

# Molecular Cloning and Functional Characterization of Three Distinct *N*-Methyltransferases Involved in the Caffeine Biosynthetic Pathway in Coffee Plants<sup>1</sup>

Hiroataka Uefuji, Shinjiro Ogita, Yube Yamaguchi, Nozomu Koizumi, and Hiroshi Sano\*

Research and Education Center for Genetic Information, Nara Institute of Science and Technology, Nara 630-0192, Japan

Caffeine is synthesized from xanthosine through *N*-methylation and ribose removal steps. In the present study, three types of cDNAs encoding *N*-methyltransferases were isolated from immature fruits of coffee (*Coffea arabica*) plants, and designated as *CaXMT1*, *CaMXMT2*, and *CaDXMT1*, respectively. The bacterially expressed encoded proteins were characterized for their catalytic properties. *CaXMT1* catalyzed formation of 7-methylxanthosine from xanthosine with a  $K_m$  value of 78  $\mu\text{M}$ , *CaMXMT2* catalyzed formation of 3,7-dimethylxanthine (theobromine) from 7-methylxanthine with a  $K_m$  of 251  $\mu\text{M}$ , and *CaDXMT1* catalyzed formation of 1,3,7-trimethylxanthine (caffeine) from 3,7-dimethylxanthine with a  $K_m$  of 1,222  $\mu\text{M}$ . The crude extract of *Escherichia coli* was found to catalyze removal of the ribose moiety from 7-methylxanthosine, leading to the production of 7-methylxanthine. As a consequence, when all three recombinant proteins and *E. coli* extract were combined, xanthosine was successfully converted into caffeine in vitro. Transcripts for *CaDXMT1* were predominantly found to accumulate in immature fruits, whereas those for *CaXMT1* and *CaMXMT2* were more broadly detected in sites encompassing the leaves, floral buds, and immature fruits. These results suggest that the presently identified three *N*-methyltransferases participate in caffeine biosynthesis in coffee plants and substantiate the proposed caffeine biosynthetic pathway: xanthosine  $\rightarrow$  7-methylxanthosine  $\rightarrow$  7-methylxanthine  $\rightarrow$  theobromine  $\rightarrow$  caffeine.

Caffeine (1,3,7-trimethylxanthine) and theobromine (3,7-dimethylxanthine) are purine alkaloids, which are produced by a variety of plant species including coffee (*Coffea arabica*), tea (*Camellia sinensis*), maté (*Ilex paraguariensis*), guaraná (*Paullinia cupana*), cola (*Cola nitida*), and cacao (*Theobroma cacao*; Ashihara and Crozier, 2001). Accumulating in seeds, cotyledons, and/or young leaves, its biological role appears to be in chemical defense, serving as an antiherbivory (Nathanson, 1984; Bernays et al., 2000) and allelopathic compound (Waller, 1989; Smyth, 1992). In human societies, it has been widely used as a stimulant and an ingredient in drugs.

The biosynthesis pathway of caffeine in coffee plants is predicted to involve successive synthesis from purine precursors such as AMP and GMP through multiple steps catalyzed by several enzymes (Fig. 1; Ashihara et al., 1996). The proposed pathway of the final stage of caffeine biosynthesis is illustrated in Figure 1. The first step of the final stage of caffeine synthesis is methylation of xanthosine by xanthosine methyltransferase (XMT), yielding 7-methylxanthosine (Ashihara et al., 1996). An

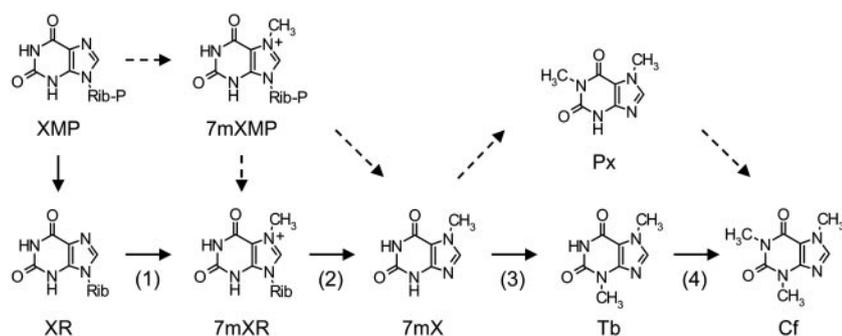
alternative entry, which is the conversion of xanthosine 5'-monophosphate (XMP) via channeled formation of 7-methylXMP to 7-methylxanthine, has also been proposed (Schulthess et al., 1996). After removal of its Rib residue by 7-methylxanthosine nucleosidase, the resulting 7-methylxanthine is methylated at the 3-*N* position by 7-methylxanthine methyltransferase (MXMT or theobromine synthase), producing 3,7-dimethylxanthine (theobromine). This latter is further methylated at the 1-*N* position by 3,7-dimethylxanthine methyltransferase (DXMT or caffeine synthase) to give 1,3,7-trimethylxanthine (caffeine). Some bypass pathways featuring 1,7-dimethylxanthine (paraxanthine) have also been proposed in tea plants (Kato et al., 1996), but the major pathway is suggested to be through theobromine in coffee plants (Fig. 1; Roberts and Waller, 1979; Ashihara et al., 1996). All methylation reactions require *S*-adenosyl-L-Met (Ado-Met) as a methyl donor.

