Targeted metabolomics and transcript profiling of methyltransferases in three coffee species

Andrea Montis\textsuperscript{a,b}, Cédric Delporte\textsuperscript{a,b}, Yusaku Noda\textsuperscript{c}, Piet Stoffelen\textsuperscript{d}, Caroline Stévigny\textsuperscript{a}, Christian Hermans\textsuperscript{e}, Pierre Van Antwerpen\textsuperscript{a,b,*,}\textsuperscript{f}, Florence Souard\textsuperscript{g,h}

\textsuperscript{a} RD3 Unit of Pharmacognosy, Bioanalysis and Drug Discovery, Faculty of Pharmacy, Université libre de Bruxelles, Campus Plaine, CP 205/05, Brussels 1050, Belgium
\textsuperscript{b} APFP Analytical platform of the Faculty of Pharmacy, Faculty of Pharmacy, Université libre de Bruxelles, Campus Plaine, CP 205/5, Brussels 1050, Belgium
\textsuperscript{c} The National Institutes for Quantum Science and Technology (QST), Takasaki Institute for Advanced Quantum Science, Gunma, 370-1292, Japan
\textsuperscript{d} Meise Botanic Garden, Domain van Bouchout, Nieuwe laan 38, Meise 1860, Belgium
\textsuperscript{e} Crop Production and Biostimulation Laboratory, Brussels Bioengineering School, Université libre de Bruxelles, Campus Plaine, CP 245, Brussels 1050, Belgium
\textsuperscript{f} Département de Pharmacochimie Moléculaire, UMR 5063 CNRS, Université Grenoble Alpes, 470 rue de la chimie, Saint-Martin d'Hères 38400, France
\textsuperscript{g} DPP Department – Unit of Pharmacology, Pharmacotherapy and Pharmaceutical care, Faculty of Pharmacy, Université libre de Bruxelles, Campus Plaine, CP 205/07, Brussels 1050, Belgium

\textbf{A P I C L E I N F O}

\textbf{Keywords:}
Coffee
Xanthines
O- and N-methyltransferases
LC-(HR)MS
Metabolomics
Transcriptomics

\textbf{A B S T R A C T}

Coffee plants contain well-known xanthines as caffeine. Three \textit{Coffee} species grown in a controlled greenhouse environment were the focus of this research. \textit{Coffee arabica} and \textit{C. canephora} are two first principal commercial species and commonly known as arabica and robusta, respectively. Originating in Central Africa, \textit{C. anthonyi} is a novel species with small leaves. The xanthine metabolites in flower, fruit and leaf extracts were compared using both targeted and untargeted metabolomics approaches. We evaluated how the xanthine derivatives and FQA isomers relate to the expression of biosynthetic genes encoding N- and O-methyltransferases. Theobromine built up in leaves of \textit{C. anthonyi} because caffeine biosynthesis was hindered in the absence of synthase gene expression. Despite this, green fruits expressed these genes and they produced caffeine. Given that \textit{C. anthonyi} evolved successfully over time, these findings put into question the defensive role of caffeine in leaves. An overview of the histolocalisation of xanthines in the different flower parts of \textit{Coffee arabica} was also provided. The gynoecium contained more theobromine than the flower buds or petals. This could be attributed to increased caffeine biosynthesis before fructification. The presence of theophylline and the absence of theobromine in the petals indicate that caffeine is catabolized more in the petals than in the gynoecium.

1. Introduction

Many plant species contain alkaloids, a large family of secondary metabolites with a nitrogen-containing heterocyclic ring (Dey et al., 2020). Caffeine (1,2,3-trimethylxanthine) is a purine alkaloid that plays a role in plant protection. It accounts for one or two percents dry weight of coffee (Ashihara et al., 2008). This molecule is synthesized from xanthosine in a four-step pathway involving three methylation reactions and one ribose hydrolysis (Fig. 1). The methylation steps are sequentially catalysed by the xanthosine 7-methyltransferase (XMT), 7-methylxanthine methyltransferase/theobromine synthase (MXMT) and 3, 7-dimethylxanthine methyltransferase/caffeine synthase (DXMT). The coenzyme involved in these methyl group transfers is the S-Adenosyl Methionine (SAM). In coffee plants, different isoforms with dual functions are characterized: MXMT1 and MXMT2/CTS2 are essentially involved in theobromine biosynthesis, while DXMT1 and DXMT2/CCS1 in both theobromine and caffeine biosynthesis (Mizuno et al., 2003; Uefuji et al., 2003) (Fig. 1).

