Plant Physiology and Biochemistry 111 (2017) 340–347

Contents lists available at ScienceDirect

Plant Physiology and Biochemistry

journal homepage: www.elsevier.com/locate/plaphy

Research article

Diterpenes biochemical profile and transcriptional analysis of cytochrome P450s genes in leaves, roots, flowers, and during *Coffea arabica* L. fruit development

Suzana T. Ivamoto [a, b], Leonardo M. Sakuray [a, b], Lucia P. Ferreira [a], Cíntia S.G. Kitzberger [c], Maria B.S. Scholz [c], David Pot [d], Thierry Leroy [d], Luiz G.E. Vieira [e], Douglas S. Domingues [f], Luiz F.P. Pereira [a, g, *]

[a] Instituto Agronômico do Paraná (IAPAR), Laboratório de Biotecnologia Vegetal, CP 481, CEP 86001-970, Londrina, PR, Brazil
[b] Universidade Estadual de Londrina (UEL), CP 6001, CEP 86051-980, Londrina, PR, Brazil
[c] Instituto Agronômico do Paraná (IAPAR), Laboratório de Ecofisiologia Vegetal, CP 481, CEP 86001-970, Londrina, PR, Brazil
[d] Centre de Coopération Internationale en Recherche Agronomique Pour le Développement, (CIRAD), UMR AGAP, 34398, Montpellier, France
[e] Universidade do Oeste Paulista (UNOESTE), Rodovia Raposo Tavares, Km 572, CEP 19067-175, Presidente Prudente, SP, Brazil
[f] Universidade Estadual Paulista (UNESP), Instituto de Biociências de Rio Claro, Avenida 24-A, 1515, CEP 13506-900, Rio Claro, SP, Brazil
[g] Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA Café), CEP 70770-901, Brasília, DF, Brazil

ARTICLE INFO

Article history:
Received 21 July 2016
Received in revised form
1 December 2016
Accepted 2 December 2016
Available online 3 December 2016

Keywords:
Coffee
Cafestol
Kahweol
Gene expression
RT-qPCR
HPLC

ABSTRACT

Lipids are among the major chemical compounds present in coffee beans, and they affect the flavor and aroma of the coffee beverage. Coffee oil is rich in kaurene diterpene compounds, mainly cafestol (CAF) and kahweol (KAH), which are related to plant defense mechanisms and to nutraceutical and sensorial beverage characteristics. Despite their importance, the final steps of coffee diterpenes biosynthesis remain unknown. To understand the molecular basis of coffee diterpenes biosynthesis, we report the content dynamics of CAF and KAH in several *Coffea arabica* tissues and the transcriptional analysis of cytochrome P450 genes (P450). We measured CAF and KAH concentrations in leaves, roots, flower buds, flowers and fruit tissues at seven developmental stages (30–240 days after flowering - DAF) using HPLC. Higher CAF levels were detected in flower buds and flowers when compared to fruits. In contrast, KAH concentration increased along fruit development, peaking at 120 DAF. We did not detect CAF or KAH in leaves, and higher amounts of KAH than CAF were detected in roots. Using P450 candidate genes from a coffee EST database, we performed RT-qPCR transcriptional analysis of leaves, flowers and fruits at three developmental stages (90, 120 and 150 DAF). Three P450 genes (*CaCYP76C4*, *CaCYP82C2* and *CaCYP74A1*) had transcriptional patterns similar to CAF concentration and two P450 genes (*CaCYP71A25* and *CaCYP701A3*) have transcript accumulation similar to KAH concentration. These data warrant further investigation of these P450s as potential candidate genes involved in the final stages of the CAF and KAH biosynthetic pathways.

1. Introduction

Coffee is appreciated worldwide as a beverage due to its aroma, flavor and stimulant properties. Beverage quality is highly related to the chemical compounds in coffee beans (Oestreich-Janzen, 2010; Scholz et al., 2016). In addition to caffeine, other important components of coffee beans include lipids, sugars, trigonelline, chlorogenic acids and volatile compounds from the plant secondary metabolism (Speer and Kölling-Speer, 2006; Oestreich-Janzen, 2010). A complex combination of these chemicals determines beverage sensory characteristics (Oestreich-Janzen, 2010).

Lipids are mainly found in coffee endosperm, with only a small amount in the fruit outer layer (wax). They represent nearly 15% of the total chemical compounds in *Coffea arabica* and 10% in *Coffea canephora* green beans (Speer and Kölling-Speer, 2006; Oestreich-Janzen, 2010). These compounds occur in beans before and after

* Corresponding author. Embrapa Café, Plant Biotechnology Laboratory Rodovia Celso Garcia Cid, 375 Km, CP: 481, CEP 86001-970, Londrina, Paraná, Brazil.
E-mail address: filipe.pereira@embrapa.br (L.F.P. Pereira).

http://dx.doi.org/10.1016/j.plaphy.2016.12.004
0981-9428/

roasting processes (Oestreich-Janzen, 2010), and they are probably related to tolerance against oxidative stresses caused by temperature variation, acting as membrane stabilizers (DaMatta and Ramalho, 2006).

The diterpenes cafestol (CAF) and kahweol (KAH) are the major components of the unsaponifiable lipid fraction in green and roasted coffee beans (Kitzberger et al., 2013; Scholz et al., 2016). They also affect beverage quality (Del Terra et al., 2013). In addition, they are commonly associated with anti-inflammatory, anticarcinogenic and antioxidant properties in human health (Chu et al., 2011). The industrial potential of these compounds is very high: there are more than 700 patents related to CAF and more than 400 for KAH (October 2016, Google patents survey), including their use in pharmaceuticals, cosmetic products (oils and sunscreens) and soluble coffee flavoring. CAF is also related to volatile phenolic compounds released by flowers, which could be responsible for attracting pollinators (Del Terra et al., 2013). However, CAF is also associated with increased cholesterol levels in the blood plasma (hyperlipidemia) at high ingestion doses, depending on the filter process method used to prepare the coffee drink (Naidoo et al., 2011).

