.3 Mg_M327
143.1 Me_ssr122056 **
166.3 Mg_M509
169.4 Mg_ssrR14
178.1 Mg_M783
179 Mg_CMA0031
179.3 Mg_M838
180.3 Mg_M208
184.8 Mg_M784
200.8 Me_ssr120263
207.9 Me_ssr124612
235.1 Mg_M384
238 Mg_M362
241 Me_ssr120329
243.3 Mg_M495
C
0 Mg_M403
37.2 Me_ssr121336
55.7 Mg_ssrR196
59.9 Mg_ssrR108
61.9 Me_ssr127380
64 Mg_M804
73.9 Me_ssr122383
76.3 Me_BE_SPS_M09
78.3 Mg_M767
83 G_C18332cenar1 **
85 Mg_CMA0008 ****
90.1 Mg_M478 **
98.7 Mg_M457
103.2 Mg_ssrR270
114.5 Me_ssr123669
119.2 Mg_ssrR126
129.9 Mg_ssrR222_A
156.5 Mg_M341
162.6 Mg_M446
178 Mg_M379
D
0 Mg_M742
3 Me_BE_SUSY_DL025
17.7 Mg_M837
18.5 Me_BE_CWI_DL013
21 G_CWI_SSR05
21 Me_BE_CWI_DL004
22 Mg_M429
23.8 Me_BE_SUSY_DL026
24.6 Me_BE_SUSY_DL007
26.5 Me_BE_CWI_DL011
30.7 Me_BE_SUSY_DL028
44.6 Mg_M392
48.7 Mg_M825
58.8 Mg_M461
62.9 Me_ssr131583
67.7 Mg_CMA0034
68.5 Mg_M755
70.8 Mg_M790
76 Mg_M853
81.8 Me_ssr122656
82.3 Me_ssr123888
84.7 Mg_M355
99.9 Mg_M471
E
0 Mg_ssrR278
26.4 Mg_M443
38.1 Me_ssr129793
43.3 Mg_ssrR105
47 Mg_M821
62 Me_ssr123964
73.5 Mg_M309
82.3 Mg_ssrR220
83.5 Mg_M491
85.2 Me_ssr125933
87.2 Mg_M859
87.5 Mg_M500
89 Mg_CMA0075
93.2 Me_ssr120239
95.1 Me_ssr119559
98.3 Mg_M373
98.4 Me_BE_SUSY_DL010
F
0 Mg_M383
12.6 Me_BE_gA71_DL003
16 Mg_M782
22.9 GC03_CCoAMT_intron1
29.3 Mg_M402
32.4 Mg_M350
34.7 Mg_ssrR127
35.6 Mg_CMA0007
37.2 Mg_M883
42.8 Mg_M416
44.4 Mg_CMA0055
47.6 Mg_M477
59.3 G_SUSY_SSR14
61.4 Me_BE_SUSY_DL018
62.9 G_SUSY_SSR12
64.6 G_promSUSY_SSR09
65.5 Me_BE_SUSY_M02
66.5 Mg_M381
67.8 Mg_M873 ****
68.8 Mg_M334 ****
69.6 GC25_SDD_3'UTR_1
72.9 Me_BE_gA71_DL020
74.2 G_promSUSY_SSR08 **
80.2 Mg_M331
88.8 Me_BE_gA10_DL032
93 G_CWI_SSR01
94.1 Me_BE_gA10_DL021
105.5 Mg_CMA0018
111.7 Me_ssr124468
114.4 Mg_M328
125.4 Mg_M353
G
0 Mg_M486
2 Me_ssr124727
5 Mg_ssrR20
5.5 Me_ssr121097
6.5 Me_ssr123579 *
14.2 Mg_ssrR159 ****
18.5 G_4CL_C6 **
20.4 Me_ssr122403 **
23.7 Me_ssr124765 **
34.8 Mg_M358
38.2 Mg_M460
40.9 GC32_RBCS_intron1
41.8 Mg_ssrR157
42.8 Me_BE_BAC19_M03
45.1 Mg_BAC19_M15
46 Mg_BAC19_M73
47 Mg_BAC19_M40
48.6 Mg_ssrR342
57.7 Mg_ssrR251
61.9 Mg_M840
71.1 Mg_M404
H
0 Mg_ssrA20
32.9 Mg_M344
36.1 Me_ssr124365
45.1 Mg_M488
65.9 G_4CL_C5
74.7 Mg_ssrR226
84.2 Mg_M521
85.9 Mg_M302
87.9 Mg_M367
91.2 Mg_ssrR257
93.3 Mg_M314
100.7 Me_ssr125944
118 Mg_M305
124.2 Mg_M360
126 Mg_M418
136.2 Mg_M270
142.7 Mg_M456
144 Mg_M368
144.4 Mg_M338
147.1 Mg_M507
I
0 Mg_M441 ***
4.4 Mg_M445 ****
10.7 Me_ssr122793 ****
25.5 Mg_M300 *****
35.7 G_NMT_A *******
38 Mg_ssrR250 ********
45.5 Mg_M277 *******
47.6 Mg_CMA0017 *******
48.7 Mg_CMA0019 ******
52.2 Me_ssr124358 *****
53.5 Mg_ssrR372 ****
57.3 Mg_ssrR313 **
J
0 Mg_CMA0041
5.5 Mg_M414
10.4 Me_ssr124503
15.5 G_KO_sing2_1
16.6 G_KO_C1_2
22.2 Mg_M779
24.7 Mg_M337
27.7 Mg_CMA0028
32.8 Mg_ssrR40
41 Mg_ssrR290
49 Mg_M378
57.6 Me_ssr119463
60.8 Mg_M279
62 Mg_M811
67.5 Me_BE_65O2_M01
76.5 Me_BE_62K8_M05
82 Mg_M745
90.6 Mg_M492
91.5 GC17_CP26_intron1
92.5 Me_ssr127762
93.1 Mg_M257
103.2 Mg_M501
K
0 Mg_M773
11.2 Mg_M395
29 Mg_M313
48 Mg_M448
52 Mg_M473
61 Mg_ssrR111
66 Me_ssr124886
75 Me_ssr131504
76 Me_ssr120538
77 Mg_ssrR06
83 Mg_M260