Caffeine is primarily synthesized in the young leaves and developing endosperms of fruits (Zheng and Ashihara, 2004). Every tissue expresses \textit{DXMT2} and \textit{MXMT2}, while flower buds lowly express \textit{DXMT2} (Mizuno et al., 2003). In \textit{Coffee arabica} (Arabica), transcripts of \textit{DXMT1} are abundant during fruit development, while those of \textit{MXMT1} low (Perrois et al., 2015). Mature leaves contain lower \textit{DXMT1} transcripts compared to younger ones. In \textit{C. canephora} (Robusta), the expression of the same genes is showing a different pattern (Perrois et al., 2015). For instance,
Caffeine is catabolised into theophylline in mature tissues, after being synthesized in young beans and leaves. The catabolic pathway involves a multistep reaction wherein theophylline is produced after one demethylation (Fig. 1). That catabolite loses the two remaining methyl groups and is subsequently converted into 3-methylxanthine and xanthine (Ashihara et al., 2008). An alternative catabolic pathway has been proposed in certain Coffea species, including *abekukatue*, *deweverei* and *liberica*. The mature leaves of these species lose caffeine in favour of theacrine (4-methylxanthine), which is then converted into liberine (Petermann and Baumann, 1983).

Most studies on the expression of genes involved in caffeine biosynthesis have been conducted in *C. arabica* and *C. canephora*. Indeed, these two major species are commercially cultivated for beverage production (Montis et al., 2021). Other species have received less attention. This is the case for *C. anthonyi*, an endemic to Cameroon, Congo and Gabon. It closely relates to *C. eugenioides*, the maternal progenitor of *C. arabica* (Stoffelen et al., 2009), while *C. canephora* is the paternal progenitor of the latter one (Bawin et al., 2021). Furthermore, it grows in lower altitude and could be an alternative to the two other species in the context of the climate change and temperature increase.

Besides, the bitterness of coffee is attributed to the breakdown of chlorogenic acids (CGAs) into phenolic compounds during roasting. The polyphenolic fraction of coffee tissues is primarily composed of cafefoylquinic acids (CQAs), feruloylquinic acids (FQAs) and coumaroylquinic acids (3-CoQA), in monomeric or dimeric forms (Chen, 2019).

To identify the metabolic pathways of *C. arabica* and *C. canephora* in leaves, we compared specifically the metabolomes of *C. canephora* and two *C. arabica* genotypes, unlike the two other ones. Moreover, CQA was present in the genotype, unlike the two other ones. Moreover, CQA was present in *C. anthonyi*, while caffeine was absent from leaves and phloem sap in *C. arabica*.

Indeed, these observations with the expression of genes encoding N- and O-methyltransferases, both members of the methyltransferase super-family. In addition, we want to clarify whether the caffeine accumulation in the fruits of *C. arabica* is primarily due to phloem translocation from leaves or to efficient biosynthetic activity in different flower parts.

### 2. Materials and methods

#### 2.1. Chemicals and reagents

Ultra-high purity water was obtained by filtration using a Milli-Q system from Millipore (Bedford, MA, USA). MS quality acetonitrile, formic acid (FA) and trifluoroacetic acid (TFA) were purchased from Fisher Scientific, while caffeine (> 99 %), theobromine (> 98 %) and theophylline (> 99 %) were purchased from Sigma-Aldrich (Steinheim, Germany).

#### 2.2. Plant material

The coffee trees were grown in the tropical greenhouses of the Meise Botanic Garden (Meise, Belgium) with the same environmental conditions: natural daylight, temperature > 20 °C and relative humidity. The codes of coffee trees were archived, and vouchers deposited in the Herbarium of the Garden. All samples were harvested in the morning between 9 am and 12 am. Some 30 samples of leaves from *C. anthonyi*, *C. arabica* and *C. canephora* were collected. Two developmental stages were distinguished: L1 consisted of young leaves, and L2 of mature leaves. The first ones were recently expanded with light green shade (Salgado et al., 2008) and they measured respectively 2 cm wide and 3 cm long in *C. arabica* and *C. anthonyi*, and 6 cm wide and 10 cm long in *C. canephora*. The second ones were fully developed without marginal necrosis (Souard et al., 2018), and they measured respectively 6 cm wide and 12 cm long, and 10 cm wide and 18 cm long. Ten samples of fruits were picked from *C. anthonyi* and *C. arabica*. Flower parts were also collected from *C. arabica* before fruitication: six buds from one plant, 18 gynoecia and 30 petals from five different plants. For metabolomic analysis, only the mature leaves (L2) and the ripened fruits (RF) were considered. For RNA extraction and expression study, the same organs, the young leaves (L1) and the green fruits (GF) were harvested.