The quantitative analysis of CAF and KAH in *C. arabica* genotypes showed significant intra-species variability in their concentrations, suggesting genetic control of their biosynthesis (Kitzberger et al., 2013). CAF ($C_{20}H_{28}O_3$) and KAH ($C_{20}H_{26}O_3$) have a 20-carbon kaurane skeleton, and their biosynthesis begins from two isoprenoid pathways (Dudareva et al., 2013): i) the cytosolic mevalonate pathway (MVA) and/or ii) the chloroplastic non-mevalonate pathway (MEP). Both pathways produce isopentenyl pyrophosphate (IPP) and dimethylallyl diphosphate (DMAPP), which are used as substrates for geranylgeranyl diphosphate (GGPP) synthesis. GGPP is one of the substrates used to generate terpenoid compounds (Dudareva et al., 2013) and is probably used as a substrate to produce CAF and KAH in coffee plants. Multiple enzymes and substrates for diterpenoid biosynthesis have been described in recent years (Wang et al., 2012a; Dudareva et al., 2013). However, to the best of our knowledge, there is no report identifying genes involved in the final stages of CAF and KAH synthesis/degradation.

Cytochromes P450 (P450s) can recognize and modify the kaurane skeleton leading to diterpenes production (Wang et al., 2012a; Zerbe et al., 2013). P450 genes are also responsible for several secondary metabolites, phytohormones and plant defense compounds (Nelson and Werck-Reichhart, 2011; Wang et al., 2012a; Pateraki et al., 2015).

The chemical compounds present in coffee beans are synthetized during fruit development, which can be partitioned into three stages: initial perisperm development, replacement of the perisperm by the endosperm and fruit ripening (De Castro and Marraccini, 2006). Perisperm development is very important for determining the chemical composition and coffee bean size (De Castro and Marraccini, 2006). CAF and KAH concentrations have mainly been evaluated in ripe fruit and in green or toasted coffee grains (Dias et al., 2010; Kitzberger et al., 2013). Most studies were performed to discriminate commercial roasted and ground coffees according to chemical composition (De Souza and Benassi, 2012). However, there is no study that addresses the accumulation dynamics of CAF and KAH during *C. arabica* fruit development, particularly during their initial growth stages. This information is crucial for addressing the molecular basis of diterpenes synthesis in coffee plants.

In this study, we detail the content dynamics of CAF and KAH during coffee fruit development, as well as their quantification in leaves, roots, flower buds and flowers. In addition, we analyzed the transcriptional profile of seven P450 genes potentially involved in diterpene biosynthetic pathways.

2. Materials and methods

2.1. Plant material

Leaves, roots, flower buds, flowers and fruits of *C. arabica* cv. IAPAR59 were collected from field grown plants at the Agronomic Institute of Paraná (IAPAR) in Londrina (Brazil). Fruit samples were harvested monthly after bloom from 30 to 240 Days After Flowering (DAF). All tissues were immediately frozen in liquid nitrogen after harvesting. Perisperm and endosperm tissues were separated from the pulp, ground with liquid N_2 and stored at −80°C for further diterpenes quantification and RNA extraction. We used 9 biological replicates (tissues from one plant as a single biological replicate) and two technical injections for HPLC quantification analysis. For RT-qPCR analyses, we used bulked plant material in 3 biological replicates, where each replicate was represented by a pool of tissues from 3 plants, and we performed three technical replicates for each biological replicate.

2.2. Cafestol and kahweol quantification by HPLC

To determine CAF and KAH content, we followed the protocol described by Dias et al. (2010) using a reversed-phase Spherisorb ODS 1 column (250 mm × 4.6 mm id 5 mm) (Waters, Milford, USA). A Surveyor Plus high-pressure liquid chromatography instrument (Thermo Scientific, San Jose, USA) was used to determine the diterpenes concentrations. HPLC grade methyl tert-butyl ether (Acros Organics, New Jersey, USA), analytical grade potassium hydroxide KOH (Quimex, São Paulo, Brazil), HPLC grade acetonitrile (J.T. Baker, Xalostoc, Mexico), and kahweol and cafestol standards (Axxora, San Diego, USA) were also used for this analysis. Samples (0.2 g) were saponified at 80°C for 1 h with 2 mL of 2.5 mol L^{-1} potassium hydroxide (96% ethanol, v/v). After adding 2 mL of ultrapure water, the unsaponifiable fraction was extracted with 2 mL of methyl terc-butyl ether, and after agitation and centrifugation (3 min at 3000 rpm), the organic phase was collected. This extraction procedure was repeated three times. Distilled water (2 mL) was added, and the extract was evaporated to dryness in a water bath (70°C). After resuspension in 4.0 mL of the mobile phase (acetonitrile/water 55:45 v/v), the extract was filtered through a 0.45 μm nylon membrane (Millipore, Billerica, USA) and we used a flow rate of 0.9 mL min^{-1}. The detection of cafestol and kahweol was performed at 220 and 290 nm, respectively. An oven temperature of 25°C was applied for 20 min. The identification of the compounds was based on retention time comparisons and co-elution with known standards. All samples were submitted for duplicate extraction and injection. Data were analyzed by two-way ANOVA and Tukey's test ($p < 0.05$) using Assistat software (www.assistat.com).

2.3. Selection of P450 candidate genes

Seven P450 genes (Table 1) were selected from a previous study (Ivamoto et al., 2015) based in the function of putative orthologs and *in silico* expression patterns. The P450 candidate gene sequences were selected from the Brazilian Coffee Expressed Sequenced Project (CafEST) (http://www.lge.ibi.unicamp.br/cafe).

2.4. RNA extraction

The total RNA of *C. arabica* cv. IAPAR59 mature leaves, flowers and fruit perisperm at three developmental stages (90, 120 and 150 DAF) was extracted as described by a previous study from our group (Geromel et al., 2006). Total RNA samples were purified using the Pure Link Micro to Midi Total RNA Purification System (Invitrogen,

Table 1
P450 candidate genes identified and selected by *in silico* approaches.

Coffee P450 name	CafEST	*C. canephora* genome	Putative P450 function	Reference
CaCYP72A15	Contig16992	Cc05_g08890	triterpene biosynthesis	Koo et al., 2011
CaCYP94B1	GW461079	Cc01_g18610	jasmonic acid catabolism	Heitz et al., 2012
CaCYP76C4	Contig146	Cc02_g36410	monoterpenoid biosynthesis	Boachon et al., 2015
CaCYP74A1	Contig7490	Cc10_g03570	lipoxygenase pathway	Zhu et al., 2012
CaCYP82C2	Contig17481	Cc04_g10600	homoterpene biosynthesis	Tholl et al., 2011
CaCYP701A3	Contig11456	Cc10_g03710	ent-kaurene oxidase	Wang et al., 2012a
CaCYP71A25	Contig14459	Cc04_g11300	monoterpenes hidroxylation	Li et al., 2010

Carlsbad, CA, USA) and treated with DNase (Invitrogen, Carlsbad, CA, USA). The RNA integrity was verified in 1% agarose gel electrophoresis, and its concentration and purity were determined using a NanoDrop ND-100 spectrophotometer (Thermo Scientific, San Jose, USA). The absence of genomic DNA contamination was confirmed by polymerase chain reaction (PCR) using glyceraldehyde 3-phosphate dehydrogenase (GAPDH) primers following the protocol described by Cruz et al. (2009) with 100 ng of RNA (Electronic Supplementary Fig. 1). Complementary DNA (cDNA) was synthesized using SuperScript III Reverse Transcriptase (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions with a final volume of 20 μl using 5 μg of total RNA. The final cDNA products were diluted tenfold prior to use in quantitative polymerase chain reaction (RT-qPCR).