it can be considered that these distortions have a biological basis. As has been noted for other plants (Billote et al. 2005; Lu et al. 2002), the presence of lethal genes or reproduction regulating genes could explain these distortions (Zamir and Tadmor 1986). In *C. canephora*, the self-incompatibility gene S should be located in the LG designated LG 9 by Lashermes et al. (1996), and this LG has been then confirmed to be highly distorted (Lashermes et al. 2001), suggesting a link between segregation distortion and the S locus. Our LG I corresponds to LG 9 of Lashermes. In our study, if we consider that this simple gametophytic factor is responsible of the distortion observed, estimates of the recombination frequency should not be affected (Lu et al. 2002). In the other LGs, only small clusters of markers presented segregation distortion (Fig. 1).

## QTL identification and gene discovery

We were able to identify QTLs for most traits with large effects and consistent detection over the years, especially for yield, technological traits, and chlorogenic acid content, thus indicating ontogenic stability. These QTLs, also referred to as stable QTLs, are important because they can be excellent candidates for MAS work (Kenis et al. 2008). For other traits, such as organoleptic traits and other biochemical compounds related to quality, the consistency over the years is less obvious. LG I is highly distorted, and it could, thus, affect the quality of QTL identification. Recent reports (Xu 2008; Zhang et al. 2010) have noted that distortion can decrease the possibility of QTL detection, but it should not increase the rate of false positives. Zhang et al. (2010) also noted the importance of the population size for avoiding modifications of QTLs due to distortion. Our coffee population of more than 240 trees can be considered as a large population, thus limiting the decrease in the power of QTL detection. However, Xu (2008) suggested that the power of QTL mapping could be artificially increased when a map is sparse, as in our study. In conclusion, the QTLs identified in our distorted LG I should not prevent further investigation on these regions of the genome.

In this study, several zones of interest were identified. One zone in LG K (ranging from 45 to 60 cM) is of great interest for cumulative yield, and the QTL on this LG co-localized with trigonelline content in the 2003 sample. As previously reported for the organoleptic quality of fresh tomato fruits (Causse et al. 2000; Saliba et al. 2001), we identified large clusters of QTLs related to quality in LGs I and B. The main QTL zone for quality in LG I is associated with correlated traits like acidity and bitterness, caffeine and chlorogenic acids A key gene involved in caffeine biosynthesis is co-localized with these QTLs. It was demonstrated here for the first time that the genetic variability in chemical compounds, such caffeine and chlorogenic acids, is related to the genetic variability of beverage quality (acidity and bitterness).The establishment of this clear relationship can be considered as the main result of this work. We also observed co-localization between bean size and quality traits in LG B. Finally, we identified a co-localization between a QTL for bean size, a *SUS1* gene and a *CWI* gene coding for the cell wall invertase in LG D. This result is of extreme importance, as several invertase- and sucrose synthase-encoding genes have been mapped along with a QTL related to fruit size in tomatoes (Fridman et al. 2000; Causse et al. 2004). These genes are involved in the control of tomato fruit size (Klann et al. 1996; D´Aoust et al. 1999). In shaded coffee beans, Geromel et al. (2008) also observed higher invertase activity during the latest stages of perisperm development that could be related to the larger size of beans in comparison to those grown under full sun conditions.

Uncovering co-localization of QTLs with mapped candidate genes from various biological pathways and subsequent use of these genes in association mapping will present possibilities for determining the causes of variation of quality in coffee (Henery et al. 2007). We can explain the co-localization between QTLs for organoleptic traits and genes involved in caffeine or CGA biosynthesis based on the fact that both CGA and caffeine are involved in conferring bitterness to the beverage.

Our study is the first to report the identification of QTLs involved in quality traits and yield in coffee. We identified major QTLs for yield and quality-related traits. We confirmed the relationships between some traits and their involvement in determining quality. Investigation of chemical and biochemical traits related to quality presents an alternative approach to improve quality, as few QTLs were discovered for direct beverage quality traits, except for acidity and bitterness. These types of indirect traits, such as caffeine, chlorogenic acid, and sugar content, are important for breeders because they are relatively easy to quantify and because their variability is high, and they presented significant correlations with beverage quality traits. We also mapped genes involved in the biosynthesis of crucial compounds, including caffeine, sugars and chlorogenic acids; they co-localize with QTLs. Our results constitute a first step toward the detection of favorable alleles related to beverage quality in coffee. Finally, we identified several regions of interest that will be the target for future association mapping studies. The regions of the genome connected to yield determination and to quality are different,

**Fig. 2** Main QTL localizations in the linkage groups for the agronomical, technological, biochemical, and organoleptic traits. For each trait, the QTLs are represented by *boxes* for a confidence interval of $LOD_{max} \pm 1$. QTLs were confident at 10% (*normal*), 5% (*italics*), or 1% (*bold*) at the genome-wide level. The names of the QTLs are presented in Table 4