Enzymatic activities of *N*-methyltransferase for caffeine biosynthesis in coffee plants have been detected in cell-free extracts prepared from immature fruits, cultured cells, and young leaves (Roberts and Waller, 1979; Baumann et al., 1983; Schulthess et al., 1996). *N*-methyltransferase proteins have also been purified (Mazzafera et al., 1994; Möslí Waldhauser et al., 1997b; Moisyadi et al., 1998). Subsequently, cDNAs for MXMT (*CaMXMT*, *CTS1*, and *CTS2*) were successfully cloned from coffee plants (Ogawa et al., 2001; Mizuno et al., 2003), although *CTS1* and *CaMXMT* were later found to be identical. In addition, a cDNA encoding a

<sup>1</sup> This work was supported by the New Energy and Industrial Technology Development Organization (grant) and by the Research for the Future Program (grant no. JSPS-RFTF00L01604) from the Japan Society for the Promotion of Science.

\* Corresponding author; e-mail sano@gtc.aist-nara.ac.jp; fax 81-743-72-5659.

Article, publication date, and citation information can be found at [www.plantphysiol.org/cgi/doi/10.1104/pp.102.019679](http://www.plantphysiol.org/cgi/doi/10.1104/pp.102.019679).



**Figure 1.** Caffeine biosynthetic pathway in coffee plants. Solid arrows indicate major routes, and broken arrows indicate minor or predicted routes. The first (1), third (3), and fourth (4) steps are *N*-methylation, and the second (2) step is Rib removal. 7 mXMP, 7-methylxanthosine 5'-monophosphate; XR, xanthosine; 7 mXR, 7-methylxanthosine; 7 mX, 7-methylxanthine; Tb, theobromine; Cf, caffeine.

dual-functional *N*-methyltransferase (CCS1) possessing MXMT and DXMT activities was recently cloned (Mizuno et al., 2003). This property resembles that of caffeine synthase from tea plants (TCS1; Kato et al., 2000).

In the present report, we describe isolation of three genes encoding XMT, MXMT and DXMT in coffee plants. Functional characterization of these enzymes and *in vitro* reconstitution partly substantiated the caffeine biosynthetic pathway suggested by several investigators in the past.

## RESULTS

### Isolation of cDNAs

To isolate cDNAs for enzymes involved in caffeine biosynthesis, we made two assumptions. First, we assumed that the three *N*-methyltransferases would resemble each other in nucleotide and amino acid sequences. Second, they would be expressed in immature fruits, which exhibit high caffeine synthesis (Suzuki and Waller, 1984). As a consequence, we attempted to isolate these genes from the cDNA pool derived from immature fruits by PCR using a primer set designed on the basis of the *CaMXMT* sequence. Among several PCR products obtained, cDNAs encoding CaMTL3 and CaMXMT proteins, and two novel proteins were identified. *CaMTL3* was renamed *CaXMT1*, and the other two cDNAs were designated *CaMXMT2* and *CaDXMT1* based on the respective *N*-methyltransferase properties of the corresponding recombinant proteins (see the following section). To avoid confusion, *CaMXMT* was also renamed as *CaMXMT1* in this article.

### Phylogenetic Analysis

The deduced amino acid sequences of CaXMT1 (41.8 kD), CaMXMT2 (43.4 kD), and CaDXMT1 (43.3 kD) showed significant similarity with each other and with CaMXMT1 (42.7 kD), exhibiting more than 82% identity (Fig. 2A). They were not found to possess any recognizable N-terminal signal sequences, but blocks of several residues were inserted or deleted in the C-terminal region. Phylogenetic analysis indicated that CaMXMT2 and CaDXMT1 were distinct from

