Journal of Heredity Advance Access published October 29, 2011

Journal of Heredity
doi:10.1093/jhered/esr120

Homeologous Gene Expression in Response to Growing Temperature in a Recent Allopolyploid (*Coffea arabica* L.)

Marie-Christine Combes, Alberto Cenci, Hélène Baraille, Benoît Bertrand, and Philippe Lashermes

From the Institut de Recherche pour le Développement (IRD), UMR RPB (IRD, Centre de coopération Internationale en Recherche Agronomique pour le Développement [CIRAD] Université Montpellier II), BP 64501, 34394 Montpellier Cédex 5, France (Combes, Cenci, Baraille, and Lashermes); and the CIRAD, UMR RPB (IRD, CIRAD, Université Montpellier II), BP 64501, 34394 Montpellier Cédex 5, Montpellier, France (Bertrand).

Address correspondence to Marie-Christine Combes at the address above, or e-mail: marie-christine.combes@ird.fr.

Abstract

Allopolyploidy is considered as a major factor contributing to speciation, diversification, and plant ecological adaptation. In particular, the expression of duplicate genes (homeologs) can be altered leading to functional plasticity and to phenotypic novelty. This study investigated the influence of growing temperatures on homeologous gene expression in *Coffea arabica* L., a recent allopolyploid involving 2 closely related diploid parental species. The relative expression of homeologs of 13 genes all located in the same genomic region was analyzed using an SNP ratio quantification method based on dideoxy-terminated sequences of cDNA amplicons. The relative expression of homeologous genes varied depending on the gene, the organ, and the growing condition. Nevertheless, expression of both homeologs was always detected (i.e., no silencing). Although the growing conditions were suitable for one or other of the parental species, neither subgenome appeared preferentially expressed. Furthermore, relative homeologous expression showed moderate variations across organs and conditions and appeared uncorrelated between adjacent genes. These results indicate the absence of signs of subfunctionalization suggesting *C. arabica* has not undergone noticeable diploidization. Furthermore, these results suggest that the expression of homeologous genes in *C. arabica* is regulated by a shared *trans*-regulation mechanism acting similarly on the 2 subgenomes and that the observed biases in the relative homeolog expression may result from *cis* fine-scale factors.

Key words: *adaptation, coffee, duplicate gene expression, introgression, polyploidy, transcription*

Polyploidy, a feature of organisms that have multiple sets of chromosomes, occurs in eukaryotes and is particularly widespread in flowering plants (angiosperms) including many major crops (wheat, maize, sugarcane, potato, soybean, coffee, and cotton) (Leitch AR and Leitch IJ 2008). Whereas an autopolyploid results from doubling a diploid genome, an allopolyploid is formed by the combination of 2 or more sets of distinct genomes (homeologous chromosomes) after interspecific hybridization and chromosome doubling (Chen 2007). In the last decade, in numerous studies on polyploids formed at different timescales (from a few years up to several million years ago and most recently on synthetic polyploids), dynamic and stochastic changes in genomic and epigenetic organization and in gene expression have been documented (Chen and Ni 2006; Chen 2007). Polyploidy is considered as a major factor contributing to speciation, diversification, and ecological adaptation (Osborn et al. 2003; Adams 2007; Peng et al. 2008; Jackson and Chen 2010). Genomic plasticity (Leitch AR and Leitch IJ 2008) enables polyploids to combine multiple sets of genetic material in the same nucleus (Chen 2007; Hegarty and Hiscock 2009), whereas functional plasticity of duplicate genes enables them to regulate gene expression to adapt to their environment (Comai 2005; Jackson and Chen 2010; Buggs et al. 2011).

In several plant allopolyploid species such as *Gossypium hirsutum*, *Triticum aestivum*, *Tragopogon miscellus*, and *Arabidopsis suecica*, variations in gene expression have been investigated by either comparing the expression of the diploid parents and the allopolyploids or comparing the ratio of transcripts derived from each homeolog. Changes in homeologous expression have been explored at the genome scale in a single organ (Flagel et al. 2008; Hovav et al. 2008; Vidal et al. 2010) and in one to several dozen genes in different organs (Adams et al. 2003, 2004; Mochida et al. 2003; Nomura et al. 2005; Chaudhary et al. 2009; Flagel et al. 2009; Buggs et al. 2010). However, few studies have explored homeologous gene expression in several organs in relation to different types of

stress (Liu and Adams 2007; Stamati et al. 2009; Dong and Adams 2011). In the allopolyploid, locally or throughout the genome, homeologous genes may exhibit unequal expression patterns. The homeologous expression ratio can vary with the organ, its development, and in response to abiotic stresses. In some instance, complete silencing of one copy or reciprocal silencing (silencing of a homeologous gene in one organ and silencing of the other homeologous gene in another organ) is observed. Complex genetic and epigenetic regulatory mechanisms govern genome-specific expression biases in allopolyploid plant species. No single unifying factor has been reported suggesting that each allopolyploid plant has its own regulatory mechanism (Chaudhary et al. 2009). Along these lines, the homeologous expression ratio contributes to the fate of duplicated genes and to the evolution of allopolyploid plants after polyploidization. For instance, studies of gene expression patterns in different organs revealed adaptation mechanisms for evolution (neofunctionalization, subfunctionalization, and nonfunctionalization) (Buggs et al. 2010). All these findings on changes in homeologous gene expression attest to the capacity of polyploids for functional and adaptative diversification.

Coffee is one of the most important agricultural commodities and provides the main livelihood for more than 80 million people worldwide. Among the around 100 *Coffea* species identified (Davis et al. 2006), 2, *Coffea arabica* and *C. canephor*a, are cultivated and account for 65% and 35% of world coffee production, respectively. *Coffea* is a relatively young genus (whose origin was estimated at 100 000–450 000 years before present; Anthony et al. 2010), and all *Coffea* species are diploid, except for *C. arabica*, which is allotetraploid ($2n = 4\times = 44$) and derived from a recent (10 000–50 000 years ago) interspecific hybridization between 2 diploid species: *C. eugenioides* and *C. canephora* (Cenci et al. forthcoming; Lashermes et al. 1999). Homeologous genomes in *C. arabica* are designated E^a and C^a according to their parental origin. The 2 parental species exhibit different agroecological adaptations. *Coffea canephora* and *C. eugenioides* are endemic in regions where the annual mean temperature ranges from, respectively, 22 to 26 °C and 18 to 23 °C with no substantial oscillations (DaMatta and Cochicho Ramalho 2006; Davis et al. 2006). In contrast, *C. arabica*, whose optimum mean annual temperature range from 20 to 24 °C, can be grown in regions with marked variations in thermal amplitude. The 2 parental species are closely related, and the 2 subgenomes exhibit low sequence divergence (Cenci et al. forthcoming). Nevertheless, *C. arabica* displays diploid like meiotic behavior (Lashermes 2000). Preliminary results of a comparison of genome-wide expression patterns in *C. arabica* and its ancestral parents (Bardil et al. 2011) showed genomic expression dominance like that reported in nascent *Gossypium* allopolyploids (Rapp et al. 2009). Moreover, an expressed sequence tag (EST)–based study of differential homeologous gene expression in *C. arabica* (Vidal et al. 2010) showed both equal and biased homeologous genes expression.