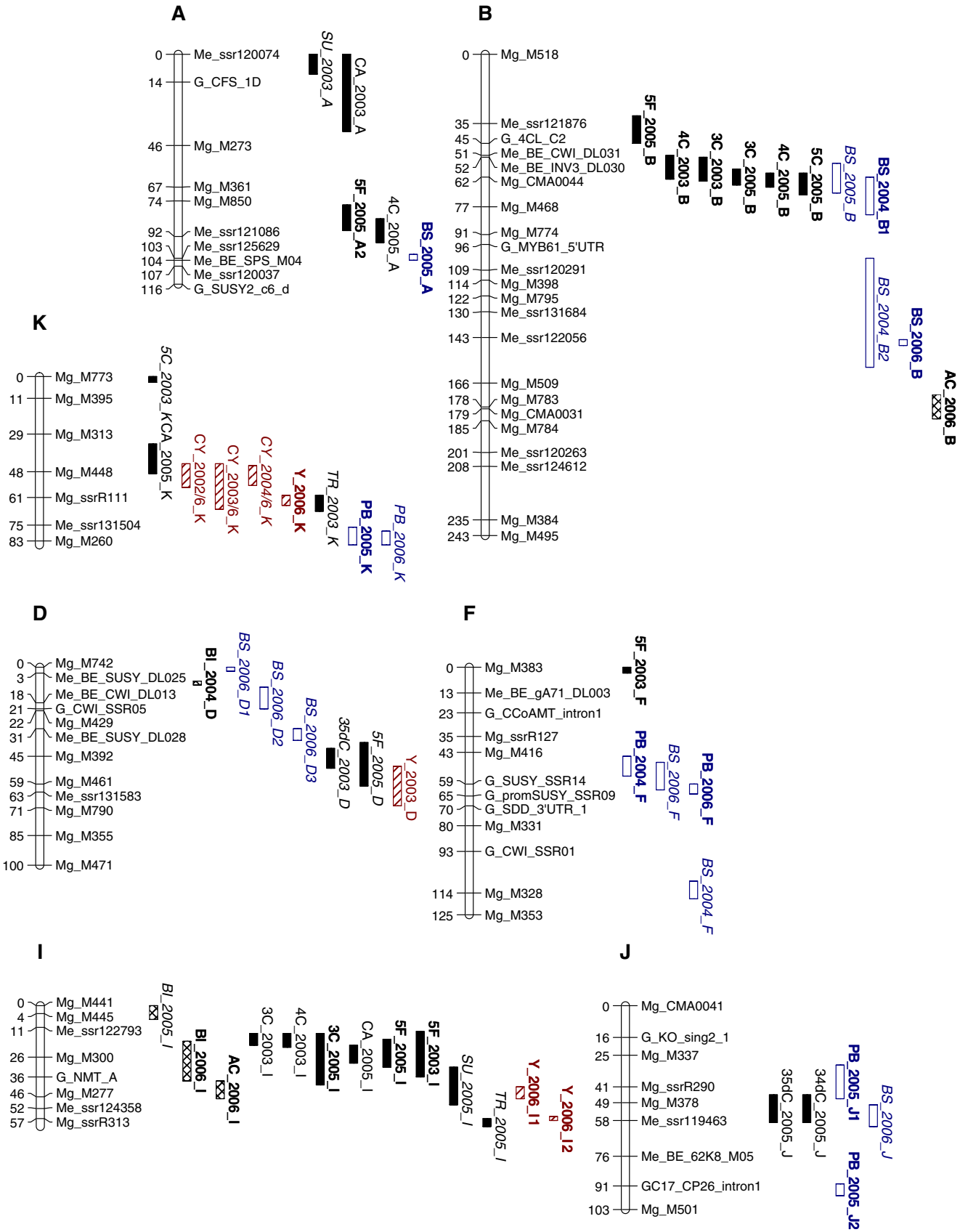

A
0 Me_ssr120074
14 G_CFS_1D
46 Mg_M273
67 Mg_M361
74 Mg_M850
92 Me_ssr121086
103 Me_ssr125629
104 Me_BE_SPS_M04
107 Me_ssr120037
116 G_SUSY2_c6_d
SU_2003_A
CA_2003_A
5F_2005_A2
4C_2005_A
BS_2005_A
B
0 Mg_M518
35 Me_ssr121876
45 G_4CL_C2
51 Me_BE_CWI_DL031
52 Me_BE_INV3_DL030
62 Mg_CMA0044
77 Mg_M468
91 Mg_M774
96 G_MYB61_5'UTR
109 Me_ssr120291
114 Mg_M398
122 Mg_M795
130 Me_ssr131684
143 Me_ssr122056
166 Mg_M509
178 Mg_M783
179 Mg_CMA0031
185 Mg_M784
201 Me_ssr120263
208 Me_ssr124612
235 Mg_M384
243 Mg_M495
5F_2005_B
4C_2003_B
3C_2003_B
3C_2005_B
4C_2005_B
5C_2005_B
BS_2005_B
BS_2004_B1
BS_2004_B2
BS_2006_B
AC_2006_B
K
0 Mg_M773
11 Mg_M395
29 Mg_M313
48 Mg_M448
61 Mg_ssrR111
75 Me_ssr131504
83 Mg_M260
5C_2003_K
CA_2005_K
CY_2002/6_K
CY_2003/6_K
CY_2004/6_K
Y_2006_K
TR_2003_K
PB_2005_K
PB_2006_K
D
0 Mg_M742
3 Me_BE_SUSY_DL025
18 Me_BE_CWI_DL013
21 G_CWI_SSR05
22 Mg_M429
31 Me_BE_SUSY_DL028
45 Mg_M392
59 Mg_M461
63 Me_ssr131583
71 Mg_M790
85 Mg_M355
100 Mg_M471
BI_2004_D
BS_2006_D1
BS_2006_D2
BS_2006_D3
35dC_2003_D
5F_2005_D
Y_2003_D
F
0 Mg_M383
13 Me_BE_gA71_DL003
23 G_CCoAMT_intron1
35 Mg_ssrR127
43 Mg_M416
59 G_SUSY_SSR14
65 G_promSUSY_SSR09
70 G_SDD_3'UTR_1
80 Mg_M331
93 G_CWI_SSR01
114 Mg_M328
125 Mg_M353
5F_2003_F
PB_2004_F
BS_2006_F
PB_2006_F
BS_2004_F
I
0 Mg_M441
4 Mg_M445
11 Me_ssr122793
26 Mg_M300
36 G_NMT_A
46 Mg_M277
52 Me_ssr124358
57 Mg_ssrR313
BI_2005_I
BI_2006_I
AC_2006_I
3C_2003_I
4C_2003_I
3C_2005_I
CA_2005_I
5F_2005_I
5F_2003_I
SU_2005_I
TR_2005_I
Y_2006_I1
Y_2006_I2
J
0 Mg_CMA0041
16 G_KO_sing2_1
25 Mg_M337
41 Mg_ssrR290
49 Mg_M378
58 Me_ssr119463
76 Me_BE_62K8_M05
91 GC17_CP26_intron1
103 Mg_M501
35dC_2005_J
34dC_2005_J
PB_2005_J1
PB_2005_J2
BS_2006_J