CTS1 and CTS2, and CCS1, respectively (Fig. 2B). Phylogenetic analysis also indicated that these coffee *N*-methyltransferases were more closely related to carboxyl-methyltransferases, jasmonic acid carboxyl methyltransferase (*Arabidopsis*, AY008434), salicylic acid carboxyl methyltransferase (*C. breweri*, AF133053), and benzoic acid carboxyl methyltransferase (*A. majus*, AF198492), than to other *N*-methyltransferases, phosphoethandamine *N*-methyltransferase (*S. oleracea*, AF237633), coclaurine *N*-methyltransferase (*C. japonica*; AB061863), and putrescine *N*-methyltransferase 1 (*N. tabacum*; BAA05867). The identity of the coffee *N*-methyltransferases with TCS1 (tea, AB031280) was around 38% (Fig. 2B). These analyses suggested that coffee *N*-methyltransferases, together with CaMTL1 and CaMTL2, form a closely related group and that they constitute a branch in the plant methyltransferase family (Fig. 2B).

### Catalytic Specificity

CaMXMT1, CaMXMT2, and CaDXMT1 were expressed in *E. coli* as glutathione *S*-transferase (GST) fusion proteins and examined by thin-layer chromatography (TLC) for *N*-methyltransferase activity using xanthosine derivatives as the substrates. As the enzyme source, crude and purified protein samples were prepared (Fig. 3A). Purification through a glutathione-Sepharose column yielded expected recombinant proteins, although some degraded products of low molecular mass were present. The apparent molecular masses of the CaXMT1, CaMXMT2, and CaDXMT1 recombinant proteins were estimated to be 69, 74, and 81 kD, respectively, whereas the calculated value for each was approximately 70 kD. Purified CaDXMT1 appeared to contain more impurities than the others, probably due to degradation products.

Using these recombinant proteins, catalytic specificity was determined with the following substrates: XMP, xanthosine, xanthine, 1-methylxanthine, 3-methylxanthine, 7-methylxanthine, paraxanthine, theobromine, and theophylline (Fig. 3B–D; Table I). Purified CaXMT1 solely catalyzed conversion of xanthosine to 7-methylxanthosine (Fig. 3B, lane 2). In