2.5. Transcriptional profiles of P450 candidate genes by qPCR

Primers were designed using Primer Express software version 3.0 (Electronic Supplementary Table 1). Primer specificity was verified using dissociation curve analysis and the amplicon length verified in 1% agarose gel electrophoresis. Primer efficiency was calculated using the LinRegPCR software (Ruijter et al., 2009), and only reactions over 90% efficiency were selected for further analysis. The transcriptional profile was analyzed by RT-qPCR in a 7500 Fast Real-Time PCR System (Applied Biosystems, Carlsbad, CA, USA) using SYBR Green PCR Master Mix (Applied Biosystems, Carlsbad, CA, USA). The reaction mixture contained 12.5 μl of SYBR Green PCR Master mix, 10 μM of each primer, 20 ng of cDNA and Milli-Q water to a total volume of 25 μl. Thermal conditions were 95°C for 10 min, followed by 40 cycles of 95°C for 30 s and 60°C for 60 s. Melting curves were analyzed to verify the presence of a single product, with a negative control. All reactions were performed using three technical and biological replicates. Relative expression was calculated as $(1 + E)^{-\Delta\Delta Cq}$, where $\Delta Cq_{target} = Cq_{target\ gene} - Cq_{UBQ}$ and $\Delta\Delta Cq = \Delta Cq_{target} - \Delta Cq_{reference\ sample}$, with polyubiquitin gene (*UBQ*) as the normalizer gene, as recommended for *C. arabica* plants (Cruz et al., 2009). Leaf samples were used for calibration. Data were analyzed by two-way ANOVA and Tukey's test ($p < 0.05$) using the software Assistat (www.assistat.com).

3. Results

3.1. Cafestol and kahweol quantification by HPLC

CAF and KAH profiles were determined by HPLC in the leaves, roots, flowers buds, flowers and fruits (perisperm and endosperm) collected monthly from 30 to 240 days after flowering (DAF; Fig. 1). We did not detect CAF or KAH in *C. arabica* leaves. CAF and KAH concentration patterns in roots and flower buds were contrasting: we observed higher levels of KAH (289 mg/100 g FW) and lower levels of CAF (16.78 mg/100 g FW) in roots, but a higher amount of CAF (263.75 mg/100 g FW) than KAH (8.74 mg/100 g FW) in flower buds. Flowers had the lowest total diterpenes levels (69 mg/100 g FW), with a predominance of CAF over KAH.

In fruits, there was a continued increase of both diterpenes from the beginning of their development (30 DAF) up to 120 DAF. After this phase, total diterpenes concentration started to decline. There was five times more KAH than CAF in *C. arabica* fruits at 120 DAF and 150 DAF.

In the perisperm, KAH concentration (Fig. 2A) increased more than 12-fold from 30 DAF to 120 DAF (81.77–1009.48 mg/100 g FW). After this period, there was a prompt decline, with KAH returning to baseline levels at 210 DAF (14 mg/100 g FW) during the final stages of fruit maturation. Higher levels of KAH were also detected in the endosperm at 120 DAF (721.37 mg/100 g FW), with a slight but steady decrease at the subsequent sample times, followed by an increase at 240 DAF (577.48 mg/100 g FW).

In contrast (Fig. 2B), we did not observe significant differences in CAF content in the perisperm from 30 DAF to 90 DAF. An increase in CAF levels was observed only at 120 DAF, with a sharp decline after this period. In the final stage of fruit development, similarly to the results for KAH, CAF also returned to basal levels in the perisperm. Meanwhile, the endosperm showed an increase in CAF levels from 120 DAF (48.7 mg/100 g FW) to 180 DAF (137.59 mg/100 g FW). After this point, there was a small but significant decrease in the CAF level in the endosperm at 210 DAF, ending with a similar content to the value observed in the previous sampling period (124.24 mg/100 g FW at 240 DAF).

KAH values in endosperm and perisperm were similar to the amount observed in green coffee beans, in a range of 370 to almost 1000 mg/100 g FW (Kitzberger et al., 2013). However, the values observed for CAF were below the ones usually observed (Kitzberger et al., 2013, 2016).

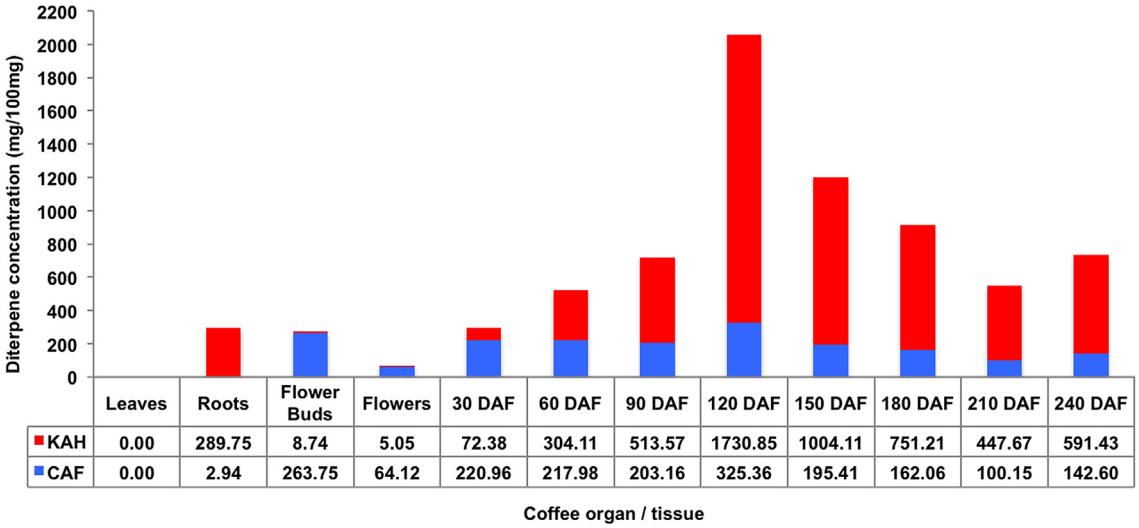
3.2. Transcriptional patterns of P450 candidate genes

Seven P450 candidate genes described as related to secondary metabolite biosynthesis (Ivamoto et al., 2015, Table 1) were selected for transcriptional analysis using RT-qPCR in the leaves, flowers and perisperm at 90, 120 and 150 DAF. These tissues were chosen for further analysis based on displaying the most contrasting CAF and KAH contents. All candidate genes have orthologs in the *C. canephora* genome (Denoeud et al., 2014, Table 1).