**Table 4** List of QTLs identified by MQM for agronomic, technological, biochemical, and organoleptic traits

| Traits | LG | QTL name | Cofactors[a] | $LOD_{max}$ | $LOD_{max}$ position[b] | $R^2$ [c] | Af | Am | D |
|---|---|---|---|---|---|---|---|---|---|
| Cumulative yield, 2002–2006 | K | CY_2002/6_K | – | 3.96 | 49.2 | 17.4 | 0.06 | 0.06 | −0.09 |
| Cumulative yield, 2003–2006 | K | CY_2003/6_K | – | 4.07 | 50.2 | 16.7 | 0.06 | 0.06 | −0.09 |
| Cumulative yield, 2004–2006 | K | CY_2004/6_K | – | 3.98 | 50.2 | 16.2 | 0.07 | 0.07 | −0.09 |
| Yield, 2003 | D | Y_2003_D | – | 4.07 | 66.1 | 8.1 | 0.26 | −0.02 | −0.18 |
| Yield, 2006 | I | **Y_2006_I1** | G_NMT_A, Me_ssr122793, Mg_M445, Mg_M313 | 6.33 | 43.8 | 34.0 | −0.34 | −0.41 | −0.29 |
| | I | **Y_2006_I2** | G_NMT_A, Me_ssr122793, Mg_M445, Mg_M313 | 9.46 | 57.9 | 42.4 | 0.03 | 0.57 | −0.03 |
| | K | **Y_2006_K** | G_NMT_A, Me_ssr122793, Mg_M445, Mg_M313 | 9.93 | 62.4 | 56.8 | 0.24 | 0.25 | −0.25 |
| Bean size, 2004 | B | **BS_2004_B1** | Me_ssr120037, Mg_M328 | 4.64 | 71.9 | 10.5 | 0.38 | −0.30 | −0.20 |
| | B | *BS_2004_B2* | Me_ssr120037, Mg_M328 | 4.48 | 130.8 | 8.1 | 0.42 | 0.07 | 0.18 |
| | F | *BS_2004_F* | Me_ssr120037, Mg_M328 | 4.26 | 113.2 | 7.9 | 0.24 | 0.42 | 0.01 |
| Bean size, 2005 | A | **BS_2005_A** | Me_ssr125629, Me_BE_SPS_M04, Mg_M328, Mg_M521[d] | 16.44 | 103.7 | 35.3 | −0.63 | 0.00 | 0.01 |
| | B | *BS_2005_B* | Me_ssr125629, Me_BE_SPS_M04, Mg_M328, Mg_M521[d] | 4.24 | 64.3 | 6.4 | −0.05 | −0.25 | −0.49 |
| Bean size, 2006 | B | **BS_2006_B** | Mg_M783, Mg_CMA0031, Mg_BAC19_M40[d], Me_ssr119463 | 10.30 | 145.8 | 25.3 | 0.50 | −0.01 | −0.15 |
| | D | *BS_2006_D1* | Mg_M783, Mg_CMA0031, Mg_BAC19_M40[d], Me_ssr119463 | 4.27 | 0 | 6.1 | −0.12 | 0.00 | 0.45 |
| | D | *BS_2006_D2* | Mg_M783, Mg_CMA0031, Mg_BAC19_M40[d], Me_ssr119463 | 4.92 | 15.3 | 10.9 | −0.09 | 0.00 | 0.60 |
| | D | *BS_2006_D3* | Mg_M783,Mg_CMA0031, Mg_BAC19_M40[d], Me_ssr119463 | 5.12 | 33.3 | 14.1 | 0.04 | 0.15 | 0.66 |
| | F | *BS_2006_F* | Mg_M783,Mg_CMA0031, Mg_BAC19_M40[d], Me_ssr119463 | 4.49 | 52.8 | 11.0 | −0.33 | 0.36 | −0.37 |
| | J | *BS_2006_J* | Mg_M783, Mg_CMA0031, Mg_BAC19_M40[d], Me_ssr119463 | 5.00 | 57 | 7.2 | −0.46 | −0.19 | −0.01 |
| Pea berry, 2004 | F | **PB_2004_F** | – | 6.46 | 49.8 | 19.4 | −6.05 | 0.91 | −0.18 |
| Pea berry, 2005 | J | **PB_2005_J1** | – | 3.89 | 39.6 | 11.5 | 3.45 | 0.12 | 1.50 |
| | J | **PB_2005_J2** | – | 3.94 | 92.8 | 8.1 | 3.03 | 1.23 | 0.59 |
| | K | **PB_2005_K**[d] | – | 4.49 | 83.3 | 16.6 | −3.71 | 1.48 | −3.12 |
| Pea berry, 2006 | F | **PB_2006_F** | Mg_ssrR127, G_SSR09_promSUSY | 5.21 | 61.5 | 10.5 | −4.33 | 0.27 | 1.15 |
| | K | *PB_2006_K*[d] | Mg_ssrR127, G_SSR09_promSUSY | 4.76 | 85.5 | 13.5 | −3.83 | 2.00 | −2.60 |
| 34dCQA, 2005 | J | 34dC_2005_J | – | 4.15 | 53 | 13.1 | 0.02 | −0.08 | −0.40 |
| 35dCQA, 2003 | D | *35dC_2003_D* | – | 2.52 | 45.2 | 10.7 | -0.70 | −0.20 | 0.52 |
| 35dCQA, 2005 | J | 35dC_2005_J | – | 4.10 | 52 | 13.0 | 0.00 | −0.13 | −0.43 |
| 3CQA, 2003 | B | **3C_2003_B** | – | 6.58 | 57.3 | 26.7 | 1.02 | −0.04 | 0.06 |
| | I | 3C_2003_I | – | 4.02 | 15.6 | 21.1 | 0.83 | −0.46 | -0.49 |
| 3CQA, 2005 | B | **3C_2005_B** | Mg_M774, Mg_M398 | 9.18 | 63.3 | 16.0 | 0.97 | −0.33 | -0.06 |