To investigate the influence of growing temperatures on homeologous gene expression, we analyzed homeologous transcript ratio in different organs of *C. arabica* grown in 2 contrasted culture conditions known to be stressful for one or the other parental species. The relative expression of homeologs of 13 genes, all located in the same genomic region (Cenci et al. 2010; Lashermes et al. 2010), was determined using the SNP ratio quantification method based on dideoxy-terminated sequences of cDNA amplicons. Among previous studies on polyploids, this study completes the analysis of the variations in homeologous gene expression and informs about the behavior of duplicated genes in a recent allopolyploid. This study also provides new evidences regarding the *C. arabica*'s ability to adapt to environmental variations.

Materials and Methods

Plant Materials, RNA Extractions, and Reverse Transcription

Individuals of 2 *C. arabica* inbred varieties, Caturra and S795, were used. Caturra is a regular cultivar (genotype $E^aE^aC^aC^a$), whereas S795 was derived from an interspecific cross between *C. arabica* and *C. liberica* and carries several *C. liberica*–introgressed segments (Prakash et al. 2004). In particular, an introgression occurred on the subgenome C^a in a region named S_H3 (Lashermes et al. 2010), resulting locally in a genotype E^aE^aLL.

For the first study, seedlings of *C. arabica* var. Caturra were cultivated for 2 months in 2 sets of contrasted growing conditions with different day/night temperatures: 33 °C/30 °C (warmer conditions) and 23 °C/20 °C (colder conditions). The conditions were chosen to suit either *C. canephora* or *C. eugenioides*. In the 2 climatic chambers, the photoperiod, the hygrometry, and the luminosity were set at 12 h per day, 80–90%, and 600 $\mu E/m^2/s^1$, respectively. In each climatic chamber, the plants were grown in a randomized complete block design. The cotyledons, young leaves, leaves, stems, and roots were harvested on 3 plants (biological replicates) for each culture condition.

For the second study, *C. arabica* var. Caturra and S795 were cultivated in a greenhouse. The temperatures were set at 26 ± 3 °C in the day and at 24 ± 3 °C at night. The hygrometry was maintained over 60%. Young leaves, leaves, stems, and roots were collected from 3 S795 plants and 2 Caturra plants of similar size.

In both studies, the collected samples were immediately flash frozen in liquid nitrogen and stored at −80 °C until RNA extraction.

RNA was extracted from all tissues using the Qiagen (Valencia, CA) RNeasy kit according to the manufacturer's instructions with a few modifications. The tissues (100 mg) were ground in liquid nitrogen. The tubes were centrifuged at 4 °C and kept in ice throughout extraction and DNAse treatment. RNA was eluted by impregnating the purification column with 65 and 25 µl of sterile water for, respectively, 30 and 5 min before centrifugation. All RNA samples were quantified and visually assessed for degradation and DNA contamination using a Nanodrop (ThermoFisher Scientific Inc, Waltham, MA) and gel analysis.

Reverse Transcription was performed using 1 μg of total RNA with the ImProm-II Reverse Transcription System (Promega, Madison, WI). As controls for DNA contamination, reactions were performed without reverse transcriptase.

Design of Primers for Gene-Specific Amplification and Sequencing

Based on both *C. arabica* subgenome sequences from the S_H3 region (Cenci et al. 2010, forthcoming, Lashermes et al. 2010), SNPs differentiating single-copy homeologous genes were identified. For 13 different genes spanning 350 kb of the S_H3 region, primer pairs were designed on a single exon to amplify a fragment larger than 100 bp containing at least one SNP (Table 1).

PCR reactions were performed twice in a volume of 20 μl with either 1 μl of the diluted (one-fifth) cDNA generated by the first-strand synthesis or 25 ng of genomic DNA, 0.4 μM of each primer, 2.5 mM $MgCl_2$, and 1 U of *Taq* DNA polymerase. Cycling was done in a GeneAmp PCR 9700 thermocycler for 2 min at 94 °C followed by 5 cycles of 10 s at 94 °C and 30 s at 60–55 °C (−1 °C per cycle) and by 30 cycles of 10 s at 94 °C, 30 s at 55 °C, 30 s at 72 °C, and a final 8-min extension at 72 °C. Sequencing was performed by Beckman Coulter Genomics (Takeley, UK).

Relative Expression of Homeologs

The relative expression of homeologs was estimated using a method similar to that of Rauscher et al. (2002) on sequence chromatograms of exon portions amplified on cDNA. Chromatogram files were analyzed using BIOEDIT (Hall 1999). The height of both peaks at an SNP position indicating the presence of the 2 homeoalleles was measured

Table 1 Homeologous genes investigated, SNP, and pairs of oligonucleotides used to study their relative contribution to the transcriptome of *Coffea arabica*

Genes	Putative function (BLASTP e value)	Primer sequences 5′ > 3′	Size of the amplified product (bp)	Position of SNPs (differentiating bases E[a] and C[a])
S_H3-44	Serine–threonine protein kinase, plant type ($E < 0.05$)	For: ggctacattgcccctgaata Rev: caagcaagtgcgacagtcat	269	173 bp (A/G)
S_H3-48	Calcineurin B-like calcium sensor protein-interacting serine/threonine protein kinase (8×10^{135})	For: aaaacgggatgcttcatacg Rev: gtcggattaggatccagcaa	265	154 bp T/C
S_H3-51a	Beta-ketoacyl-CoA synthase ($E < 0.05$)	For: aaattgcatctttcggatgg Rev: aactaatggcccgagtgttg	247	63 bp G/A, 83 bp A/G
S_H3-55	Hypothetical protein (2×10^{123})	For: gacaaccgttcctcctcaaa Rev: ttgatacttcgccggaaatc	317	131 bp C/T
S_H3-56	Pyruvate kinase ($E < 0.05$)	For: gagcacacggaggactaagc Rev: tcctcattcaacctccgaac	171	101 bp A/G
S_H3-61	Chloroplast inner envelope protein (9×10^{156})	For: tcaaggccagtttctcaacc Rev: tagcaagttgatgcggtgac	274	84 bp T/C, 131 bp G/A
S_H3-66	Sulfate adenylyltransferase ($E < 0.05$)	For: gagaccactgtggtatctatattt Rev: gtaccttcttcccatgatctg	187	129 bp C/G
S_H3-67	Eukaryotic translation initiation factor 2c ($E < 0.05$)	For: agtgccagagctcaatg Rev: caccttgatcagatttccg	171	85 bp T/C, 129 bp T/C
S_H3-74	Putative early light-induced protein 2 (2×10^{151})	For: gccatgatagggtttgttgc Rev: aagggcacctcctttgagat	270	119 bp C/T
S_H3-76	DNA-binding protein (2×10^{99})	For: gctaatgaagggttcttaaggg Rev: catcatctgctgctgtttat	296	134 bp C/T
S_H3-78	Acireductone dioxygenase (4×10^{91})	For: gatgaggttgtttgtgggaga Rev: gcactcccaccttctcagtg	110	75 bp G/A
S_H3-81	Chlorophyllase-2, chloroplast (5×10^{122})	For: cttatcaaagctggcgcta Rev: gcctgatccgatgactaatac	217	147 bp A/G
S_H3-83	Ice-binding protein (3×10^{176})	For: tctgcacagggaaagga Rev: agctgttaaggtcatgatgtc	186	142 bp C/T, 147 bp T/C