**A**

```

CaMXMT1 MELQEVLMHNEGEGDTSYAKNASYN-LALAKVKPFLEQCIRELLRANLPNINKCIKVADLGCASGPNTLLTVRDIVQSID 79
CaMXMT2 MELQEVLMHNEGEGDTSYAKNASYN-LALAKVKPFLEQCIRELLRANLPNINKCIKVADLGCASGPNTLLTVRDIVQSID 79
CaXMT1 MELQEVLMHNEGEGDTSYAKNASYN-LVLAQVKPFLEQCIRELLRANLPNINKCIKVADLGCASGPNTLLTVRDIVQSID 80
CaDXMT1 MELQEVLMHNEGEGDTSYAKNSFYN-LFLIRVKPFLEQCIRELLRANLPNINKCIKVADLGCASGPNTLLTVRDIVQSID 79

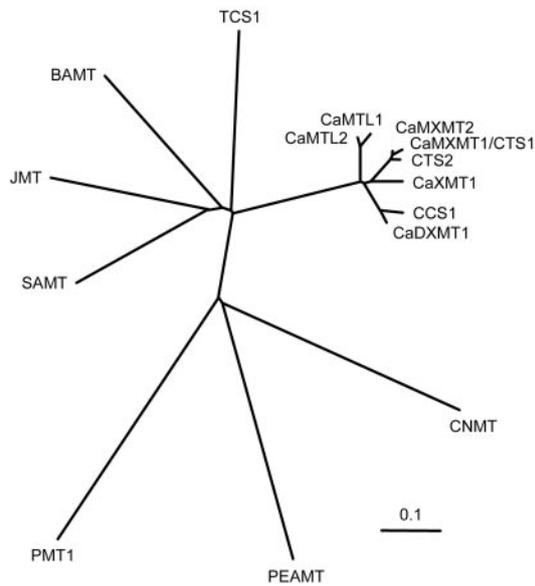
CaMXMT1 KVGQEEKNELEERPTIQIFLNDLFQNDNFNSVFKLLPSFYRKLEKENGSRKIGSCLISAMPGSFYGRLPPEESMHFLHSCYSV 159
CaMXMT2 KVGQEEKNELEERPTIQIFLNDLFQNDNFNSVFKLLPSFYRKLEKENGSRKIGSCLISAMPGSFYGRLPPEESMHFLHSCYSV 159
CaXMT1 KVGQEEKNELEERPTIQIFLNDLFQNDNFNSVFKLLPSFYRKLEKENGSRKIGSCLISAMPGSFYGRLPPEESMHFLHSCYCL 160
CaDXMT1 KVGQEEKNELEERPTIQIFLNDLFQNDNFNSVFKLLPSFYRKLEKENGSRKIGSCLISAMPGSFYGRLPPEESMHFLHSCYCL 159

CaMXMT1 HWLSQVPSGLVTELGICANKGSIYSSKGCRRPPVQKAYLDOFTKDFTTFLRIHSEELFSRGRMLLTCICKVDEEDDPNPLD 239
CaMXMT2 HWLSQVPSGLVTELGICANKGSIYSSKASRPPVQKAYLDOFTKDFTTFLRIHSEELFSRGRMLLTCICKVDEEDDPNPLD 239
CaXMT1 QNLSQVPSGLVTELGISTNKGSIYSSKASRRPPVQKAYLDOFTKDFTTFLRIHSEELFSRGRMLLTCICKVDEEDDPNPLD 240
CaDXMT1 HWLSQVPSGLVTELGISANKGCIYSSKASRPPVQKAYLDOFTKDFTTFLRIHSEELFSRGRMLLTCICKVDEEDDPNPLD 239

CaMXMT1 LLDMAINDLIVEGLEEBEKLDSFNIFPFFTPSAEEVKCIVEEEGSCFEILYLETFRAHYDAAFSIDDDYPVR-----SHEQ 313
CaMXMT2 LLDMAINDLIVEGHLEEBEKLDSFNIFPFFTPSAEEVKCIVEEEGSCFEILYLETFRAHYDAGFSIDDDYPVRSHFQVYGDEH 319
CaXMT1 LLEMAINDLIVEGHLEEBEKLDSFNIFVYIIPSAEEVKCIVEEEGSCFEILYLETFRKVLVDAGFSIDDD-----EH 310
CaDXMT1 LLEMSINDLIVEGHLEEBEKLDSFNIFVIYAPSTEEVKCIVEEEGSCFEILYLETFRKVPYDAGFSIDDDYQGRSHSPVSCDEH 319

CaMXMT1 IKAEYVASLIRSVYEPILASHFGEAIMPDLFHRLLAKHAAKVLHMGKGCYNNLIISLAKKPEKSDV 378
CaMXMT2 IKAEYVASLIRSVYEPILASHFGEAIMPDLFHRLLAKHAAKVLHMGKGCYNNLIISLAKKPEKSDV 384
CaXMT1 IKAEYVASLIRSVYEPILASHFGEAIMPDLFHRLLAKHAAKVLHMGKGCYNNLIISLAKKPEKSDV 372
CaDXMT1 ARAAHVASVRSIFEPIVASHFGEAIMPDLFHRLLAKHAAKVLHMGKGCYNNLIISLAKKPEKSDV 384
    
```