**Table 4** (continued)

| Traits | LG | QTL name | Cofactors[a] | $LOD_{max}$ | $LOD_{max}$ position[b] | $R^2$ [c] | Af | Am | D |
|---|---|---|---|---|---|---|---|---|---|
| | I | **3C_2005_I** | Mg_M774, Mg_M398 | 5.31 | 20.4 | 6.5 | 0.47 | −0.26 | -0.05 |
| 4CQA, 2003 | B | **4C_2003_B** | – | 6.10 | 53.1 | 28.8 | 1.18 | −0.10 | 0.32 |
| | I | 4C_2003_I | – | 4.04 | 15.6 | 21.6 | 1.00 | −0.55 | −0.50 |
| 4CQA, 2005 | A | 4C_2005_A | Mg_M774 | 3.96 | 86.5 | 8.3 | −0.64 | 0.33 | −0.17 |
| | B | **4C_2005_B** | Mg_M774 | 19.62 | 64.3 | 40.6 | 1.41 | −0.25 | −0.05 |
| 5CQA, 2003 | K | *5C_2003_K*[d] | – | 4.31 | 0 | 18.5 | 1.89 | 0.90 | 1.45 |
| 5CQA, 2005 | B | 5C_2005_B | – | 4.26 | 65.3 | 10.6 | −1.83 | 0.86 | -0.43 |
| 5FQA, 2003 | F | **5F_2003_F** | – | 5.21 | 0 | 35.3 | 0.92 | 0.59 | 1.36 |
| | I | **5F_2003_I** | – | 5.12 | 27.8 | 23.4 | −1.47 | −0.43 | 0.04 |
| 5FQA, 2005 | A | **5F_2005_A2** | – | 5.47 | 82.5 | 13.7 | 0.88 | −0.65 | −0.35 |
| | B | **5F_2005_B** | – | 5.89 | 41.1 | 15.2 | −1.10 | 0.46 | −0.11 |
| | D | *5F_2005_D* | – | 4.38 | 48.2 | 10.8 | −0.66 | 0.53 | −0.49 |
| | I | **5F_2005_I** | – | 7.08 | 24.8 | 16.7 | −1.26 | 0.12 | −0.27 |
| Caffeine, 2003 | A | CA_2003_A | Me_BE_gA71_DL003, Mg_300 | 3.78 | 22.9 | 16.2 | −0.18 | −0.09 | 0.11 |
| Caffeine, 2005 | C | **CA_2005_C**[d] | Me_ssr121336[d], G_NMT_A, Mg_M448 | 5.08 | 2 | 13.2 | 0.15 | 0.11 | 0.01 |
| | I | CA_2005_I | Me_ssr121336[d], G_NMT_A, Mg_M448 | 4.05 | 22.8 | 7.4 | −0.14 | 0.07 | 0.08 |
| | K | CA_2005_K | Me_ssr121336[d], G_NMT_A, Mg_M448 | 3.99 | 38.2 | 7.2 | −0.10 | 0.07 | −0.07 |
| Sucrose, 2003 | A | *SU_2003_A* | Me_ssr120074 | 4.42 | 0 | 28.8 | −0.40 | −0.58 | 0.15 |
| Sucrose, 2005 | I | *SU_2005_I* | – | 4.39 | 40.8 | 12.9 | 0.40 | 0.10 | 0.02 |
| Trigonelline, 2003 | K | *TR_2003_K* | Mg_M313 | 11.73 | 64.4 | 81.2 | 0.16 | 0.15 | 0.11 |
| Trigonelline, 2005 | I | *TR_2005_I* | Mg_M441 | 4.07 | 60.9 | 41.8 | −0.04 | 0.04 | −0.08 |
| Acidity, 2006 | B | **AC_2006_B** | Mg_M300, Me_ssr122793 | 11.91 | 173.2 | 29.8 | 0.00 | −0.05 | 0.00 |
| | I | **AC_2006_I** | Mg_M300, Me_ssr122793 | 8.22 | 42.8 | 54.8 | 0.05 | 0.04 | 0.05 |
| Bitterness, 2004 | D | **BI_2004_D** | Mg_M429 | 5.07 | 7.7 | 31.7 | 0.36 | −0.16 | −0.03 |
| Bitterness, 2005 | I | *BI_2005_I* | – | 4.37 | 0 | 16.8 | −0.23 | −0.06 | 0.00 |
| Bitterness, 2006 | I | **BI_2006_I** | – | 5.24 | 28.8 | 15.3 | −0.17 | −0.02 | −0.02 |
| Global, 2005 | H | GL_2005_H[d] | – | 4.03 | 73.3 | 18.6 | −0.03 | 0.09 | −0.14 |

The trait, the linkage group where the QTL is located, the QTL name, the markers used as eventual cofactors in the MQM analyzed, the maximum LOD and its position, the percentage of phenotypic variation explained for the trait ($R^2$), allelic female and male additivity and dominance are indicated. All of these QTLs present a confidence under 1% at the LG level and 10% at the genome-wide (GW) level. QTLs that present a confidence under 5% GW at the genome level are indicated in italics, and for less than 1% confidence, they are shown in bold

*MQM* multiple QTL mapping, *TR* trigonelline, LG linkage group, *Af* female additive effects, *Am* male additive effects, *D* dominance effects

[a] Cofactors were selected using the "Automatic Cofactor Selection" procedure in MapQTL 4.0 software

[b] Position of the LOD maximum expressed in cM on the reference map

[c] Percentage of phenotypic variation explained by the QTL

[d] Not displayed in Fig. 2

[e] As presented in "Materials and methods"

indicating that simultaneous selection can be conducted on both types of traits, as previously suggested by Montagnon et al. (1998). The ongoing work on coffee genome sequencing will allow for more accurately defining the location of genes of interest in relation to the QTLs identified in this study.