Three coffee organs (leaves, flowers and fruits) were chosen for transcriptional profile analysis based on contrasting diterpenes contents (Fig. 1). In addition, we chose three initial fruit development stages (90, 120 and 150 DAF) showing high concentrations of CAF and KAH, especially at 120 DAF (Fig. 1).

Two genes, *CaCYP72A15* and *CaCYP94B1*, showed no expressive differences in transcriptional profile in the leaves and neither in fruit perisperm throughout all sampling times (Fig. 3A and B). These two P450 genes also showed the lowest transcript levels of all genes analyzed in this study.

CaCYP76C4 (Fig. 3C) and *CaCYP74A1* (Fig. 3D) showed similar transcriptional profiles. We observed higher transcriptional activity in flowers compared to leaves, and a gradual increase during fruit



	Leaves	Roots	Flower Buds	Flowers	30 DAF	60 DAF	90 DAF	120 DAF	150 DAF	180 DAF	210 DAF	240 DAF
KAH	0.00	289.75	8.74	5.05	72.38	304.11	513.57	1730.85	1004.11	751.21	447.67	591.43
CAF	0.00	2.94	263.75	64.12	220.96	217.98	203.16	325.36	195.41	162.06	100.15	142.60

Fig. 1. Cafestol and kahweol contents in leaves, roots, flower buds, flowers and fruits (30–240 DAF) in mg/100 g FW. Kahweol content is shown in red and cafestol in blue. Samples are the sum of the perisperm and endosperm. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

perisperm development (from 90 to 150 DAF). There was a significant difference between these two genes in terms of relative expression values, where *CaCYP74A1* transcripts were approximately 10 times more abundant than *CaCYP76C4*.

CaCYP82C2 showed the highest transcript levels in the flowers compared with all other tissues (Fig. 3E). The peak of transcripts in fruit tissues occurred at 150 DAF (1354-fold), similar to what was obtained for *CaCYP76C4* and *CaCYP74A1*.

CaCYP701A3 and *CaCYP71A25* showed higher transcriptional activity in the perisperm than in the leaves and flowers (Fig. 3F and G). For both genes, transcriptional activity peaked in the perisperm at 90 DAF. However, *CaCYP701A3* showed a gradual decrease in transcript abundance from 90 to 150 DAF, while no significant *CaCYP71A25* gene expression was observed in the perisperm at 120and 150 DAF in comparison with 90 DAF.

4. Discussion

Coffee is a popular beverage due to its stimulant characteristics and its flavor. Several metabolites related to those characteristics are synthetized and transported during perisperm development (Geromel et al., 2006; De Castro and Marraccini, 2006; Joët et al., 2009). CAF and KAH are relevant components of coffee lipidic fraction but, despite their importance, very little is known related to their accumulation during fruit development. Most studies have quantified CAF and KAH in green and roasted beans at the cherry stage (Dias et al., 2010; Kitzberger et al., 2013) when fruits are ready to be harvested (around 240 DAF). In our study, CAf and KAH concentration in mature fruits (240 DAF) were similar to previous observations on *C. arabica* green beans, where diterpenes could be roughly divided in 20% of CAF and 80% of KAH (Kitzberger et al., 2013). We also quantified CAF and KAH in several vegetative and reproductive tisues: roots, leaves, flower buds, flowers, perisperm and endosperm during fruit development (30–240 DAF) to address the accumulation of these compounds along the fruit maturation process. Moreover, five P450 genes had transcriptional profiles with similar patterns to CAF and KAH content, as observed by HPLC quantification.

In roots, KAH is present in higher concentration than CAF, with concentration similar to the observed in perisperm (60 DAF). This is the first report on the occurrence of CAF and KAH in coffee root tissues. In *Arabidopsis*, the synthesis in roots of the semi volatile diterpene rhizathelene is involved in plant herbivory against root-feeding insects (Vaughan et al., 2013). In rice, diterpenoid phytoalexins are produced in the roots in response to *Magnaporthe oryzae* attack (Yamamura et al., 2015). Altogether, these results raises questions on the importance of CAF and KAH accumulation in coffee root, where we can speculate their participation in herbivore defense or microbial interactions.

CAF levels were higher in flower buds and flower organs than in fruit tissues except at 120 DAF. The presence of high levels of CAF in flower buds and flowers might be attributed to pollinators floral attractiveness (Farré-Armengol et al., 2015). This hypothesis is based on a previous study showing the grindelic acid accumulation profile, an important diterpenoid compound also produced in high levels by flowers from *Grindelia robusta* plants (Zerbe et al., 2015).

KAH content was lower in floral buds and flower tissues but considerably higher than CAF during the fruit developmental stages, peaking at 120 DAF. Interestingly, KAH was the main diterpene observed during fruit maturation but had very low level in flower and flower buds. Therefore, it is possible that this diterpene is not related to attraction of insects. Based on our CAF and KAH profiles among organs and tissues, we can also speculate that CAF is synthetized firstly, and then catalyzed into KAH, but this hypothesis needs to be further investigated.