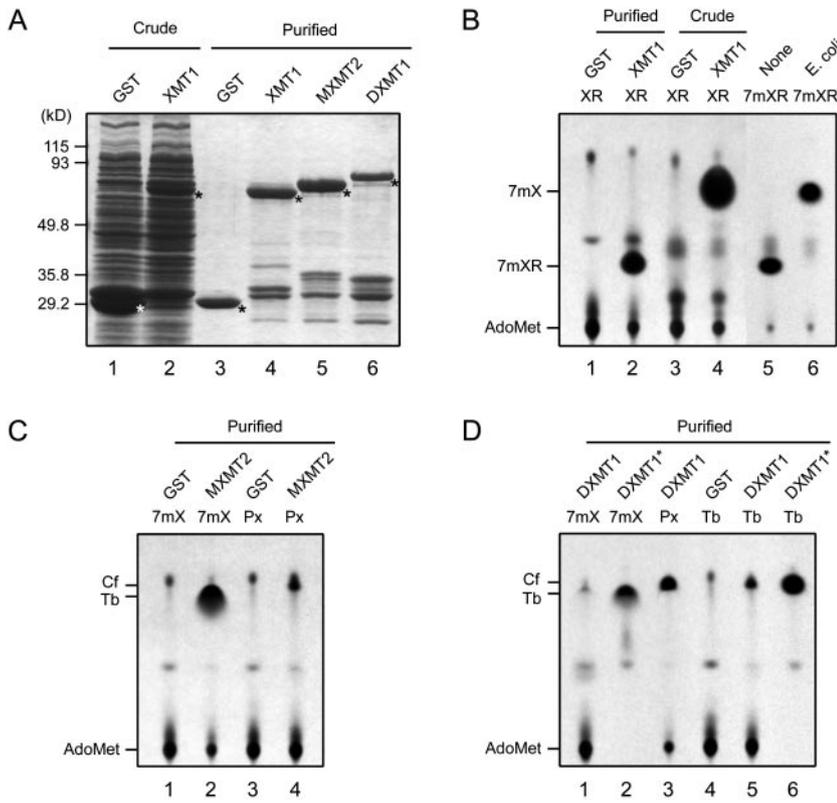
**B**



**Figure 2.** Comparison of deduced amino acid sequences of methyltransferases. A, Alignment of coffee *N*-methyltransferases. Genes, substrates, and accession numbers are as follows: CaXMT1, xanthosine, AB048793; CaMXMT1, 7-methylxanthine, AB048794; CaMXMT2, 7-methylxanthine, AB084126; and CaDXMT1, theobromine, AB084125. The sequences were aligned by ClustalW (Thompson et al., 1994). Identical residues shared by at least three proteins are shaded in black, and conserved substitutes found in at least three proteins are shaded in gray. B, Unrooted phylogenetic tree of coffee *N*-methyltransferases with related proteins. Genes, substrates, species nomenclatures, and accession numbers are as follows: CaMTL1, unknown, coffee, AB039725; CaMTL2, unknown, coffee, AB048792; CTS1, 7-methylxanthine, coffee, AB034700; CTS2, 7-methylxanthine, coffee, AB054841; CCS1, theobromine, coffee, AB086414; TCS1, theobromine, tea, AB031280; JMT, jasmonic acid methyltransferase, *Arabidopsis*, AY008434; SAMT, salicylic acid carboxyl methyltransferase, *Clarkia breweri*, AF133053; BMT, benzoic acid carboxyl methyltransferase, *Antirrhinum majus*, AF198492; PEAMT, phosphoethanolamine *N*-methyltransferase, *Spinacia oleracea*, AF237633; CNMT, coclaurine *N*-methyltransferase, *Coptis japonica*, AB061863; and PMT1, putrescine *N*-methyltransferase 1, *Nicotiana tabacum*, BAA05867. The sequences were compared using ClustalW, and the phylogenetic tree was generated by TreeView 1.5.2 (Page, 1996). The branch lengths represent numbers of substituted residues per site.

contrast, crude CaXMT1 catalyzed conversion of xanthosine to 7-methylxanthine, suggesting that conversion of xanthosine to 7-methylxanthosine and that of

7-methylxanthosine to 7-methylxanthine successively occurred (Fig. 3B, lane 4). Although the former reaction was catalyzed by CaMXT1, the latter was



**Figure 3.** N-methyltransferase activity. A, SDS-PAGE analysis of recombinant proteins. GST and CaXMT1, CaMXMT2, and CaDXMT1 fusion proteins with GST (XMT1, MXMT2, and DXMT1) were expressed in *Escherichia coli* and prepared as crude and purified samples. The samples were separated on a 9% (w/v) SDS-polyacrylamide gel and visualized by Coomassie Brilliant Blue staining. Asterisks indicate recombinant proteins. Lane numbers are given under the panel. B to D, TLC analysis of reaction products. Each recombinant protein was subjected to reaction with 500  $\mu\text{M}$  methyl group acceptor and 16  $\mu\text{M}$  [methyl- $^{14}\text{C}$ ] Ado-Met for 16 h. Reaction mixtures were subjected to TLC analysis, and methylated products were visualized by autoradiography. Proteins and methyl-acceptors are indicated above the panel. Identity of methylated products and Ado-Met are indicated on the left of the panel. Nonindicated spots are impurities. In B, [methyl- $^{14}\text{C}$ ] 7-methylxanthosine (7 mXR) produced by purified CaXMT1 was further applied to the Rib removal reaction by non-transformed *E. coli* crude extract (*E. coli*) and the extraction buffer (None) for 16 h. In D, concentrated samples are shown in the lanes marked with asterisks. XR, Xanthosine; 7 mXR, 7-methylxanthosine; 7 mX, 7-methylxanthine; Px, paraxanthine; Tb, theobromine; Cf, caffeine. Lane numbers referred to in the text are given under each panel.

found to be catalyzed by a nonspecific purine-nucleoside phosphorylase derived from *E. coli* (Fig. 3B, lane 6). CaMXMT2 catalyzed conversion of 7-methylxanthine to theobromine (Fig. 3C, lane 2). In addition, it catalyzed conversion of paraxanthine to caffeine with low activity (Fig. 3C, lane 4; Table I), consistent with the properties of CaMXMT1 (Table I).

**Table I.** Substrate specificities of the recombinant N-methyltransferases

All substrates were tested at a 500- $\mu\text{M}$  concentration, and methyltransferase activity was determined by measuring the radioactivity of the transferred  $^{14}\text{C}$ -labeled methyl group from AdoMet. The relative activity of each recombinant protein with the most preferred substrate was set at 100%. The no. in parentheses indicates the observed value of each activity. Values are the averages of three independent measurements. -, Not detected; XR, xanthosine; X, xanthine; 1mX, 1-methylxanthine; 3mX, 3-methylxanthine; 7mX, 7-methylxanthine; Px, paraxanthine; Tb, theobromine; Tp, theophylline; Cf, caffeine.