## Consequences for breeding

The perspectives arising from the results of our study are important in relation with the improvement of the quality of *C. canephora*. This study was based on a "backcross"

progeny between the Congolese and the Guinean groups on which the selection program in Côte d'Ivoire is established. These QTL studies are specific to the pedigree, site, and ontogenic stage of the individuals investigated. Thus, we will have to validate the stability of the QTLs obtained here in other locations and in different progenies, particularly in the offspring of crosses generated between the diversity groups used in the selection scheme. For the offspring of crosses generated within the diversity groups, we will have to perform complementary QTL studies. Association mapping studies within genetic groups or on large coffee collections will also allow precisely determining the zones of the genome and the genes implicated in the establishment of quality. The results presented here and future studies to be implemented in other populations will contribute developing marker-assisted selection for quality improvement, considering the favorable alleles of the markers involved in the QTLs for quality. A marker-assisted selection could be applied directly to our progeny for the selection of varieties presenting good yield and quality.

Finally, our work is of interest for assisting breeders attempting to control the introgression of resistance genes from *C. canephora* to *C. arabica* without lowering quality (Bertrand et al. 2003a).

**Acknowledgments** This work was supported by EU grant ICA4-CT-2001-10068. The University of Trieste (Italy) kindly provided 16 SSR markers.

## References

Akaffou DS, Ky CL, Barre P, Hamon S, Louarn J, Noirot M (2003) Identification and mapping of a major gene (Ft1) involved in fructification time in the interspecific cross *Coffea pseudozanguebariae* x *C. liberica* var. Dewevrei: impact on caffeine content and seed weight. Theor Appl Genet 106:1486–1490

Alcazar A, Jurad JM, Martin MJ, Pablos F, Gonzales AG (2004) Enzymatic-spectrophotometric determination of sucrose in coffee beans. Talanta 67:760–766

Anthony F, Clifford MN, Noirot M (1993) Biochemical diversity in the genus *Coffea* L.: chlorogenic acids, caffeine and mozambioside contents. Genet Res Crop Evol 40:61–70

Anthony F, Bertrand B, Quiros O, Wilches A, Lashermes P, Berthaud J, Charrier A (2001) Genetic diversity of wild coffee (*Coffea arabica* L.) using molecular markers. Euphytica 118:53–65

Barre P, Akaffou S, Louarn J, Charrier A, Hamon S, Noirot M (1998) Inheritance of caffeine and heteroside contents in an interspecific cross between a cultivated coffee species *Coffea liberica* var dewevrei and a wild species caffeine-free *C. pseudozanguebariae*. Theor Appl Genet 96:306–311

Baruah A, Naik V, Hendre PS, Rajkumar R, Rajendrakumar P, Aggarwal K (2003) Isolation and characterization of nine microsatellite markers from *Coffea arabica* L. showing wide cross-species amplifications. Mol Ecol Notes 3:647–650

Berthaud J (1986) Les ressources génétiques pour l'amélioration des caféiers africains diploïdes. Evaluation de la richesse génétique des populations sylvestres et de ses mécanismes organisateurs. Conséquences pour l'application. ORSTOM, Paris

Bertrand B, Guyot B, Anthony F, Lashermes P (2003a) Impact of the *Coffea canephora* gene introgression on beverage quality of *C.* Arabica. Theor Appl Genet 107:387–394

Bertrand C, Noirot M, Doulbeau S, de Kochko A, Hamon S, Campa C (2003b) Chlorogenic acid content swap during fruit maturation in *Coffea pseudozanguebariae*. Quantitative comparison with leaves. Plant Sci 165:1355–1361

Billote N, Marseillac N, Risterucci AM et al (2005) Microsatellite-based high density linkage map in oil palm (*Elaeis guineensis* Jacq.). Theor Appl Genet 110:754–765

Bouchet S, Marraccini P, Jourdan I, Leroy T, Vieira LGE, Ferreira LP, Musoli P, Pot D (2005) Nucleotide diversity and molecular evolution of five genes involved in the sucrose biosynthesis pathway of *Coffea canephora*. In: Proceedings of the 4th Plant Genomics European Meeting, Amsterdam, 20–23 Sept 2005

Campa C, Noirot M, Bourgeois M, Pervent M, Ky CL, Chrestin H, Hamon S, de Kochko A (2003) Genetic mapping of a caffeoyl-coenzyme A 3-0-methltransferase gene in coffee trees. Impact on chlorogenic acid content. Theor Appl Genet 107:751–756

Causse M, Saliba V, Lesschaeve I, Buret M (2000) Genetic analysis of organoleptic quality in fresh-market tomato. 2. Mapping QTLs for sensory attributes. Theor Appl Genet 102:257–272

Causse M, Duffe P, Gomez MC, Buret M, Damidaux R, Zamir D, Gur A, Chevalier C, Lemaire-Chamley M, Rothan C (2004) A genetic map of candidate genes and QTLs involved in tomato fruit size and composition. J Exp Bot 55:1671–1685

Charrier A, Berthaud J (1988) Principles and methods in coffee plant breeding *Coffea canephora* Pierre. In: Clarke RJ, Macrae R (eds) Coffee, vol 4, Agronomy. Elsevier Applied Science, London, pp 167–197