For, forward; Rev, reverse.

using the coordinate system. Because the sequence process generated variations in peak size for a given position, the height of the different haplotypes was calibrated using chromatograms obtained from the amplification of *C. arabica* genomic DNA, where the E^a and C^a haplotypes are expected to be equally represented. For each SNP position, correction factor thus determined was applied to the peak heights obtained from different cDNA samples. The contribution of the E^a homeolog was expressed as the height of the E^a peak over the sum of the heights of the E^a and C^a peaks and conversely.

Two-way ANOVA and correlation analysis, performed with R software, were used to identify significant differences in homeologous gene expression as a function of the organ and growing condition concerned. The matrices of distances between genes based on pairwise correlations of relative homeologous gene expression and physical distances were compared by the Mantel test (Mantel 1967; Piepho 2005) using GENETIX software version 4.01 (Laboratoire Génome et Populations, CNRS UPR 9060, Université de Montpellier II, Montpellier, France).

Method Validation

Two experiments were performed to validate the sequencing-based method used to determine the relative contribution of the homeologs. First, the relative proportion of SNP haplotypes was measured for all genes in amplicons from DNA samples showing different genome combinations: *C. arabica* var. Caturra (tetraploid, 1/2 E^a, 1/2 C^a), H855 triploid hybrid: *C. arabica* × *C. canephora* (1/3 E^a, 2/3 C^a + C,), and Arabusta tetraploid hybrid: *C. arabica* × "tetraploid" *C. canephora* (1/4 E^a, 3/4 C^a + C,). Second, DNAs from bacterial artificial chromosome (BAC) clones containing the S_H3 region derived from both E^a and C^a subgenomes were mixed with different proportions of the targeted region (0.84 E^a/0.16 C^a, 0.63 E^a/0.37 C^a, and 0.36 E^a/0.64 C^a). The relative proportion of the 2 SNP haplotypes in the amplicons was measured after amplification of the mixed solutions for 3 genes: S_H3-66, S_H3-76, and S_H3-81.

Results

Assessment of the Relative Contribution of Homeologs to the Transcriptome

In order to evaluate the reliability of the quantitative analysis of SNP peaks on sequence chromatograms, the expected and estimated proportion of SNP haplotypes in the amplicons were compared in 2 experiments. In the first experiment, the proportion of SNP haplotypes was analyzed in amplicons from genomic DNA extracted from genotypes with different genomic composition; the results obtained for the S_H3-66 gene are presented to illustrate the experiment (Figure 1A). In the second experiment, the proportion of SNP haplotypes for the S_H3-66, S_H3-76, and S_H3-81 genes was estimated in amplicons from mixed homeologous BAC DNA solutions (Figure 1B). In both experiments, the observed proportions of SNP were very close to expected values (in the first experiment, r ranged from 0.94 to 0.98 depending on the genes considered with $P < 0.001$, in the second experiment for S_H3-66, $r = 0.99$ with $P < 0.003$, for S_H3-76, $r = 0.99$ with $P < 0.002$, for S_H3-81, $r = 0.99$ with $P < 0.002$), indicating

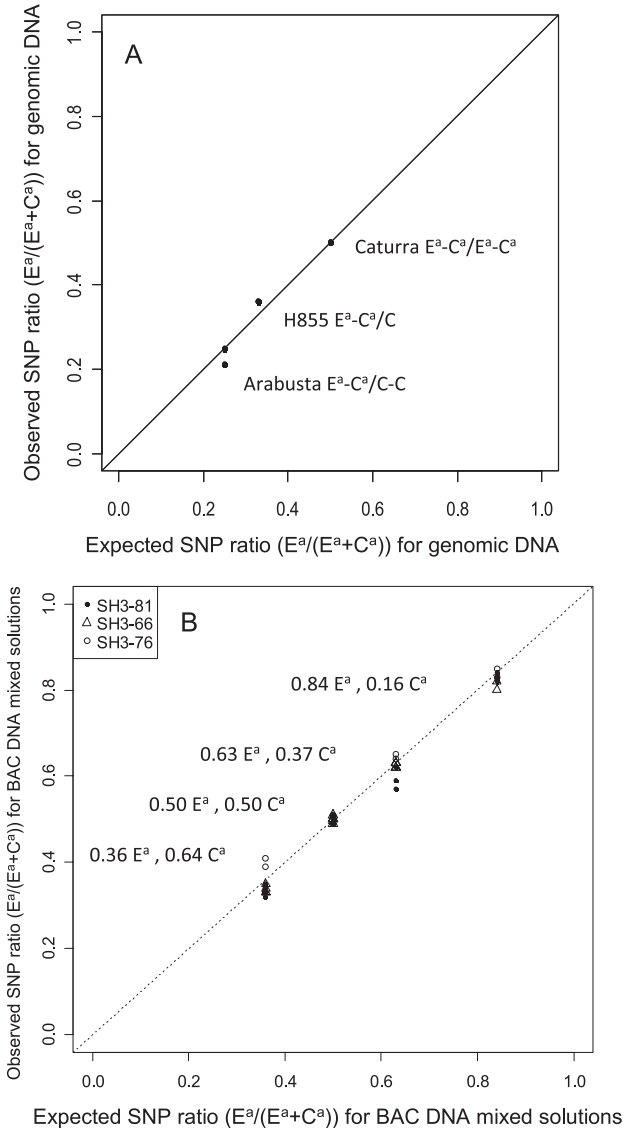

Figure 1. Validation of the sequencing-based method to determine the relative contribution of the homeologs using 2 methods. (**A**) Example (gene S_H3-66) for measure of proportion of SNP haplotypes in amplicons from 3 types of genomic DNA: *Coffea arabica* Caturra (tetraploid, 1/2 E^a, 1/2 C^a), H855 triploid hybrid *C. arabica* × *C. canephora* (2/3 C^a + C, 1/3 E^a), and Arabusta tetraploid hybrid from a cross between *C. arabica* and "tetraploid" *C. canephora* (3/4 C^a + C, 1/4 E^a). (**B**) The relative proportion of 2 types of the SNP in amplicons from different mixed BAC DNA solutions (0.84 E^a/0.16 C^a, 0.63 E^a/0.37 C^a, and 0.36 E^a/0.64 C^a) and *C. arabica* var. Caturra (tetraploid, 1/2 E^a, 1/2 C^a) for 3 genes: S_H3-66, S_H3-76, and S_H3-81. In the 2 figures, the values obtained were compared with the theoretical ratio curve.

that the method based on dideoxy-terminated sequencing accurately determined the ratios of homeologs in an amplicon.