---

Fig. 1. The major biosynthetic pathway of caffeine in coffee plants. The pathway shows metabolites and genes encoding enzymes investigated in this work. (1) 7-methylxanthosine synthase; (2) N-methylxanthosine synthase; (3) theobromine synthase (enzymatic activity exerted by DXMT1, DXMT2, MXMT1 and MXMT2); (4) caffeine synthase (enzymatic activity exerted by DXMT1 and DXMT2). S-adenosyl-L-methionine (SAM) is converted into 5-adenosyl-L-homocysteine (SAH) during steps catalysed by N-methyltransferases. The ribose unit (Rib) is lost upon conversion of 7-Methylxanthosine into 7-Methylxanthine. Adapted from (Montis et al., 2021).

---

MXMT1 is not expressed in young leaves, while DXMT1/2 is more in that species. Fig. 1. The major biosynthetic pathway of caffeine in coffee plants. The pathway shows metabolites and genes encoding enzymes investigated in this work. (1) 7-methylxanthosine synthase; (2) N-methylxanthosine synthase; (3) theobromine synthase (enzymatic activity exerted by DXMT1, DXMT2, MXMT1 and MXMT2); (4) caffeine synthase (enzymatic activity exerted by DXMT1 and DXMT2). S-adenosyl-L-methionine (SAM) is converted into 5-adenosyl-L-homocysteine (SAH) during steps catalysed by N-methyltransferases. The ribose unit (Rib) is lost upon conversion of 7-Methylxanthosine into 7-Methylxanthine. Adapted from (Montis et al., 2021).
2.3 Sample preparation and storage conditions

Samples for metabolomic analysis were dried after collection in sealed plastic bags filled with silica gel, which was replaced daily for seven days. Then, the leaves and flower petals were powdered with an ULTRA-TURRAX® Tube Drive control (Q-Lab, Vilvoorde, Belgium) during 10 min at 6 102 min-1. The gynoecia and flower buds were first grinded with a mortar and pestle, and submitted the same treatment during 20 min. The extraction consisted of suspending 15 mg of homogenized sample powder in 1.5 mL of milliQ water, and placing for 5 min in 55 kHz ultrasonic bath (Souard et al., 2018). Three to eight extraction replicates were obtained. Samples were filtered through a cellulose acetate membrane (0.2 μm) and stored at −20 °C until proceeding to analysis. Fresh plant tissues for RNA extraction were flash-frozen liquid nitrogen and stored at −80 °C.

2.4. LC-HRMS analysis

Semi-polar metabolite fingerprints were monitored using a 1200 series rapid resolution LC (RRLC) system coupled to a 6520 series electrospray ionization (ESI)-quadrupole time-of-flight (QTOF) high-resolution mass spectrometer (HRMS) from Agilent Technologies (Waldbronn, Germany). A Poroshell 120 EC-C18 column (2.7 μm, 100 mm × 2.1 mm) from the same supplier was used to carry out the chromatographic separation. The column temperature was set to 55.0 °C. The mobile phases were composed of 0.025 % of TFA and 0.075 % of FA in water (solvent A), and in acetonitrile (solvent B) for positive ion mode analysis. The same LC-MS conditions were applied to the negative ionisation mode, by slightly changing the composition of the solvents. Solvent A was 20 mM ammonium formate pH 5.5 and solvent B acetonitrile. The flow rate was set to 0.5 mL min-1 and the applied gradient was: 0 min, 0 % B; 0–8 min, 0–10 % B; 8–9 min, 10–12.5 % B; 9–11 min, 12.5–15 % B; 11–17 min, 15–80 % B; 17–18 min, 80–100 % B; 18–19 min, 100 % B; 19–20 min, 100–0 % B; post-run 8 min at 0.5 mL min-1. ESI-QTOF parameters are detailed in supplemental data 1.

2.5. Quantification of caffeine, theobromine and theophylline

Mass Hunter Quantitative Analysis software version B.10 (Agilent Technologies) was used to determine the concentration of caffeine, theobromine and theophylline in the extracts. Caffeine concentration was determined by interpolation of caffeine calibration curve (y = −0.01141 x² + 606.5 x + 826739.8; R² = 0.9983) drawn from 0.001 to 200 μg mL⁻¹ using a caffeine standard. Theobromine and theophylline concentrations were determined by referring to calibration curves (y = 1715.437707 x + 0.708100; R² = 0.9933; and y = 2448.170265 x + 0.731876; R² = 0.9945) drawn from 0.001 to 20 μg mL⁻¹.