Combes MC, Andrzejewski S, Anthony F, Bertrand B, Rovelli P, Grasiozi G, Lashermes P (2000) Characterization of microsatellite loci in *Coffea arabica* and related coffee species. Mol Ecol Notes 9:1171–1193

Coulibaly I, Revol B, Noirot M, Poncet V, Lorieux M, Carasco-Lacombe C, Minier J, Hamon P (2003) AFLP and SSR polymorphism in a *Coffea* interspecific backcross progeny ((*C. heterocalyx* x *C. canephora*) x *C. canephora*). Theor Appl Genet 107:1148–1155

D'Aoust MA, Yelle S, Nguyen-Quoc B (1999) Antisense inhibition of tomato fruit sucrose synthase decreases fruit setting and the sucrose unloading capacity of young fruit. Plant Cell 11:2407–2418

Dussert D, Lashermes P, Anthony F, Montagnon C, Trouslot P, Combes MC, Berthaud J, Noirot M, Hamon S (1999) Le caféier, *Coffea canephora*. In: Hamon P, Seguin M, Perrier X, Glaszmann JC (eds) Diversité génétique des plantes tropicales cultivées. CIRAD, Montpellier, pp 175–194

Freire LP, Vieira NG, Vinecky F, Alvas GSC, Leroy T, Pot D, Elbelt S, Marques T, Rodrigues GC, Marraccini P, Andrade AC (2010). Expression analysis and nucleic polymorphism of candidate genes for drought tolerance in coffee (in press)

Fridman E, Pleban T, Zamir D (2000) A recombination hotspot delimits a wild-species quantitative trait locus for tomato sugar content to 484 bp within an invertase gene. Proc Natl Acad Sci USA 97:4718–4723

Geromel C, Ferreira LP, Guerreiro SMC, Cavalari AA, Pot D, Pereira LFP, Leroy T, Vieira LGE, Mazzafera P, Marraccini P (2006) Biochemical and genomic analysis of sucrose metabolism during coffee (*Coffea arabica*) fruit development. J Exp Bot 57:3243–3258

Geromel C, Ferreira LP, Davrieux F, Guyot B, Ribeyre F, dos Santos Scholz MB, Pereira LFP, Vaast P, Pot D, Leroy T, Androcioli Filho A, Vieira LGE, Mazzafera P, Marraccini P (2008) Effects of shade on the development and sugar metabolism of coffee (*Coffea arabica* L.) fruits. Plant Physiol Biochem 46:569–579

Gomez C, Dussert S, Hamon P, Hamon S, de Kochko A, Poncet V (2009) Current genetic differentiation of *Coffea canephora* Pierre ex A. Froehn in the Guineo-Congolian African zone: cumulative impact of ancient climatic changes and recent human activities. BMC Evol Biol 9:167

Grattapaglia D, Sederoff R (1994) Genetic maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-test cross mapping strategy and RAPD markers. Genetics 137:1121–1137

Henery ML, Moran GF, Wallis IR, Foley WJ (2007) Identification of quantitative trait loci influencing foliar concentrations of terpenes and formylated phloroglucinol compounds in Eucalyptus nitens. New Phytol 176:82–95

Jansen RC, Stam P (1994) High resolution of quantitative traits into multiple loci via interval mapping. Genetics 136:1447–1455

Joët T, Laffargue A, Salmona J, Doulbeau S, Descroix F, Bertrand B, de Kochko A, Dussert S (2009) Metabolic pathways in tropical dicotyledonous albuminous seeds: *Coffea arabica* as a case study. New Phytol 182:146–162

Joët T, Salmona J, Laffargue A, Descroix F, Dussert S (2010) Use of the growing environment as a source of variation to identify the quantitative trait transcripts and modules of co-expressed genes that determine chlorogenic acid accumulation. Plant Cell Environ 33:1220–1233

Kenis K, Keulemans J, Davey MW (2008) Identification and stability of QTLs for fruit quality traits in apple. Tree Genet Gen 4:647–661

Klann EM, Hall B, Bennett AB (1996) Antisense acid invertase (*TIV1*) gene alters soluble sugar composition and size in transgenic tomato fruit. Plant Physiol 112:1321–1330

Ky CL, Louarn J, Guyot B, Charrier A, Hamon S, Noirot M (1999) Relations between and inheritance of chlorogenic acid contents in an interspecific cross between *Coffea pseudozanguebariae* and *Coffea liberica* var 'dewevrei'. Theor Appl Genet 98:628–637

Ky CL, Doulbeau S, Guyot B, Akaffou S, Charrier A, Hamon S, Louarn J, Noirot M (2000a) Inheritance of coffee bean sucrose content in the interspecific cross *Coffea pseudozanguebariae* x *Coffea liberica* 'dewevrei'. Plant Breed 119:165–168

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

Ky CL, Louarn J, Dussert S, Guyot B, Hamon S, Noirot M (2001a) Caffeine, trigonelline, chlorogenic acids and sucrose diversity in wild *Coffea arabica* L. and *C. canephora* P. accessions. Food Chem 75:223–230

Ky CL, Guyot B, Louarn J, Hamon S, Noirot M (2001b) Trigonelline inheritance in the interspecific *Coffea pseudozanguebariae* x *C. liberica* var. dewevrei cross. Theor Appl Genet 102:630–634

Lashermes P, Couturon E, Moreau N, Paillard M, Louarn J (1996) Inheritance and genetic mapping of self -incompatibility in *Coffea canephora* Pierre. Theor Appl Genet 93:458–462

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

Lefebvre-Pautigny F, Wu F, Philippot M, Rigoreau M, Priyono ZM, Frasse P, Bouzayen M, Broun P, Pétiard V, Tanksley SD, Crouzillat D (2010) High resolution syntheny maps allowing direct comparisons between the coffee and tomato genome. Tree Genet Gen 6:565–577