For 3 genes (S_H3-51a, S_H3-61, and S_H3-83), which displayed several SNPs in the analyzed amplicons, the homeologous gene expression ratio for the overall set of samples (60 samples) was estimated on 2 SNPs. The relative contributions of homeologs to the transcriptome at both SNP positions were compared. The 2 measures were found to be significantly correlated (S_H3-51a: $r = 0.93$ with $P < 2.2 \times 10^{-16}$, S_H3-61: $r = 0.80$ with $P < 9.4 \times 10^{-11}$, and S_H3-83: $r = 0.87$, $P < 2.2 \times 10^{-16}$).

Finally, the variation coefficients of the estimation of the homeologous gene expression on replicates for all samples were calculated. The deviation between the technical replicates was less than 10% for 90% of the measures.

Homeologous Expression of Genes in the S_H3 Region of *C. arabica*

The relative expression of the homeologs was assayed for the overall set of samples representing 5 organs and 2 growing conditions (Figure 2). The relative expression of the homeologs was represented by the E^a relative expression ratio [ratio $E^a/(E^a + C^a)$]. For all 13 genes analyzed in the S_H3 region, both homeologous genes appeared to be expressed although their respective contribution varied. When all the 13 genes were considered, neither of the 2 subgenomes was preferentially expressed. Based on the overall relative expression of the homeologs for each gene, 4 categories of genes were distinguished. The first category contained genes S_H3-51a, S_H3-66, S_H3-74, and S_H3-67 whose relative contribution to the transriptome was biased toward the E^a genome. Inversely, the second category contained genes S_H3-44, S_H3-48, S_H3-81, and S_H3-83 for whose relative contribution was biased toward the C^a genome. For genes S_H3-55, S_H3-56, and S_H3-61, the expression of the homeolog E^a was practically equal to the expression of the homeolog C^a. Finally, for genes S_H3-76 and S_H3-78, the predominantly expressed homeolog differed with the organ and the culture condition concerned.

Effect of the Organ and Growing Conditions

The relative expression of homeologs varied with the organ analyzed. ANOVA (Table 2) indicated that 10 genes out of 13 had the statistically significant probability of the "organ" factor ($P < 0.001$). Depending on the gene, the variation in the homeologous gene expression depended either on one organ (S_H3-66) or on overall heterogeneity between organs (S_H3-76 and S_H3-78). No organ-specific effect was observed.

According to the ANOVA, 4 and 6 genes out of 13 showed statistically significant differences between "growing conditions" at probabilities $P < 0.001$ and $P < 0.05$, respectively. For some genes (Figure 2), the E^a relative expression ratio was stronger or weaker in colder and warmer conditions depending on the organ concerned, whereas for other genes (S_H3-55 and S_H3-76) exhibited a higher contribution by the homeolog E^a in colder conditions for all organs than in warmer conditions or conversely (S_H3-44).

Comparison of Neighboring Genes

To compare the behavior of genes spread over 350 kb of the S_H3 region, expression profiles representing the homeologous gene expression for the 13 genes per organ were superimposed on the warmer and colder conditions (Figure 3). The correlation between these organ expression profiles was estimated. In warmer conditions, all correlation coefficients were statistically significant (5 coefficients, $P < 0.001$; 4 coefficients, $P < 0.01$; 1 coefficient, $P < 0.05$). In colder conditions, the situation was different, 6 correlation coefficients were statistically significant (1 coefficient, $P < 0.001$; 3 coefficients, $P < 0.01$; 2 coefficients, $P < 0.05$). The deviations between tissue expression profiles were greater in colder conditions (S_H3-48, S_H3-74, S_H3-76, S_H3-78, and S_H3-81) than in warmer conditions.

We performed statistical analyses on the S_H3 region to determine if physical distance had an effect on pairwise correlations of relative homeologous gene expression across tissues and culture conditions. The significance of this factor was assessed using permutation tests with physical distance measured in base pairs. These distance data were subjected to 1000 random permutations, and physical distance was found to be nonsignificantly correlated with the relative homeologous gene expression for the S_H3 region (Pearson's $r = 0.098$; $P = 0.21$).

Homeologous Gene Expression in an Introgressed Genotype

To analyze the impact of introgression on homeologous gene expression, the relative contribution of homeologs to the transcriptome was analyzed in genes S_H3-81 and S_H3-83 in 2 *C. arabica* inbred varieties carrying (var. S795) or not (var. Caturra) a *C. liberica*–derived introgression. The ratios of the relative homeologous gene expression $E^a/(E^a + C^a)$ for var. Caturra and $E^a/(E^a + L)$ for the introgressed var. S795 were compared in 4 organs (Figure 4). Depending on the gene, introgressed homeologs were expressed differently from the E^a homeolog in almost all organs. For gene S_H3-81, the contribution to the transcriptome of the E^a homeolog was higher in the introgressed variety than in var. Caturra. In contrast, gene S_H3-83 showed a lower contribution of the homeolog E^a in the introgressed plants than in var. Caturra. As in the previous experiment, the organ factor was statistically significant ($P < 0.001$ for the gene S_H3-81 and $P < 0.01$ for the gene S_H3-83).