2.6. Data processing

ProteoWizard MSConvert tools (Version 3.03.9393, 64-bit) was used to convert raw data (.d Agilent ones) to.mzXML file. Peak Picking was chosen as a filter option. All data processing was performed on the Workflow4Metabolomics infrastructure (https://workflow4metabolomics.org) (accessed on 15/04/22). Detailed steps and parameters that were used for data processing are publicly available on the W4M workflow repository (https://usegalaxy.fr/u/amontis/f/coffee-publish ed-2022) (accessed on 28/11/23) and as supplemental data 1. Each metabolite submitted to statistical analysis were called according to their m/z ratio (M) and retention time in sec (T); for example, M195T507 stands for caffeine.

 Principle Component Analysis (PCA) and loadings plot were used to illustrate clustering and markers of this clustering. Biosigner algorithm (Rinaudo et al., 2016) was used in workflow4metabolomics/galaxy platform to confirm the markers of each cluster considering three binary classifiers: RF (random forest), SVM (Support vector machine) and PLS-DA (Partial least squares-discriminant analysis).

The annotation process involved first carefully reviewing the information about the different m/z values in the available literature, and then by computing the ion elemental formulas based on the m/z values with MassHunter Qualitative Analysis® (Version B.10) from Agilent Technologies. Standard samples were also injected with the same LC conditions to confirm the identity of the most significant metabolites. The workflow showing sample preparation from harvest to LC-MS analysis is shown in supplemental data 2.

2.7. RNA extraction, reverse transcription and gene expression

Total RNA was extracted from plant tissues with the Maxwell RSC Plant RNA Kit (Promega Benelu, the Netherlands) and reverse transcription was done with the GoScript Reverse cDNA Synthesis Kit (Promega). Quantitative PCR was performed using the GoTag qPCR master mix (Promega) and the PicoReal Real-time PCR System (Thermo Fisher Scientific, Waltham, MA, USA). The thermal cycling conditions were as follows: 95 °C for 2 min, 40 cycles of 95 °C for 15 s, 60 °C for 1 min. The efficiency method, which compares the PCR efficiency of the reference genes with that of the target gene (Pfaffl, 2001), and the 2-ΔΔCt method (Livak and Schmittgen, 2001) were used to calculate the relative expression levels. The primers are listed in supplemental data 3. The workflow showing sample preparation for RT-qPCR experiments is shown in supplemental data 4.

3. Results

3.1. Quantification of caffeine, theobromine and theophylline in leaves and fruits

Caffeine, theobromine and theophylline were measured in the mature leaves (L2) and ripened fruits (RF) (Tables 1 and 2).

Caffeine was detected in C. arabica and C. canephora but not in C. anthonyi leaves. Theobromine concentration in leaves ranged from lowest in C. arabica to greatest in C. anthonyi. Theophylline was only found in C. anthonyi leaves (Table 1).

Caffeine concentration was greater in C. arabica than in C. anthonyi fruits (Table 2). Theobromine followed opposite pattern, and theophylline was undetected in both species. Quantification was not done in C. canephora fruits because that auto-sterile species did not fructify in the greenhouse, due to the absence of pollinators.

3.2. Expression of N-METHYLTRANSFERASES

As caffeine was detected in fruits but not in leaves of C. anthonyi (Montis et al., 2022), we looked at the relationship between levels of xanthines and transcripts of four genes (MXMT1, MXMT2, DXMT1 and DXMT2) involved in caffeine biosynthesis. The survey was conducted in young (L1) and L2 leaves, as well as in green (GF) and red (RF) fruits. The MXMT1 and MXMT2 genes encoding the enzymes responsible for conversion of 7-methylxanthine into theobromine, were expressed abundantly in L1, lowly in L2 and GF, and not at all in RF of C. anthonyi and C. canephora (Fig. 2 A, B). In comparison with the other species, the transcript levels of MXMT1 were undetected and the ones of MXMT2