Lepelley M, Cheminade G, Tremillon N, Simkin A, Caillet V, McCarthy J (2007) Chlorogenic acid syntheses in coffee: An analysis of CGA content and real-time RT-PCR expression of HCT, HQT, C3H1, and CCoAOMT1 genes during grain development in *C. canephora*. Plant Sci 172:978–996

Leroy T, Montagnon C, Charrier A, Eskes AB (1993) Reciprocal recurrent selection applied to *Coffea canephora* Pierre I. Characterization and evaluation of breeding populations and value of intergroup hybrids. Euphytica 67:113–125

Leroy T, Montagnon C, Cilas C, Charrier A, Eskes AB (1994) Reciprocal recurrent selection applied to *Coffea canephora* Pierre II. Evaluation of genetic parameters. Euphytica 74:121–128

Leroy T, Marraccini P, Dufour M, Montagnon C, Lashermes P, Sabau X, Ferreira LP, Jourdan I, Pot D, Andrade AC, Glaszmann JC, Vieira LGE, Piffanelli P (2005) Construction and characterization of a *Coffea canephora* BAC library to study the organization of sucrose biosynthesis genes. Theor Appl Genet 111:1032–1041

Leroy T, Ribeyre F, Bertrand B, Charmetant P, Dufour M, Montagnon C, Marraccini P, Pot D (2006) Genetics of coffee quality. Braz J Plant Physiol 18:229–242

Lin C, Mueller LA, Mc Carthy J, Crouzillat D, Pétiard V, Tanksley S (2005) Coffee and tomato share common gene repertoires as revealed by deep sequencing of seed and cherry transcripts. Theor Appl Genet 112:114–130

Lu H, Romero-Severson J, Bernardo R (2002) Chromosomal regions associated with segregation distortion in maize. Theor Appl Genet 105:622–628

Misako K, Kouichi M (2004) Caffeine synthase and related methyltransferases in plants. Front Biosci 9:1833–1842

Moncada P, McCouch S (2004) Simple sequences repeat diversity in diploid and tetraploid *Coffea* species. Genome 47:501–509

Montagnon C, Leroy T, Yapo A (1992) Diversité génotypique et phénotypique de quelques groupes de caféiers (*Coffea canephora* Pierre) en collection. Conséquences sur leur utilisation en sélection. Café Cacao Thé 36:187–198

Montagnon C, Guyot B, Cilas C, Leroy T (1998) Genetic parameters of several biochemical compounds from green coffee, *Coffea canephora*. Plant Breed 117:576–578

Montagnon C, Leroy T, Cilas C, Charrier A (2003) Heritability of *Coffea canephora* yield estimated from several mating designs. Euphytica 133:209–218

Moschetto D, Montagnon C, Guyot B, Perriot JJ, Leroy T, Eskes AB (1996) Studies on the effect of genotype on cup quality of *Coffea canephora*. Trop Sci 36:18–31

N'Diaye A, Noirot M, Hamon S, Poncet V (2007) Genetic basis of species differentiation between *Coffea liberica* Hiern and *C. canephora* Pierre: analysis of an interspecific cross. Genet Resour Crop Evol 54:1011–1021

Paillard M, Lashermes P, Pétiard V (1996) Construction of a molecular linkage map in coffee. TheorAppl Genet 93:41–47

Poncet V, Hamon P, Minier J, Carasco C, Hamon S, Noirot M (2004) SSR cross-amplification and variation within coffee trees (*Coffea* spp.). Genome 47:1071–1081

Poncet V, Dufour M, Hamon P, Hamon S, de Kochko A, Leroy T (2007) Development of genomic microsatellite markers in *Coffea canephora* and their transferability to other coffee species. Genome 50:1156–1161

Priyono FB, Rigoreau M, Ducos JP, Sumirat U, Mawardi S, Lambot C, Broun P, Pétiard V, Wahyudi T, Crouzillat D (2010) Somatic embryogenesis and vegetative cutting capacity are under distinct control in *Coffea canephora* Pierre. Plant Cell Rep 29:343–357

Risterucci AM, Grivet L, N'Goran JAK et al (2000) A high-density linkage map of *Theobroma cacao* L. Theor Appl Genet 101:948–955

Saliba V, Causse M, Langlois D, Philouze J, Buret M (2001) Genetic analysis of organoleptic quality in fresh-market tomato. 1. Mapping QTLs for physical and chemical traits. Theor Appl Genet 102:273–283

Segura V, Durel CE, Costes E (2009) Dissecting apple architecture into genetic, ontogenetic and environmental effects: QTL mapping. Tree Genet Gen 5:165–179

Van Ooijen JW (2004) MapQTL 5, software for the mapping of quantitative trait loci in experimental populations. Kyazma BV, Wageningen

Van Ooijen JW (2006) JoinMap 4 software for the calculation of genetic linkage maps. Kyazma BV, Wageningen

Venkateswarlu M, Urs SR, Nath BS, Shashidar HE, Maheswaran M, Veeraiah TM, Sabitha MG (2006) A first genetic linkage map of mulberry (*Morus* spp.) using RAPD, ISSR and SRR markers and pseudo testcross mapping strategy. Tree Genet Gen 3:15–24

Voorrips RE (2002) Mapchart: software for the graphical presentation of linkage maps and QTLs. J Hered 93:77–78

Wintgens JN (2004) Coffee: growing, processing, sustainable production A guidebook for growers, processors, traders, and researchers. Wiley-VCG Verlag GmbH & Co., Weinheim

Xu S (2008) QTL mapping can benefit from segregation distortion. Genetics 180:2201–2208

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

Zhang L, Wang S, Li H, Deng Q, Zheng A, Li S, Li P, Li Z, Wang J (2010) Effects of missing markers and segregation distortion on QTL mapping in F2 populations. Theor Appl Genet. doi:10.1007/s00122-010-1372-z