Discussion

Variation in the Homeologous Gene Expression

In an allopolyploid, 2 diverged genomes merge in a common nucleus and contribute either equally or disproportionately to the transcriptome (Adams et al. 2003; Chaudhary et al. 2009). In this study, in *C. arabica*, we analyzed the relative contribution of 13 homeologous gene pairs to the transcript pool in 5 organs and in 2

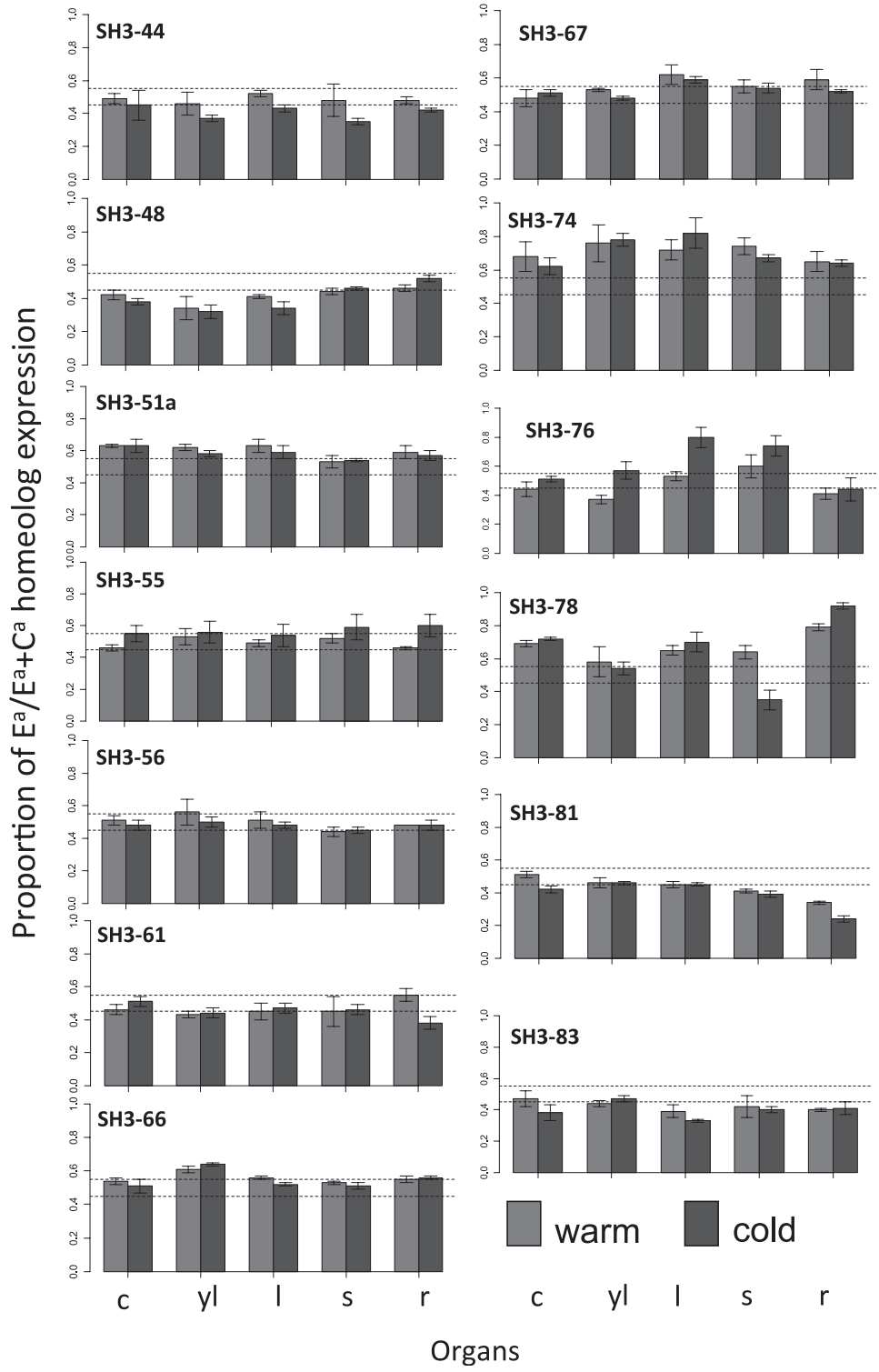

Figure 2. Representation of the variation in the relative contribution of homeologs ($E^a/E^a + C^a$) for the 13 genes as a function of the organ (c, cotyledon; yl, young leave; l, leaf; s, stem; r, root) and the culture conditions.

contrasted growth conditions. The genes analyzed in the present study are located in a 350-kb interval of the S_H3 region of *Coffea* spp. Because the sequence was available for both the subgenomes, it was possible to unequivocally assign each SNP variant to its respective subgenome (Lashermes et al. 2010). The relative contribution of the homeologous genes was estimated by the SNP ratio quantification method on dideoxy-terminated sequences

Table 2 ANOVA of the relative expression of the homeologs ($E^a/E^a + C^a$) for the 13 genes as a function of the organ and culture condition.

Studied factors	S_H3-44	S_H3-48	S_H3-51a	S_H3-55	S_H3-56	S_H3-61	S_H3-66	S_H3-67	S_H3-74	S_H3-76	S_H3-78	S_H3-81	S_H3-83
Organs													
Cotyledon	0.47	0.40	0.63	0.51	0.50	0.49	0.53	0.49	0.65	0.47	0.71	0.46	0.42
Young leaf	0.41	0.33	0.60	0.55	0.52	0.43	0.63	0.50	0.77	0.49	0.56	0.46	0.46
Leaf	0.47	0.37	0.61	0.52	0.49	0.46	0.54	0.61	0.77	0.65	0.67	0.45	0.36
Stem	0.42	0.45	0.54	0.55	0.45	0.45	0.52	0.55	0.70	0.67	0.49	0.40	0.41
Root	0.44	0.50	0.58	0.54	0.48	0.45	0.56	0.54	0.65	0.43	0.87	0.28	0.41
Probability *F*	0.019*	0.000***	0.000***	NS	0.000***	NS	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
Conditions													
Warm	0.49	0.40	0.60	0.50	0.50	0.46	0.56	0.55	0.71	0.48	0.66	0.44	0.42
Cold	0.40	0.41	0.58	0.57	0.48	0.45	0.55	0.53	0.71	0.61	0.65	0.39	0.40
Probability *F*	0.000***	NS	NS	0.000***	0.024*	NS	0.016*	0.036*	NS	0.000***	0.029*	0.000***	0.013*
Interactions													
Warm × cotyledon	0.49	0.42	0.63	0.46	0.51	0.46	0.54	0.48	0.68	0.44	0.69	0.51	0.47
Warm × young leaf	0.46	0.34	0.62	0.53	0.56	0.43	0.61	0.53	0.76	0.37	0.58	0.46	0.44
Warm × leaf	0.52	0.41	0.63	0.49	0.51	0.45	0.56	0.62	0.72	0.53	0.65	0.45	0.39
Warm × stem	0.48	0.44	0.53	0.52	0.44	0.45	0.53	0.55	0.74	0.60	0.64	0.41	0.42
Warm × root	0.48	0.46	0.59	0.46	0.48	0.55	0.55	0.59	0.65	0.41	0.79	0.34	0.40
Cold × cotyledon	0.45	0.38	0.63	0.55	0.48	0.51	0.51	0.51	0.62	0.51	0.72	0.42	0.38
Cold × young leaf	0.37	0.32	0.59	0.56	0.50	0.44	0.64	0.48	0.78	0.57	0.54	0.46	0.47
Cold × leaf	0.43	0.34	0.59	0.54	0.48	0.47	0.52	0.59	0.82	0.80	0.70	0.45	0.33
Cold × stem	0.35	0.46	0.54	0.59	0.45	0.46	0.51	0.54	0.67	0.74	0.35	0.39	0.40
Cold × root	0.42	0.52	0.57	0.60	0.48	0.38	0.56	0.52	0.64	0.44	0.92	0.24	0.41
Probability *F*	NS	0.01**	NS	NS	NS	NS	0.002**	0.029*	0.021*	0.017*	0.000***	0.000***	0.024*