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Concentration of caffeine, theobromine and theophylline in mature leaves (L2) of C. anthonyi, C. arabica and C. canephora. Values are expressed as mg g⁻¹ dry matter (mean ± SD; n = 3), N.D. = not detected.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C. anthonyi</td>
<td>C. arabica</td>
</tr>
<tr>
<td>Caffeine</td>
<td>N.D.</td>
</tr>
<tr>
<td>Theobromine</td>
<td>0.18 ± 0.02</td>
</tr>
<tr>
<td>Theophylline</td>
<td>N.D.</td>
</tr>
</tbody>
</table>
The DXMT1 and DXMT2 genes encoding enzymes converting 7-methylxanthine into theobromine, and also theobromine into caffeine, were sequentially more expressed in *C. anthonyi* leaves, *C. anthonyi* GFs and *C. arabica* leaves (Fig. 2 C,D). Interestingly, *C. anthonyi* GFs have greater transcript levels of DXMT1 and DXMT2 compared to other organs. This suggests that these genes have an equal contribution to caffeine biosynthesis in fruits of that species. Next, in both L1 and L2 of *C. arabica* and *C. canephora*, DXMT2 was more expressed than DXMT1. This indicates DXMT2 could contribute more to caffeine biosynthesis in leaves.

### 3.3. Quantification of feruloylquinic acids in leaves

As *C. anthonyi* has greater CQA quantities and it is a precursor of FQAs, the latter ones were quantified in mature leaves (L2) with a targeted metabolomic approach in a negative ionization mode. Three ions with *m/z* = 367.1038, 367.1043 and 367.1044 were detected (Fig. 3).

---

**Table 2**

Concentration of caffeine, theobromine and theophylline in ripened fruits (RF) of *Coffea anthonyi* and *C. arabica*. Values are expressed as mg g⁻¹ dry matter (mean ± SD, n = 3), N.D. = not detected.

<table>
<thead>
<tr>
<th></th>
<th><em>C. anthonyi</em></th>
<th><em>C. arabica</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Caffeine</td>
<td>2.0 ± 0.2</td>
<td>2.7 ± 0.4</td>
</tr>
<tr>
<td>Theobromine</td>
<td>0.11 ± 0.01</td>
<td>0.04 ± 0.01</td>
</tr>
<tr>
<td>Theophylline</td>
<td>N.D.</td>
<td>N.D.</td>
</tr>
</tbody>
</table>

---

![Fig. 2. Expression of genes involved in caffeine biosynthesis. Relative transcript levels of MXMT1 (A), MXMT2 (B), DXMT1 (C) and DXMT2 (D) in young (L1) and mature (L2) leaves of Coffea anthonyi, arabica and canephora, as well as in green (GF) and red fruits (RF) of *C. anthonyi* and *C. arabica*. N.D. = no data or values > 10 times lower than the lowest displayed value. ACTIN, UBQ and GAPDH were used as an internal reference genes. N = 3 technical replica ± std. Letters indicate statistical differences (p < 0.05).](image-url)
Based on the isotopic pattern and a literature survey, the formula \(C_{17}H_{19}O_9\) was assigned to three FQA isomers. The calculated \(\Delta\)ppm was about 2. As the standard compounds were not available for these three isomers, we performed a univariate Kruskal-Wallis test to detect the differences between the three Coffea species, based on the different metabolite intensities in the samples. The elution order for the three FQAs has been established according to the elution order of the three CQAs from which these compounds are originated from. Indeed, the 5-CQA eluted earlier than 4-CQA but later than 3-CQA (Montis et al., 2022) (supplemental data 2). Based on this assumption, the 5-FQA eluted earlier than 4-FQA but later than the 3-FQA. Given that 5-CQA is the most prevalent CQA in coffee leaves, that isomer should presumably be the one contributing most to the occurrence of FQAs.

The greatest levels of 3- and 5-FQA were detected in C. arabica and C. canephora leaves, while the lowest ones in C. anthonyi (Fig. 3). Both isomers were present at nearly the same level in the two first species. The level of 4-FQA was the greatest in C. arabica leaves. Even though C. anthonyi had the higher amount of CQAs and especially 5-CQA isomer (Montis et al. 2022), the amounts of FQAs in this species remained low compared to C. arabica and C. canephora.

3.4. Expression of O-methyltransferase

At the light of the quinic acid derivative quantification, the expression of the gene encoding a selective O-methyltransferase involved in the conversion of the CQA into FQA was investigated. The transcript levels of CCoAOMT were measured in L1 and L2 of the three species (Fig. 4).

In L1, the CCoAOMT transcript levels of both C. anthonyi and C. arabica were the lowest ones, and those of C. canephora the greatest ones. In L2, the transcript levels of C. anthonyi, C. canephora and C. arabica were sequentially increasing.