Both CAF and KAH had their highest concentration at 120 DAF (Fig. 1) in fruits (perisperm and endosperm), which may suggests that diterpenes are mostly synthetized in intermediary rather than in the initial or final stages of fruit development. Perisperm was previously described as an important tissue contributing to the chemical composition of coffee beans (De Castro and Marraccini, 2006; Joët et al., 2009) and it showed a higher content of CAF and KAH than endosperm. Interestingly, the perisperm undergoes intense cell division and expansion around 120 DAF, which is responsible for defining the final coffee grain size (De Castro and Marraccini, 2006). We observed that the contents of these diterpenes began to decrease in perisperm after 120 DAF and are present in endosperm from this period until the end of fruit maturation (240 DAF). During 120 to 240 DAF the perisperm is replaced by the endosperm in coffee fruits (De Castro and Marraccini, 2006). The decrease of CAF and KAH in the perisperm may be due to the disappearance of this tissue and translocation of those compounds

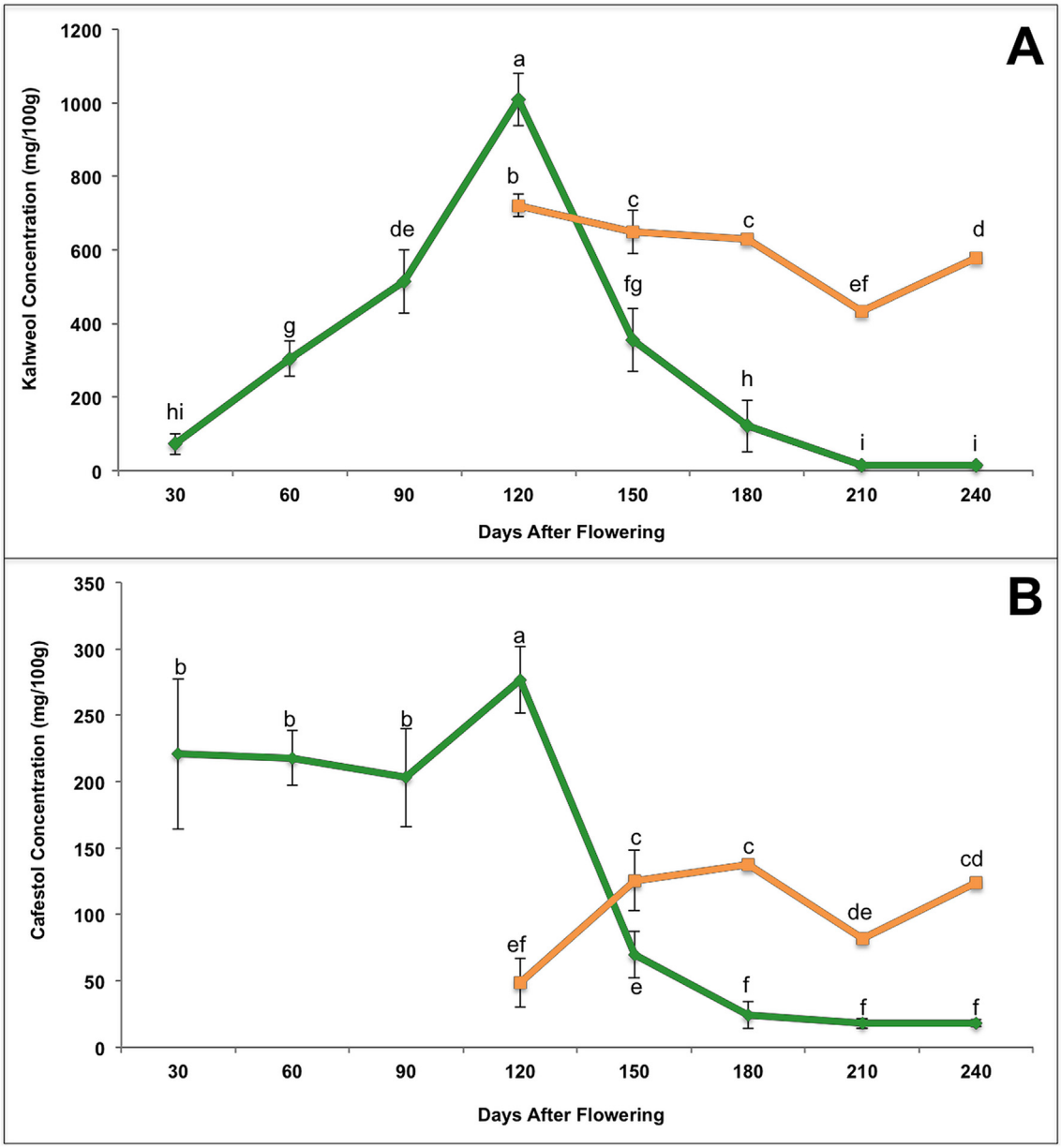


Fig. 2. Quantification of kahweol (A) and cafestol (B) content during fruit development in perisperm (green line) and endosperm (orange line) by HPLC. Fruits do not have any endosperm tissue prior to 120 DAF. Error bars represent range for replicate HPLC quantifications. Lower-case letters, from *a* to *i*, represent significant differences in the mean relative CAF and KAH values using two-way ANOVA and Tukey's test ($p < 0.05$). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

to the endosperm, similarly to what was observed in sucrose (Geromel et al., 2006) and chlorogenic acids (Joët et al., 2009) accumulation patterns.

Several studies have been describing the initial steps of diterpenes biosynthesis (Wang et al., 2012a; Zerbe et al., 2013; Mafu et al., 2016), where the precursor substrate is the isopentenyl diphosphate (IPP) molecule formed by MVA (6 steps) and MEP (7 steps) isoprenoid metabolic pathway (Dudareva et al., 2013). IPP is used as substrate to generate geranylgeranyl diphosphate (GGPP), thereafter a copalyl diphosphate synthase (CPS) catalyze the cyclization of GGPP to *ent*-copalyl diphosphate (*ent*-CPP), which is converted to *ent*-kaurene by kaurene synthase (KS) in several plants (Wang et al., 2012a; Zerbe et al., 2013; Mafu et al., 2016). Since CAF and KAH are probable derivatives of *ent*-kaurene, and there are several reports attributing to P450s the stereospecific hydroxylation and oxidation of hydroxyl groups that leads to the synthesis of several diterpenes (Zerbe et al., 2013; Pateraki et al., 2015), we hypothesized that the transcriptional evaluation of selected P450 candidate genes could give us insights on the final steps from *ent*-kaurene to CAF and KAH biosynthesis. Therefore, to identify the best P450s candidate genes possibly related to the final steps of CAF and KAH formation we combined our biochemical data with the P450s transcriptional profile obtained from a previous *in silico* analysis (Ivamoto et al., 2015). Based on CAF and KAH concentrations, we selected five contrasting samples for diterpenes accumulation (leaves, flowers and perisperm at 90, 120 and 150 DAF) and five P450s candidate genes for the transcriptional analysis by RT-qPCR.