***$P < 0.001$, **$P < 0.01$, *$P < 0.05$. NS, not significant.

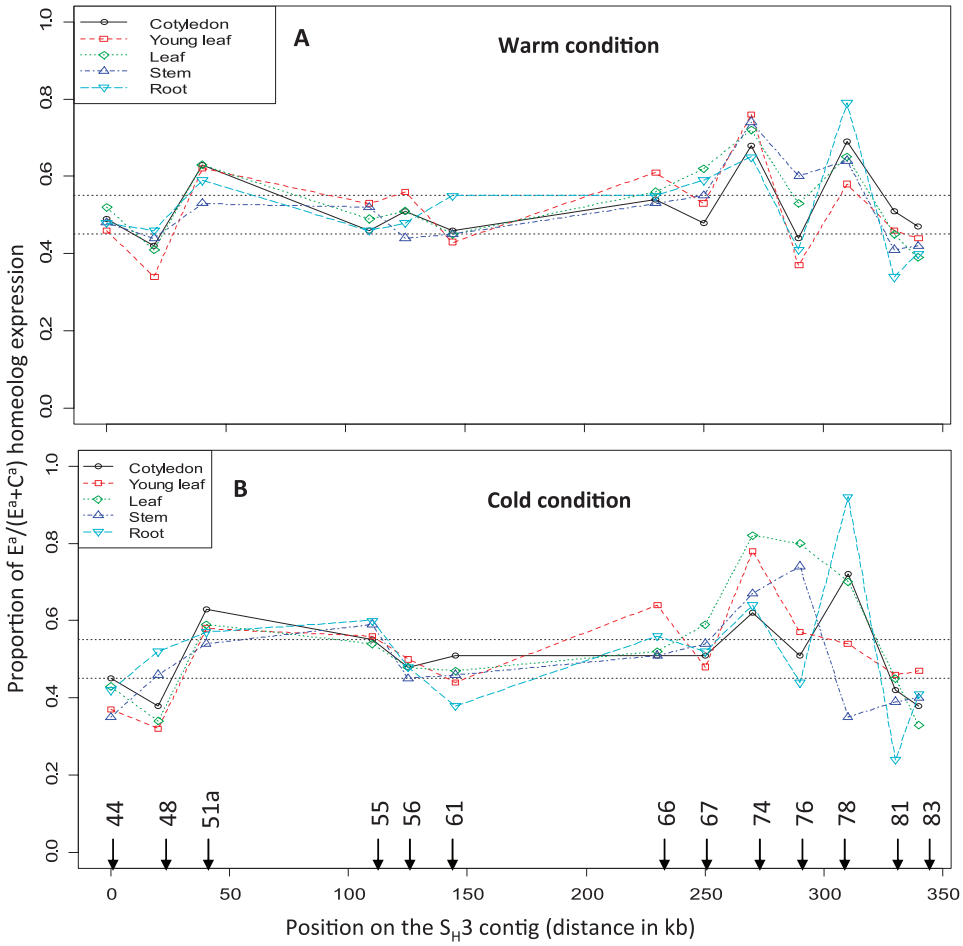

Figure 3. Representation of the average relative homeologous expression of 13 genes ($E^a/E^a + C^a$) in different organs and for 2 growth condition (**A**, warm; **B**, cold).

(Rauscher et al. 2002). Comparison of the ratio of 2 SNPs on the same gene and determination of the coefficients of variation of replicated samples of each SNP studied confirmed that the method is accurate and reliable. Although time consuming and more expensive per gene than methods enabling genome-wide analysis, this

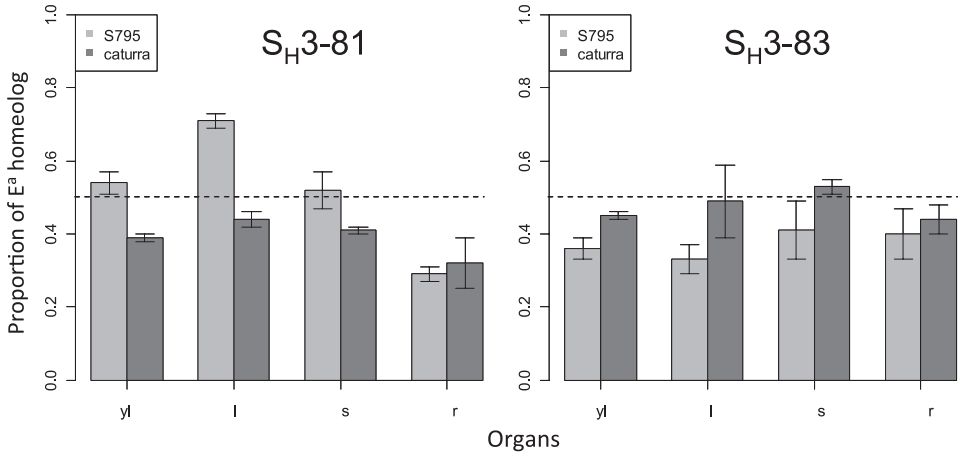

Figure 4. Variation in the relative contribution of homeologs between Caturra ($E^a/E^a + C^a$) and an introgressed variety S795 (E^a/E^a + L) for the genes S_H3-81 and S_H3-83 as a function of the organ (yl, young leaf; l, leaf; s, stem; r, root).

method appears to be suitable for the analysis of a limited gene sampling in a wide range number of conditions.

In our experiments involving combinations of 130 gene pairs/growing conditions/organs, the expression of both homeologs was always detected. As reported in several polyploids (Adams et al. 2003; Chaudhary et al. 2009; Flagel et al. 2009), the relative contribution of the homeologous genes to the transcriptome varied from equivalent to significantly biased expression depending on the organ and on the growing conditions. It is noteworthy that although the growing conditions were suitable for one or other of the parental species, neither subgenome appeared preferentially expressed. Similarly, the comparison between *C. arabica* and the *C. liberica*–introgressed *C. arabica* showed that both E^a and L homeologs were significantly expressed.