3.5. Coffee flower metabolomics

An untargeted metabolomic approach was used to detect the main xanthine derivates present in flower buds, as well as in the gynoecium.
and petals of *C. arabica*. To have a clear overview of the histolocalisation of the main annotated xanthishes, one PCA was first performed. The purpose was to highlight differences in the chemical composition between the three flower parts. A total of 1345 variables/metabolites were detected. The PCA score plots showed a fair separation of the samples, with PC1 accounting for 59 % of the total variation and PC2 for 22 % (Fig. 5A).

The loading plot showed features responsible for the clustering (Fig. 5B). Most of these showed rather low retention time values (>100 s). However, features M195T507 and M181T311 attracted notice, and were annotated as caffeine and theobromine, respectively. Since theobromine and theophylline are isomers of the same molecule, the identification of these compounds was performed by injecting both standard compounds and by checking their retention time. Interestingly, loadings nominate caffeine as a flower bud marker and theobromine as a gynoecium marker. However, theophylline was not pointed as a marker.

Caffeoylquinic acids were not detected in flowers, although previously found in fruits and leaves (Montis et al., 2022). The Biosigner results confirmed that caffeine and theobromine were the main discriminant metabolites across the samples. Surprisingly, theophylline appeared as a petal marker (Fig. 6). As a validation, these three xanthishes were simultaneously quantified in the various flower parts (Table 3).

Biosigner pointed that the main marker in the flower buds was caffeine (Fig. 6). The quantification showed caffeine in the flower buds was twice as more concentrated as in both the gynoecium and the petals (Table 3). Theobromine was returned as a gynoecium indicator (Fig. 6). The gynoecium contained three times more theobromine than the other flower parts (Table 3). Finally, theophylline was returned as a petal marker (Fig. 6). That metabolite was largely detected in the petals, but undetectable in the flower bud and the gynoecium (Table 3).

4. Discussion

4.1. The absence of caffeine correlated with the absence of DXMT1/2 expression in *C. anthonyi* leaves

The relationship between the metabolites and the transcript levels of biosynthetic genes (summarized in Fig. 7) is first discussed. The *C. anthonyi* species stood out from the other two, by an elevated concentration of theobromine and an absence of caffeine in the mature leaves. Theobromine concentration was greater in *C. canephora* compared to *C. arabica* and vice versa for caffeine (Table 1 & Fig. 7). Theophylline was only detected in *C. canephora*. Previous studies (Monteiro et al., 2019; Montis et al., 2021) confirmed that theobromine concentrations were greater in *C. canephora* (in the range 0.1 – 0.8 mg g⁻¹ dry matter) than in *C. arabica* (< 0.1 mg g⁻¹ dry matter) leaves (Fujimori and Ashihara, 1994). Nonetheless, values were ten times lower in the present study (Table 1), which may be due to theobromine being synthesized and quickly metabolised into caffeine. The concentrations of caffeine in *C. arabica* and *C. canephora* leaves were comparable to those in published reports (4–5 mg g⁻¹ dry matter) (Monteiro et al., 2019). The accumulation of caffeine during the leaf development is well known in *Coffea* species but theophylline is barely detected (Fujimori and Ashihara, 1994). The concentrations of theophylline were previously found in the range 0.1 mg g⁻¹ dry matter in both *C. arabica* and *C. canephora* leaves (Spiller, 2019). In our current study, theophylline concentration in *C. canephora* was comparable, while it was not detected in *C. arabica* (Table 1). The absence of caffeine in *C. anthonyi* leaves may be the reason why theophylline was not detected, underlying that theophylline is a catabolite of caffeine. However, despite the greater caffeine content (also reported by Fujimori and Ashihara, 1994), theophylline was not found in *C. arabica*.

Genes encoding different N-methyltransferases were found to be differently expressed in different species and in accordance with the leaf maturation stage. *C. canephora* was set apart from the two other species by exhibiting important transcript levels of MXMT1 and DXMT1/2 in the old leaves. These observations are in agreement with previous observations (Perrois et al., 2015; Uefuji et al., 2003). The most notable finding in *C. anthonyi*, was that the leaves did not express DXMT1/2, whereas the young leaves did express MXMT1/2. Theobromine is present in young leaves of *C. anthonyi* because the MXMT1/2 genes encoding the enzymes that convert 7-methylxanthine into theobromine, are highly expressed. The absence of expression of DXMT1/2 genes encoding the enzymes converting theobromine into caffeine and 7-methylxanthine into theobromine, could explain the absence of caffeine and the accumulation of theobromine in *C. anthonyi* leaves.