CaCYP72A15 and *CaCYP94B1* showed no expressive differences in their transcriptional profile among coffee tissues. Several

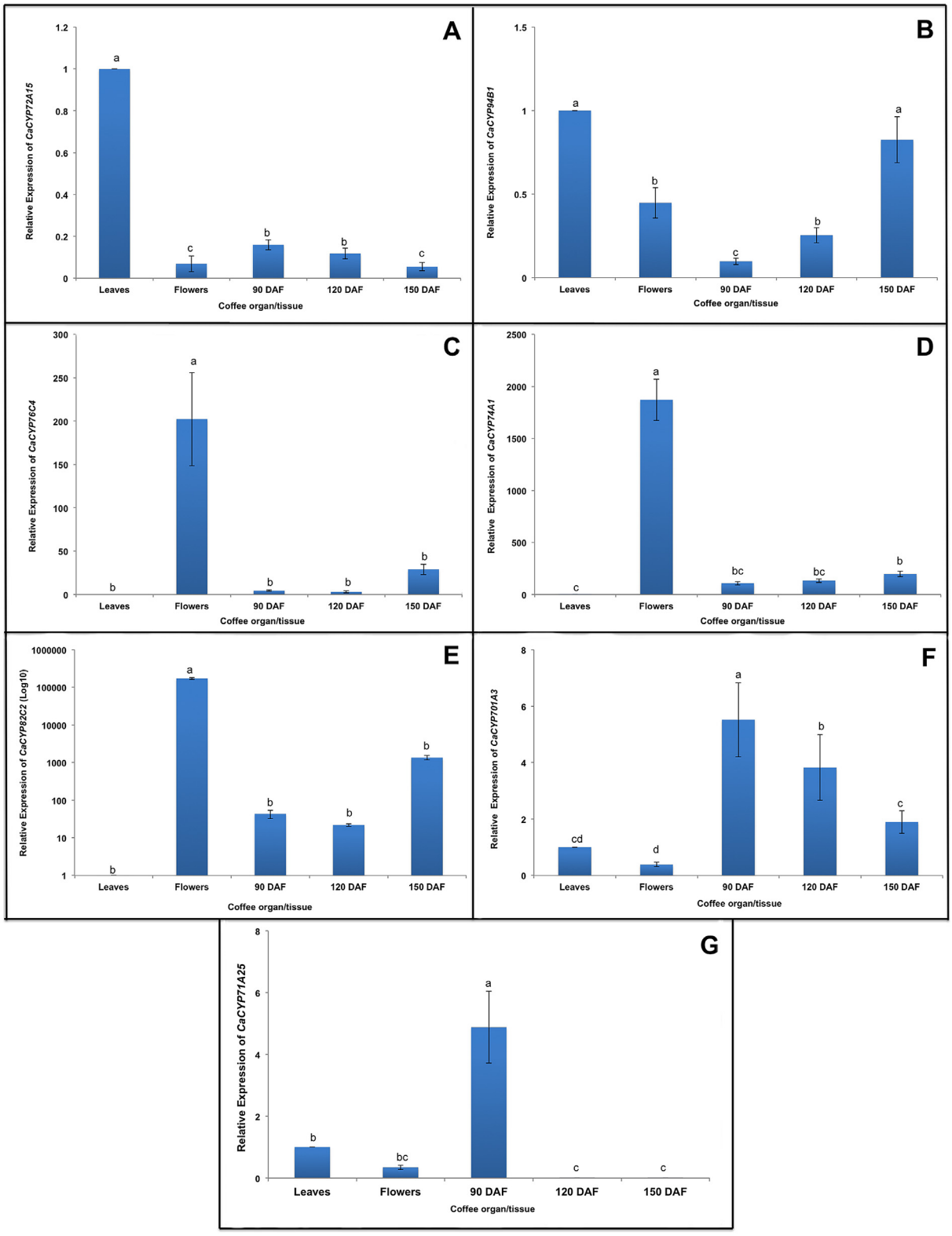


Fig. 3. Transcriptional profiles of *C. arabica* P450 candidate genes in leaves, flowers, and in the perisperm at three developmental stages (90–150 DAF). A: *CaCYP72A15*. B: *CaCYP94B1*. C: *CaCYP76C4*. D: *CaCYP74A1*. E: *CaCYP82C2*. F: *CaCYP701A3*. G: *CaCYP71A25*. Error bars represent range for replicate RT-qPCR reactions. Lower-case letters, from *a* to *d*, indicate significant differences in the mean relative expression values using two-way ANOVA and Tukey's test ($p < 0.05$).

CYP72As are involved in triterpenoid biosynthesis in plants from the order Fabales, but less than 25% of this gene family has been characterized biochemically (Prall et al., 2016). *CaCYP94B1* was reported to be involved in the biosynthesis of the phytohormone jasmonoyl-L-isoleucine (JA-Ile) and in plant development in rice (Koo et al., 2011). In *Arabidopsis*, this P450 subfamily catalyzes two successive oxidation steps of plant hormone JA-Ile turnover (Heitz et al., 2012). *CYP94B3* performed efficiently the initial hydroxylation

of JA-Ile to 12OH-JA-Ile, with little conversion to 12COOH-JA-Ile, whereas *CYP94C1* catalyzed preferentially carboxy-derivative formation (Heitz et al., 2012). Therefore, *CaCYP94B1* is more likely to be involved in jasmonate-turnover mechanisms rather than CAF and KAH biosynthesis.

In this study, *CaCYP76C4* was highly expressed in flowers, where CAF was also detected in large amounts. The *CYP76* subfamilies are involved in diterpenoid biosynthesis described in rice (Wang et al., 2012b). *OsCYP76M8* are able to catalyze C7-hydroxylation of *ent*-kaurene to produce 7α-hydroxyl-*ent*-isokaurene and three CYP76 (*OsCYP76M5, OsCYP76M6* and *OsCYP76M8*) were described to be involved in the first step of *ent*-sandaracopimaradiene catalysis into oryzalexins A-D (Wang et al., 2012b), which are diterpenoid phytoalexins. The preferential expression in flowers of several *CYP76* genes has also been reported in *Arabidopsis* (Höfer et al., 2014). As CAF is a volatile diterpene, this gene deserves further investigation as a candidate for CAF production.

The high expression of *CaCYP74A1* in coffee flowers compared with other tissues also suggest that this gene is related to CAF synthesis. P450s in the *CYP74A* subfamily have also been reported to be allene oxide synthase (AOS), which commits 13-hydroperoxy linolenic acid (13-HPOT) to the formation of plant defense hormone, jasmonic acid and hydroperoxide lyase (HPL), which converts 13-HPOT into 6-carbon aldehydes and 12-carbon ω-ketofatty acids (Guttikonda et al., 2010). In grape, they were also described as hydroperoxide lyase (Zhu et al., 2012) and being related to plant defense hormone biosynthesis (jasmonic acid) and volatile compounds.