In our study, no case of gene silencing or organ-specific silencing was detected. Nevertheless, 10 out of 13 sampled genes exhibited biased expression, 4 genes toward E^a, 4 genes toward C^a, and 2 genes toward E^a or C^a depending on the organ considered. In the natural allopolyploid *G. hirsutum* (Adams et al. 2003), the proportion of biased homeologous genes was estimated at 25% of 40 genes and in the hexaploid wheat *T. aestivum* (Mochida et al. 2003) at 83% of 90 genes. In *C. arabica*, 22% of 2069 EST contigs analyzed showed biased expression (10% overexpressed C^a and 12% E^a) (Vidal et al. 2010). The difference between these results and ours could be due to the limited number of genes analyzed in the present study and to the different sensitivity and/or accuracy of the methods used to estimate relative homeologous gene expression. In particular, our approach based on SNP ratio quantification on dideoxy-terminated sequences was able to detect significant bias in homeolog expression even when the bias was small (10%), whereas only expression biases higher than 2/1 were considered in the in silico analysis performed by Vidal et al. (2010). In addition, because homeologous gene expression varies with the organ, the probability of detecting gene silencing in a study of several organs is greater than in a study of only one organ. Biologically, these divergent results indicate that homeologous gene expression after polyploidization may occur differently depending on the polyploid species. The biased gene expression may result from a combination of factors: inherited differences from diverged parental genomes and postdivergence structural changes between homeologs. (Lin et al 2010).

Organ divergence of duplicate gene expression patterns provides information on the evolution of the plant species (Buggs et al. 2010). Indeed, after the polyploidization event, the duplicated genes appear to evolve in 2 steps at different timescales (Adams and Wendel 2005; Flagel et al. 2008). The first step occurs immediately after the genomic merger of the diploids, and some of the duplicated genes disappear in genomic rearrangements. This is followed by a long-term evolution step during which a race between subfunctionalization (which enables preservation of duplicated genes) and pseudogenization (inactivation of a gene copy) starts for the duplicated genes (Rapp and Wendel 2005). Reciprocal epigenetic silencing or biased expression of homeologs in different organs or tissues may enhance the probability of retention of homeologous genes (Lynch and Force 2000; Chaudhary et al. 2009), thus preserving the availability of these genes for variation in expression and phenotypic diversification (Rapp and Wendel 2005; Chaudhary et al. 2009; Flagel et al. 2009). Thus, each allopolyploid lineage resulting from an independent event evolves in its own unique context according its evolutionary history (Chen and Ni 2006; Hufton and Panopoulou 2009; Jackson and Chen 2010). *Coffea arabica* is a recent stable allopolyploid involving little divergent diploid genomes. Our results showing limited variation of relative homeologous gene expression between organs and the absence of signs of subfunctionalization suggest that *C. arabica* has not undergone diploidization.

Regulation of the Homeologous Gene Expression

In polyploids, a variety of changes in gene expression have been observed in comparison with their diploid parents (Adams et al. 2004; Chaudhary et al. 2009). Some of these changes could be explained by interactions of homeologous regulatory factors with their target genes. Several regulatory models of the homeologous gene expression have been proposed (Riddle and Birchler 2003; Pignatta and Comai 2009). According to these models, a slight variation in the ratio of expression of the homeologous genes could mean that both transcription factors are able to bind almost equally well to both homeologous target genes. In wheat (Mochida et al. 2003), the homeologous gene expression and silencing were shown to be regulated at the level of the gene and not at the chromosome or genome level. In cotton, 2 regions appeared to be regulated very differently (Flagel et al. 2009): coordinated for 40 kb and fine-scale control (gene by gene), whereas Adams et al. (2003) reported that epigenetic factors act at the gene level together with other underlying controls.

In *C. arabica*, nonsignificant correlations between physical distance and relative homeologous gene expression suggest a gene-specific regulatory control. Because the 2 subgenomes exhibit low sequence divergence and variations in homeologous gene expression between the 2 subgenomes are limited and consistent across organs, it is likely that homeologous regulatory factors interoperate and control the regulation of the transcription of both subgenomes. Furthermore, depending on the growing temperature, other regulatory mechanisms including localized divergence among homeologous *cis*-regulatory regions or epigenetic marks could influence the homeologous gene contribution. In addition, comparison of homeologous gene expression of 2 genes in *C. arabica* and in an introgressed *C. arabica* suggests the participation of similar *trans*-regulation factors. In this particular context of allopolyploidy, homeologous newly introgressed genes from another species of the genus *Coffea* seem regulated like other homeologous genes in the genome of *C. arabica*. These regulation mechanisms could explain why *C. arabica* is able to adapt to different growing conditions better than its diploid parental species. It would be interesting to compare gene expression of *C. arabica* with its diploid progenitor to specify the relative roles of *cis*- and *trans*-regulation (Chaudhary et al. 2009).

References

Adams KL. 2007. Evolution of duplicate gene expression in polyploid and hybrid plants. J Hered. 98(2):136–141.

Adams KL, Cronn R, Percifield R, Wendel JF. 2003. Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. Proc Natl Acad Sci U S A. 100(8):4649–4654.

Adams KL, Percifield R, Wendel JF. 2004. Organ-specific silencing of duplicated genes in a newly synthesized cotton allotetraploid. Genetics. 168(4):2217–2226.

Adams KL, Wendel JF. 2005. Polyploidy and genome evolution in plants. Curr Opin Plant Biol. 8(2):135–141.

Anthony F, Diniz LEC, Combes MC, Lashermes P. 2010. Adaptive radiation in Coffea subgenus Coffea L. (Rubiaceae) in Africa and Madagascar. Plant Syst Evol. 285(1–2):51–64.

Bardil A, Dantas de Almeida J, Combes MC, Lashermes P, Bertrand B. 2011. Genomic expression dominance in the natural allopolyploid *Coffea arabica* is massively affected by growth temperature. New Phytol. doi: 10.1111/j.1469-8137.2011.03833.x.

Buggs RJA, Elliott NM, Zhang L, Koh J, Viccini LF, Soltis DE, Soltis PS. 2010. Tissue-specific silencing of homoeologs in natural populations of the recent allopolyploid *Tragopogon mirus*. New Phytol. 186(1):175–183.

Buggs RJA, Zhang L, Miles N, Tate JA, Gao L, Wei W, Schnable PS, Barbazuk WB, Soltis PS, Soltis DE. 2011. Transcriptomic shock generates evolutionary novelty in a newly formed natural allopolyploid plant. Curr Biol. 21:551–556.

Cenci A, Combes MC, Lashermes P. 2010. Comparative sequence analyses indicate that *Coffea* (Asterids) and *Vitis* (Rosids) derive from the same paleo-hexaploid ancestral genome. Mol Genet Genomics. 283(5):493–501.