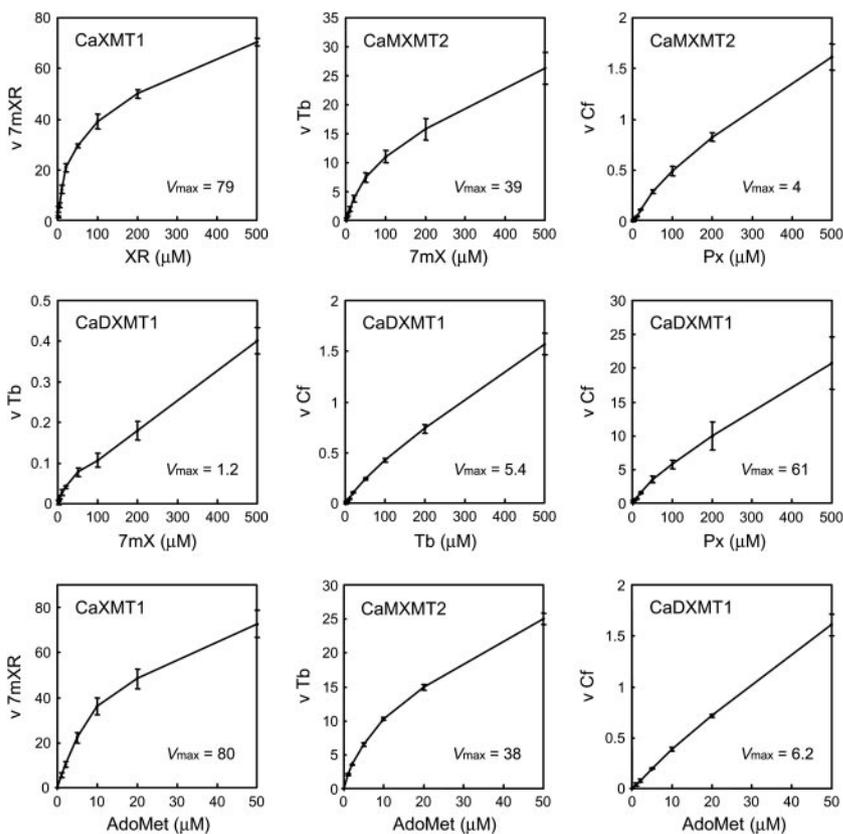
Substrate	Product	Relative Activity			
		CaXMT1	CaMXMT1	CaMXMT2	CaDXMT1
%(pmol min <sup>-1</sup> nmol protein <sup>-1</sup> )					
XMP		-	-	-	-
XR	7mXR	100 (98)	-	-	-
X		-	-	-	-
1mX		-	-	-	-
3mX		-	-	-	-
7mX	Tb	-	100 (66)	100 (63)	1.0
Px	Cf	-	5.0	5.3	100 (42)
Tb	Cf	-	-	-	3.8
Tp		-	-	-	-

CaDXMT1 catalyzed conversion of 7-methylxanthine to theobromine (Fig. 3D, lanes 1 and 2), that of theobromine to caffeine (Fig. 3D, lanes 5 and 6), and that of paraxanthine to caffeine (Fig. 3D, lane 3). However, sequential conversion of 7-methylxanthine to caffeine was not detected. Among the three substrates, paraxanthine appeared to be the most preferred by CaDXMT1 (Table I). None of the four enzymes catalyzed conversion of XMP to 7-methylXMP (Fig. 3; Table I). These results suggested that CaXMT1, CaMXMT2, and CaDXMT1 are involved in the first, third, and fourth steps of the caffeine biosynthetic pathway, respectively (Fig. 1).

**Kinetic Analysis**

To further investigate the potential role of each enzyme in detail, kinetic parameters were determined. Experiments were carefully designed to obtain reliable data by establishing the optimal reaction conditions for activity measurement of each enzyme (Fig. 4). The reaction velocity was assayed by varying the substrate concentration, and  $K_m$  values were estimated (Table II). The results revealed innate properties of each enzyme. CaXMT1 accepted only xanthosine, whereas CaMXMT1 and CaMXMT2 accepted both 7-methylxanthine and paraxanthine. The  $K_m$  values for these two substrates, however, were fewer in CaMXMT1 than in CaMXMT2, suggesting a distinct role for each. CaDXMT1 demonstrated an extremely high  $K_m$  value of around 1 mM for three acceptable