The MXMT1 and MXMT2 enzymes differ by their Michaelis constants. Indeed, the *Km* values are respectively 148 and 251 (Uefuji et al., 2003), meaning that MXMT1 has a greater substrate affinity compared.
to MXMT2. This could also account for greater theobromine in C. anthonyi compared to C. canephora leaves. Notably, MXMT1 was primarily expressed in the young leaves and theobromine was detected in greater concentrations in the mature leaves, indicating an accumulation of theobromine during development.

4.2. N-methyltransferases expression in C. anthonyi fruit explains the presence of caffeine

Fruits of C. canephora were not available due to culture conditions. Theobromine concentration was greater in mature fruits of C. anthonyi than C. arabica, while the opposite was observed for caffeine.

Fig. 6. Statistical comparison of the metabolite abundance illustrated by Biosigner. Boxplots show the normalised intensities of the main metabolites M195T507 (caffeine), M181T311 (theobromine) and M181T379 (theophylline) detected in the flower buds, gynoecium and petals. The boxplot illustrated here were obtained by Support Vector Machines (SVM).
Theophylline was not detected in the fruits. In published reports, theobromine was detected within the same range (Mehari et al., 2016; Spiller, 2019), but caffeine was more important in the range 5–23 mg g⁻¹ dry matter in C. arabica (Perrone et al., 2008; Alonso-Salces et al., 2009; Duarte et al., 2010; Gichimu et al., 2014; Mehari et al., 2016).

The DXMT1/2 transcript levels were important in green fruits of C. anthonyi when compared to C. arabica, while those of MXMT1/2 were very low in green and absent in red fruits of the two genotypes (Fig. 2A, B). These findings in C. arabica are in line with previous research. Early fruit developmental stages are shown to exhibit transient expression of MXMT1 (Perrois et al., 2015). Also, MXMT2 is early expressed but undetectable in yellow and red fruits of C. arabica (Maluf et al., 2009), and untraceable in C. canephora (Koshiro et al., 2006). A literature survey indicates DXMT1 transcripts are detected during developmental stages before C. arabica and C. canephora fruits ripen, and DXMT2 follows the same expression pattern (Koshiro et al., 2006; Ashihara et al., 2008). The C. arabica green fruits already reached their final size and were starting to turn green-yellow, which could account for the low DXMT transcript levels (Maluf et al., 2009).

We noted that DXMT1/2 were overall more expressed than MXMT1/2 across organs and species (Fig. 2). Hence, the two first genes probably have a greater contribution to theobromine accumulation in tissues, even though MXMT1/2 are solely converting 7-methylxanthine into theobromine. Presumably, low expression of DXMT1/2, rather than of these specific genes involved in theobromine biosynthesis, is the cause of caffeine biosynthesis disruption in C. anthonyi leaves. Finally, we conclude that xanthine concentration is partially influenced by genetic factors because the concentration of purine alkaloids in leaves and fruits correlated positively with the expression of biosynthetic genes across the three species. Caffeine synthesis is limited when DXMT expression is low, however there is a poor positive correlation between expression levels and the caffeine amount in organs. C. arabica had more caffeine in all organs, despite having lower transcripts than the other two species. Allelic variation in DXMT could account for that.

### 4.3. O-methyltransferase expression seems related to N-methyltransferase expression

The concentration of S-FQA is markedly low, while that of the precursor S-CQA is elevated in C. anthonyi. Notably, the conversion of CQA into FQA requires an O-methyltransferase and S-adenosyl-l-methionine (SAM) as co-factor for methylation. N- as O-methyltransferases use SAM as a cofactor and the similarity between O- and N-methyltransferases is understandable considering the residue alignment in C. arabica (Supplemental data 6). Consequently, it made sense to measure the expression of CCaoOMT transcript in leaves and indeed, it was two or three times less important in C. anthonyi leaves compared to the two other species. Considering the high CQA levels previously documented in C. anthonyi leaves (Montis et al., 2022), these findings were unexpected. A disruption in the FQAs biosynthetic pathway may occur when the O-methyltransferases are involved (Clifford et al., 2017). We conclude that altered caffeine and FQAs biosynthesis in C. anthonyi leaves is likely due to low expression of genes encoding O- and N-methyltransferases, both members of the methyltransferase superfamily.