Cenci A, Combes MC, Lashermes P. Forthcoming 2011. Genome evolution in diploid and tetraploid Coffea species as revealed by comparative analysis of orthologous genome segments. Plant Mol Biol. Mol Genet Genomics.

Chaudhary B, Flagel L, Stupar RM, Udall JA, Verma N, Springer NM, Wendel JF. 2009. Reciprocal silencing, transcriptional bias and functional divergence of homeologs in polyploid cotton (*Gossypium*). Genetics. 182(2):503–517.

Chen ZF. 2007. Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. Annu Rev Plant Biol. 58:377–406.

Chen ZF, Ni Z. 2006. Mechanisms of genomic rearrangements and gene expression changes in plant polyploids. Bioessays. 28(3):240–252.

Comai L. 2005. The advantages and disadvantages of being polyploid. Nat Rev Genet. 6(11):836–846.

DaMatta FM, Cochicho Ramalho JD. 2006. Impacts of drought and temperature stress on *Coffea* physiology and production: a review. Braz J Plant Physiol. 18(1):55–81.

Davis AP, Govaerts R, Bridson DM, Stoffelen P. 2006. An annotated taxonomic conspectus of the genus *Coffea (Rubiaceae)*. Bot J Linn Soc. 152:465–512.

Dong S, Adams KL. 2011. Differential contributions to the transcriptome of duplicated genes in response to abiotic stresses in natural and synthetic polyploids. New Phytol. 190(4):1045–1057.

Flagel L, Udall J, Nettleton D, Wendel J. 2008. Duplicate gene expression in allopolyploid *Gossypium* reveals two temporally distinct phases of expression evolution. BMC Biol. 6:1–16.

Flagel LE, Chen L, Chaudhary B, Wendel JF. 2009. Coordinated and fine-scale control of homoeologous gene expression in allotetraploid cotton. J Hered. 100(4):487–490.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. 41:95–98.

Hegarty MJ, Hiscock SJ. 2009. The complex nature of allopolyploid plant genomes. Heredity. 103(2):100–101.

Hovav R, Udall J, Chaudhary B, Rapp R, Flagel L, Wendel JF. 2008. Partitioned expression of duplicated genes during development and evolution of a single cell in a polyploid plant. Proc Natl Acad Sci U S A. 105(16):6191–6195.

Hufton AL, Panopoulou G. 2009. Polyploidy and genome restructuring: a variety of outcomes. Curr Opin Genet Dev. 19(6):600–606.

Jackson S, Chen ZJ. 2010. Genomic and expression plasticity of polyploidy. Curr Opin Plant Biol. 13(2):153–159.

Lashermes P. 2000. Brief communication. Single-locus inheritance in the allotetraploid Coffea arabica L. and interspecific Hybrid *C. arabica* x *C. canephora*. J Hered. 91(1):81–85.

Lashermes P, Combes MC, Ribas A, Cenci A, Mahé L, Etienne H. 2010. Genetic and physical mapping of the S_H3 region that confers resistance to leaf rust in coffee tree (*Coffea arabica L.*). Tree Genet Genomes. 6(6):973–980.

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(2):259–266.

Leitch AR, Leitch IJ. 2008. Genomic plasticity and the diversity of polyploid plants. Science. 320(5875):481–483.

Lin JY, Stupar RM, Hans C, Hyten DL, Jackson SA. 2010. Structural and functional divergence of a 1-Mb duplicated region in the soybean (*Glycine max*) genome and comparison to an orthologous region from *Phaseolus vulgaris*. Plant Cell. 22:2545–2561.

Liu Z, Adams KL. 2007. Expression partitioning between genes duplicated by polyploidy under abiotic stress and during organ development. Curr Biol. 17(19):1669–1674.

Lynch M, Force A. 2000. The probability of duplicate gene preservation by subfunctionalization. Genetics. 154(1):459–473.

Mantel N. 1967. The detection of disease clustering and generalized regression approach. Cancer Res. 27:209–220.

Mochida K, Yamazaki Y, Ogihara Y. 2003. Discrimination of homoeologous gene expression in hexaploid wheat by SNP analysis of contigs grouped from a large number of expressed sequence tags. Mol Genet Genomics. 270(5):371–377.

Nomura T, Ishihara A, Yanagita RC, Endo TR, Iwamura H. 2005. Three genomes differentially contribute to the biosynthesis of benzoxazinones in hexaploid wheat. Proc Natl Acad Sci U S A. 102(45):16490–16495.

Osborn TC, Pires JC, Birchler JA, Auger DL, Chen ZJ, Lee HS, Comai L, Madlung A, Doerge RW, Colot V, et al. 2003. Understanding mechanisms of novel gene expression in polyploids. Trends Genet. 19(3):141–147.

Peng H, Zhang J, Wu X. 2008. The ploidy effects in plant gene expression: progress, problems and prospects. Sci China Ser C Life Sci. 51(4):295–301.

Piepho HP. 2005. Permutation tests for the correlation among genetic distances and measures of heterosis. Theor Appl Genet. 111:95–99.

Pignatta D, Comai L. 2009. Parental squabbles and genome expression: lessons from the polyploids. J Biol. 8(4):43.

Prakash NS, Marques DV, Varzea VMP, Silva MC, Combes MC, Lashermes P. 2004. Introgression molecular analysis of a leaf rust resistance gene from *Coffea liberica* into *Coffea arabica L.* Theor Appl Genet. 109(6):1311–1317.

Rapp RA, Udall JA, Wendel JF. 2009. Genomic expression dominance in allopolyploids. BMC Biol. 7:18.

Rapp RA, Wendel JF. 2005. Epigenetics and plant evolution. New Phytol. 168(1):81–91.

Rauscher JT, Doyle JJ, Brown AHD. 2002. Internal transcribed spacer repeat-specific primers and the analysis of hybridization in the *Glycine tomentella (Leguminosae)* polyploid complex. Mol Ecol. 11(12):2691–2702.

Riddle NC, Birchler JA. 2003. Effects of reunited diverged regulatory hierarchies in allopolyploids and species hybrids. Trends Genet. 19(11):597–600.

Stamati K, Mackay I, Powell W. 2009. A quantitative genomic imbalance gene expression assay in a hexaploid species: wheat (*Triticum aestivum*). Genome. 52(1):89–94.

Vidal RO, Mondego JMC, Pot D, Ambrosio AB, Andrade AC, Pereira LFP, Colombo CA, Vieira LGE, Carazzolle MF, Pereira GAG. 2010. A high-throughput data mining of single nucleotide polymorphisms in *Coffea* species expressed sequence tags suggests differential homeologous gene expression in the allotetraploid *Coffea arabica*. Plant Physiol. 154(3):1053–1066.

Received June 20, 2011; Revised September 21, 2011; Accepted September 30, 2011

Corresponding Editor: Reid Palmer