CaCYP82C2 also showed remarkable expression in flowers compared with the other organs and tissues. The *CYP82* subfamily is described in Arabidopsis to be involved in homoterpene biosynthesis, with emissions from flowers and from vegetative tissues upon herbivore feeding (Tholl et al., 2011). In our case, *CaCYP82C2* was 14,000-fold more expressed in flowers than in leaves. Although these values seemed unusual, there are studies showing similar transcriptional profiles in *Brassica napus* and *Curcurbita maxima* (Pant et al., 2008). This gene was also expressed in the perisperm, in contrast to *CaCYP76C4* and *CaCYP74A1*, particularly at 150 DAF. These three genes warrant further investigation as putative genes involved in CAF and KAH biosynthesis, as they were not expressed in leaves.

CaCYP701A3 and *CaCYP71A25*, showed higher transcriptional values in the perisperm than in the flowers and leaves. *CaCYP701A3* was identified as the *ent*-kaurene oxidase (KO) enzyme. KO has already been established as responsible for three catalysis reactions in the early stages of gibberellin (GA) biosynthesis in rice (Wang et al., 2012a). The substrates described as GA precursor compounds, geranylgeranyl diphosphate (GGPP), *ent*-copalil diphosphate (CDP), *ent*-kaurene and *ent*-kaurenoic acid, are also used in the middle steps of *ent*-kaurene diterpene biosynthesis (Mafu et al., 2016).

In contrast to *CaCYP701A3*, the *CaCYP71A25* gene was not expressed during the late stages of fruit development. Described as the largest P450 subfamily, the *CYP71* genes are associated with several functions (Nelson and Werck-Reichhart, 2011). In tomato, high expression of *LsCYP71A2,* a gene related to the hydroxylation of monoterpenes, was detected during the early stages of tomato fruit development (Li et al., 2010). This gene is also involved in fruit development and uses isoprenoid substrates (Li et al., 2010). As the P450s of the *CYP71* subfamily are described as promiscuous enzymes, *CaCYP71A25* could be involved in CAF and KAH formation rather than monoterpenes, as observed in tomato. Based on the higher concentration of KAH in fruits when compared to flowers, *CaCYP71A25* might be involved in the middle steps of KAH biosynthesis, rather to CAF formation.

Overall, this study is a starting point to improve our knowledge of diterpenes accumulation pattern in coffee organs and tissues. The identification of KAH and CAF in the flowers and roots indicates that they are also important in plant interaction with other organisms. This report is the first to combine biochemical and transcriptional analysis to identify candidate genes involved in the middle steps of CAF and KAH biosynthesis. The results allowed us to identify five genes possibly related to the production of these diterpenes: *CaCYP71A25*, *CaCYP701A3*, *CaCYP76C4*, *CaCYP82C2* and *CaCYP74A1*. These genes presented similar transcriptional patterns to the CAF and KAH accumulation in coffee organs and tissues.

5. Conclusion

Our study is the first to point out the importance of perisperm development in the accumulation of CAF and KAH in coffee seeds. We also report five P450 candidate genes with transcriptional activity patterns closely related to the accumulation of CAF and KAH in different organs and tissues. Since P450s is a multigene family with hundreds of paralogs, we think that our results pave out further studies to test the role of specific P450s in the CAF and KAH accumulation/degradation processes.

Author's contribution statement

STI: *in silico* analysis and data interpretation; HPLC and RT-qPCR analyses; conception and writing of the manuscript; LMS, LPF, CSGK and MBSS: HPLC analysis contribution; DP and TL: revision of the final manuscript; LGEV: project coordination and revision of the manuscript; DSD: *in silico* analysis contribution and revision of the final manuscript; LFPP: project coordination; writing, elaboration and revision of the final manuscript; All authors read and approved the final manuscript.

Conflicts of interest

The authors declare that they have no conflict of interest.

Acknowledgements

We thank to the Brazilian Coffee Research Consortium (Grant #19.2004.576.00), INCT Café, Coordination for the Improvement of Higher Education Personnel (CAPES) and Brazilian Innovation Agency (FINEP– Grant #01.05.0665-00) for supporting this study. STI, LMS and LPF acknowledge National Council of Technological and Scientific Development (CNPq) and CNPq Science without borders (CNPq/CsF), CAPES and Embrapa Café for student fellowship. DSD, LGEV and LFPP are CNPq research fellows.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.plaphy.2016.12.004.

References

Boachon, B., Junker, R.R., Miesch, L., et al., 2015. CYP76C1 (Cytochrome P450)-Mediated linalool metabolism and the formation of volatile and soluble linalool oxides in *Arabidopsis* flowers: a strategy for defense against floral antagonists. Plant Cell 27 (10), 2972–2990.

Chu, Y.-F., Chen, Y., Black, R.M., et al., 2011. Type 2 diabetes-related bioactivities of coffee: assessment of antioxidant activity, NF-kB inhibition, and stimulating of glucose uptake. Food Chem. 124, 914–920.

Cruz, F., Kalaoun, S., Nobile, P., et al., 2009. Evaluation of coffee reference genes for relative expression studies by quantitative real-time RT-PCR. Mol. Breed. 23 (4), 607–616.