### 4.4. Xanthine biosynthesis occurs early in the fruit development and was observed in flowers

To understand xanthine biosynthesis in fruits, xanthine metabolites quantification was also done in the flower buds, gynoecium and petals of C. arabica. The gynoecium, the flower tissue that will eventually become a fruit, had greater theobromine precursor levels compared to the flower buds or the petals. This could be due to a greater biosynthesis of caffeine prior to fructification. The presence of theophylline and the absence of theobromine in the petals, suggest that caffeine is more catabolized there than in the gynoecium. Thus, biosynthesis of caffeine occurs happens in the gynoecium, while its catabolism occurs in the petals. As caffeine was found very low in the phloem (Montis et al., 2022), and both caffeine and theobromine elevated in the gynoecium, we assume

### Table 3

Concentration of caffeine, theobromine and theophylline in C. arabica flowers. Results are expressed as mg g⁻¹ dry matter (mean ± SD, n = 5). N.D. = not detected.

<table>
<thead>
<tr>
<th>Flower buds</th>
<th>Gynoecium</th>
<th>Petals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caffeine</td>
<td>8.6 ± 0.6</td>
<td>4.01 ± 0.07</td>
</tr>
<tr>
<td>Theobromine</td>
<td>0.07 ± 0.01</td>
<td>0.9 ± 0.2</td>
</tr>
<tr>
<td>Theophylline</td>
<td>N.D.</td>
<td>N.D.</td>
</tr>
</tbody>
</table>

---

Fig. 7. Overview of metabolites and gene expression levels measured in C. anthonyi, C. arabica and C. canephora. Samples were young (L1) and mature (L2) leaves, green (GF) and ripened (RF) fruits. The left side panel indicates the concentrations of theobromine, caffeine and theophylline. 0 = not detected. The right side panel indicates the transcript levels of MXMT1/2 and DXMT1/2. Arrows show increasing levels and brackets very low levels.
that caffeine is synthesized at an early developmental stage and stored in the fruits. The optimal defence theory may also account for a more efficient caffeine biosynthesis in the gynoecium. Accordingly, plant organs that are more susceptible to pest attacks produce greater quantities of specialized metabolites. These are often more concentrated in the fruits than in any other tissues, because they contain the seeds to be dispersed for ensuring species survival. For example, the angel’s trumpet (Brugmansia suaveolens) fruits have twice as much scopalone as the flowers and leaves (Meldau et al., 2012).

5. Conclusion

The focus of this study was on C. anthonyi, a recently described species indigenous from Central Africa (Ref, 2009) and the widely cultivated C. arabica and C. canephora species, grown in a controlled tropical greenhouse environment. The purpose was to ascertain how metabolite concentration and biosynthetic gene expression relate to one another. Theobromine accumulation in leaves of C. anthonyi likely results from a disruption in caffeine biosynthesis, caused by the absence of synthesize gene expression. Still, green fruits regain expression and produce caffeine. The organ-specificity of caffeine biosynthesis emphasizes the significance of epigenetic and gene regulation. Actually, C. anthonyi possesses functional copies of synthesize genes but does not express them, while C. hombliotiana has mutated genes and lacks caffeine in any organ (Raharimalalala et al., 2021). In contrast, the closest relative C. eugonoids synthesizes low quantity of caffeine in leaves. Given that C. anthonyi successfully evolved over time, these findings call into question the defence role of caffeine in leaves (Meldau et al., 2012). A survey of the C. anthonyi population divergence should validate this evolution in caffeine histolocalisation and the origin of such observation. Finally, after examining C. arabica flowers, we showed that caffeine synthesis already occurs in the gynoecium, at a very early fruit development stage.

Funding

This research was supported by the FNRS (Grant n° 34553.08), FER-ULB 2007 and ULB-Platform. The plant samples were provided by the Meise Botanic Garden of Belgium.

CRediT authorship contribution statement

Piet Stoffelen: Writing – review & editing, Resources. Yusuka Noda: Writing – review & editing, Supervision, Formal analysis. Christian Hermans: Writing – review & editing, Writing – original draft, Validation, Resources, Investigation, Formal analysis, Data curation. Caroline Stevigny: Resources, Funding acquisition. Florence Sourad: Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Conceptualization. Pierre Van Antwerpen: Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization. Andrea Montis: Writing – original draft, Investigation, Formal analysis. Cedric Delporte: Writing – review & editing, Validation, Resources, Methodology, Formal analysis, Data curation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Data will be made available on request.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jplante.2024.112117.

References

A. Montis, F. Sourad, C. Delporte, P. Stoffelen, C. Steväig, P. Van Antwerpen, Targeted and untargeted mass spectrometry-based metabolomics for chemical profiling of three coffee species, Molecules 27 (2022) 3152.