**Figure 4.** Effects of substrate concentration. Recombinant proteins shown in Figure 3A were subjected to reaction with 1 to 500  $\mu\text{M}$  methyl group acceptors and 1 to 50  $\mu\text{M}$  [methyl- $^{14}\text{C}$ ] Ado-Met (methyl group donor) for 1 h. Reaction velocity was measured with varied concentrations of the indicated substrate and a fixed concentration at maximum levels of the other substrate. When effects of Ado-Met concentration for CaXMT1, CaMXMT2, and CaDXMT1 were analyzed, the methyl group acceptors were XR, 7 mX, and Tb, respectively. The reaction products were separated by TLC and quantified with a bio-imaging analyzer system (BAS2500, Fuji Photo Film, Tokyo). The horizontal axis shows the concentration of the indicated substrate, and the vertical axis shows the velocity ( $v$ ) for the indicated product (picomoles per minute per nanomole of protein). Plots are the averages of three independent measurements.



substrates (7-methylxanthine, theobromine, and paraxanthine), indicating that caffeine may not be synthesized before sufficient amounts of precursors have accumulated. This is consistent with the different  $K_m$  values found for Ado-Met, being 15-fold more with CaDXMT1 than the others.

**In Vitro Reconstitution of the Caffeine Biosynthetic Pathway**

Recombinant protein samples of the three *N*-methyltransferases were analyzed in combination

**Table II.** Michaelis constants for substrates of recombinant *N*-methyltransferases

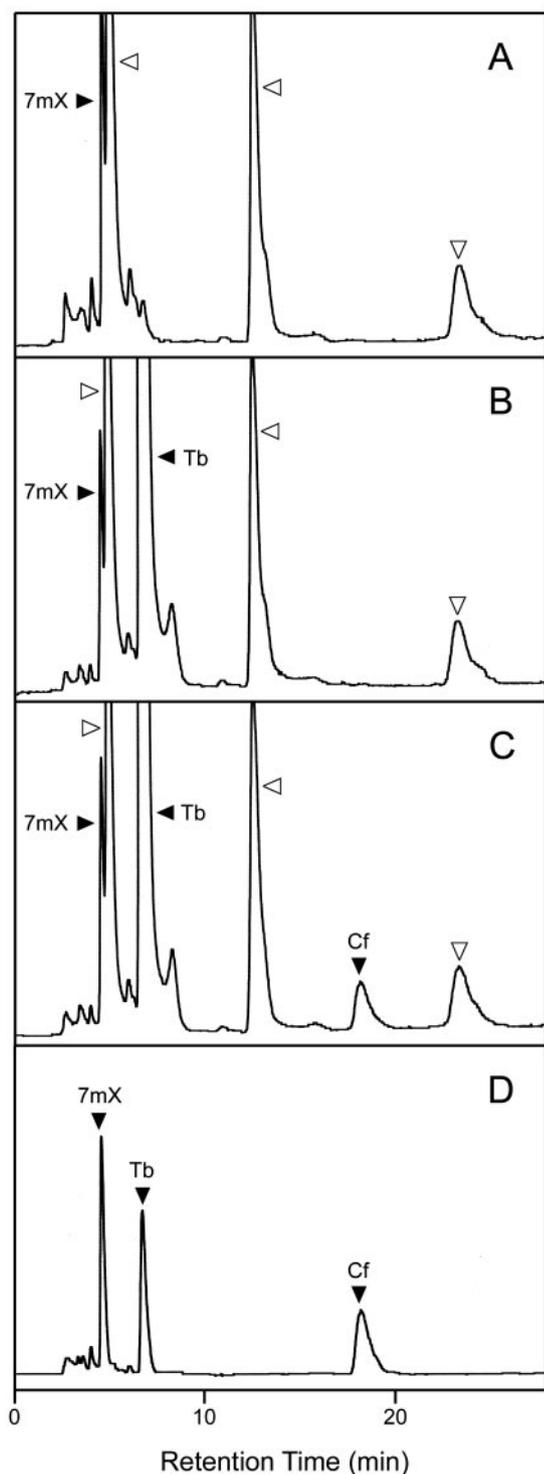
The values are the averages of the three independent measurements shown in Figure 4. Each value was obtained using Anemona (Hernandez and Ruiz, 1998).