DaMatta, F.M., Ramalho, J.D., 2006. Impacts of drought and temperature stress on

coffee physiology and production: a review. Braz. J. Plant Physiol. 18 (1), 55–81.
De Castro, R.D., Marraccini, P., 2006. Cytology, biochemistry and molecular changes during coffee fruit development. Braz J. Plant Physiol. 18, 1–25.
Del Terra, L., Lonzarich, V., Asquini, E., et al., 2013. Functional characterization of three *Coffea arabica* L. monoterpene synthases: insights into the enzymatic machinery of coffee aroma. Phytochem 89, 6–14.
Denoeud, F., Carretero-Paulet, L., Dereeper, A., et al., 2014. The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. Science 345 (6201), 1181–1184.
De Souza, R.M.N., Benassi, M.T., 2012. Discrimination of commercial roasted and ground coffees according to chemical composition. J. Braz. Chem. Soc. 23, 1347–1354.
Dias, R.C.E., Campanha, F.G., Vieira, L.G.E., et al., 2010. Evaluation of kahweol and cafestol in *Coffee* tissues and roasted coffee by a new hight-performance liquid chromatography methodology. J. Agric. Food Chem. 58, 88–93.
Dudareva, N., Klempien, A., Muhlemann, J.K., et al., 2013. Biosynthesis, function and metabolic engineering of plant volatile organic compounds. New Phytol. 198 (1), 16–32.
Farré-Armengol, G., Filella, I., Llusià, J., et al., 2015. Pollination mode determines floral scent. Biochem. Syst. Ecol. 61, 44–53.
Geromel, C., Ferreira, L.P., Guerreiro, S.M.C., et al., 2006. Biochemical and genomic analysis of sucrose metabolism during coffee (Coffea arabica) fruit development. J. Exp. Bot. 57 (12), 3243–3258.
Guttikonda, S.K., Trupti, J., Bisht, N.C., 2010. Whole genome co-expression analysis of soybean cytochrome P450 genes identifies nodulation-specific P450 monooxygenases. BMC Plant Biol. 10 (1), 243.
Heitz, T., Widemann, E., Lugan, R., et al., 2012. Cytochromes P450 CYP94C1 and CYP94B3 catalyze two successive oxidation steps of plant hormone Jasmonoyl-isoleucine for catabolic turnover. J.Biol. Chem. 287, 6296–6306.
Höfer, R., Boachon, B., Renault, H., et al., 2014. Dual function of the cytochrome P450 CYP76 family from *Arabidopsis thaliana* in the metabolism of monoterpenols and phenylurea herbicides. Plant Physiol. 166 (3), 1149–1161.
Ivamoto, S.T., Domingues, D.S., Vieira, L.G.E., et al., 2015. Depicting the transcriptionally active cytochromes P450s in *Coffea* spp. Genet. Mol. Res. 14 (1), 2399–2412.
Joët, T., Laffargue, A., Salmona, J., et al., 2009. Metabolic pathways in tropical dicotyledonous albuminous seeds: *Coffea arabica* as a case study. New Phytol. 182 (1), 146–162.
Kitzberger, C.S.G., Scholz, M.B.S., Pereira, L.F.P., et al., 2013. Diterpenes in green and roasted coffee of *Coffea arabica* cultivars growing in the same edapho-climatic conditions. J. Food Compos 30, 52–57.
Kitzberger, C.S.G., Scholz, M.B.S., Pereira, L.F.P., et al., 2016. Profile of the diterpenes lipid and protein content of different coffee cultivars of three consecutive harvests. Google. AIMS Agric. Food 1 (3), 254–264.
Koo, A.J.K., Cooke, T.F., Howe, G.A., et al., 2011. Cytochrome P450 CYP94B3 mediates catabolism and inactivation of the plant hormone jasmonoyl-L-isoleucine. P. Natl. Acad. Sci. U. S. A. 108 (22), 9298–9303.
Li, Z., Hao, T., Yang, Y., 2010. Molecular cloning and expression analysis of a cytochrome P450 gene in tomato. Plant Grown Regul. 61, 297–304.
Mafu, S., Jia, M., Zi, J., et al., 2016. Probing the promiscuity of *ent*-kaurene oxidases via combinatorial biosynthesis. P. Natl. Acad. Sci. U. S. A. 113 (9), 2526–2531.
Naidoo, N., Chen, C., Rebello, S., 2011. Cholesterol-raising diterpenes in types of coffee commonly consumed in Singapore, Indonesia and India and associations with blood lipids: a survey and cross sectional study. Nutr. J. 48, 1–10.
Nelson, D., Werck-Reichhart, D., 2011. A P450-centric view of plant evolution. Plant J. 66, 194–211.
Oestreich-Janzen, S., 2010. Chemistry of coffee. Chem. Biol. 3, 1085–1117.
Pant, B.D., Buhtz, A., Kehr, J., 2008. MicroRNA399 is a long-distance signal for the regulation of plant phosphate homeostasis. Plant J. 53, 731–738.
Pateraki, I., Heskes, A.M., Hamberger, B., 2015. Cytochromes P450 for terpene functionalisation and metabolic engineering. In: Biotechnology of Isoprenoids. Springer International Publishing, pp. 107–139.
Prall, W., Hendy, O., Thornton, L.E., 2016. Utility of a phylogenetic perspective in structural analysis of CYP72A enzymes from flowering plants. PloS one 11 (9), e0163024.
Ruijter, J.M., Ramakers, C., Hoogaars, W.M.H., et al., 2009. Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. Nucleic Acid. Res. 37 (6), e45.
Scholz, M.B.S., Pagiatto, N.F., Kitzberger, C.S.G., et al., 2016. Chemical composition in wild ethiopian Arabica coffee accessions. Euphytica 1–10.
Speer, K., Kölling-Speer, I., 2006. The lipid fraction of the coffee bean. Braz. J. Plant Physiol. 18, 201–216.
Tholl, D., Sohrabi, R., Huh, J., 2011. The biochemistry of homoterpenes - common constitutes of floral and herbivore-induced plant volatile bouquets. Phytochem 72, 1635–1646.
Vaughan, M.M., Wang, Q., Webster, F.X., et al., 2013. Formation of the unusual semivolatile diterpene rhizathalene by the Arabidopsis class I terpene synthase TPS08 in the root stele is involved in defense against belowground herbivory. Plant Cell 25 (3), 1108–1125.
Wang, Q., Hillwig, M.L., Wu, Y., 2012a. CYP701A8: a rice *ent*-kaurene oxidase paralog diverted to more specialized diterpenoid metabolism. Plant Physiol. 158, 1418–1425.
Wang, Q., Hillwig, M.L., Okada, K., et al., 2012b. Characterization of CYP76M5–8 indicates metabolic plasticity within a plant biosynthetic gene cluster. J. Biol. Chem. 287 (9), 6159–6168.
Yamamura, C., Mizutani, E., Okada, K., et al., 2015. Diterpenoid phytoalexin factor, a bHLH transcription factor, plays a central role in the biosynthesis of diterpenoid phytoalexins in rice. Plant J. 84 (6), 1100–1113.
Zerbe, P., Hamberger, B., Yuen, M.M.S., et al., 2013. Gene discovery of modular diterpene metabolism in nonmodel systems. Plant Physiol. 162, 1073–1091.
Zerbe, P., Rodriguez, S.M., Mafu, S., et al., 2015. Exploring diterpene metabolism in non-model species: transcriptome-enabled discovery and functional characterization of labda-7, 13E-dienyl diphosphate synthase from *Grindelia robusta*. Plant J. 83 (5), 783–793.
Zhu, B.-Q., Xu, X.-Q., Wu, Y.-W., 2012. Isolation and characterization of two hydroperoxide lyase genes from grape berries. Mol. Biol. Rep. 39, 7443–7455.