Substrate	$K_m$			
	CaXMT1	CaMXMT1	CaMXMT2	CaDXMT1
	$\mu\text{M}$			
Methyl group acceptor				
XR	78 $\pm$ 7			
7mX		148 $\pm$ 7	251 $\pm$ 20	916 $\pm$ 18
Px		458 $\pm$ 36	738 $\pm$ 51	973 $\pm$ 88
Tb				1,222 $\pm$ 197
Methyl group donor				
AdoMet	13 $\pm$ 2	12 $\pm$ 2	14 $\pm$ 1	153 $\pm$ 31

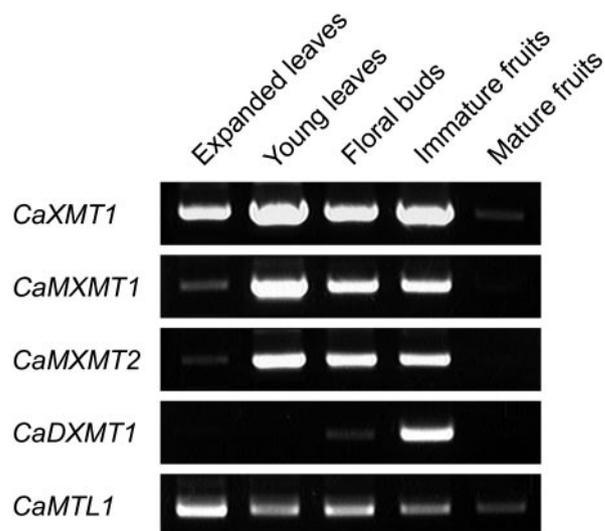
reactions using xanthosine as the starting material, and reaction products were identified by HPLC analysis. Crude CaXMT1 alone catalyzed synthesis of 7-methylxanthine (Fig. 5A), consistent with the TLC assay findings (Fig. 3B). A combination of crude CaXMT1 and purified CaMXMT2 catalyzed the production of 7-methylxanthine and theobromine (Fig. 5B). A combination of crude CaXMT1, purified CaMXMT2, and CaDXMT1 catalyzed the production of 7-methylxanthine, theobromine, and caffeine (Fig. 5C). These results clearly demonstrated that the caffeine biosynthetic pathway could be accomplished in vitro using the three *N*-methyltransferases and *E. coli* extract.

**Tissue-Specific Transcript Accumulation**

Transcript accumulation for the identified *N*-methyltransferase genes was analyzed by reverse transcriptase (RT)-PCR using gene-specific primer sets (Fig. 6). Transcripts for *CaXMT1* were found in all tissues except in mature fruits. Transcripts for *CaMXMT1* and *CaMXMT2* were identified at high levels in young leaves, floral buds, and immature fruits. In contrast, those of *CaDXMT1* were predominantly detected in immature fruits. None of them were found in mature fruits. These results support previous observation that immature fruits are the major sites of caffeine biosynthesis in coffee plants.



**Figure 5.** In vitro reconstitution of caffeine biosynthetic pathway. A to C, Single or combined recombinant protein samples shown in Figure 3A were subjected to reaction with  $500 \mu\text{M}$  xanthosine as the sole methyl group acceptor (starting material) and  $1.5 \text{ mM}$  Ado-Met for 16 h, and the reaction products were identified by HPLC analysis. The recombinant protein samples were crude CaXMT1 alone (A), crude CaXMT1 and purified CaMXMT2 (B), and crude CaXMT1, purified CaMXMT2, and purified CaDXMT1 (C). Authentic standards were also run in parallel (D). Black arrowheads indicate reaction products and standards, and white arrowheads indicate impurities.



**Figure 6.** Tissue-specific transcript accumulation. Total RNAs were prepared from indicated tissues of coffee, and RT-PCR analysis was performed using the gene-specific primer sets. *CaMTL1* was used as the internal standard (Ogawa et al., 2001).

## DISCUSSION

This paper documents isolation of three genes encoding *N*-methyltransferases, which are possibly involved in caffeine biosynthesis in coffee plants. Based on the distinct catalytic properties of the presently identified enzymes, we conclude that, in coffee plants, xanthosine is step-wise methylated at 7-, 3-, and 1-*N* positions to yield the final product, caffeine. This was directly shown by in vitro conversion of xanthosine into caffeine by a mixture of the three enzymes, although the Rib moiety of 7-methylxanthosine was removed by the host *E. coli* purine-nucleoside phosphorylase, which was present in the crude CaXMT1 preparation.

The above-stated conclusion was also partly strengthened by kinetic analyses. Supposing the  $K_m$  values to indicate the affinity for the substrate, the results point to several notable properties of each enzyme. First, the substrate affinity of the concerned enzyme proportionally decreased toward the end point, showing  $K_m$  values of  $78 \mu\text{M}$  for xanthosine (CaXMT1),  $250 \mu\text{M}$  for 7-methylxanthine (CaMXMT2), and  $1,200 \mu\text{M}$  for theobromine (CaDXMT1; Table II). This may indicate that the further down the pathway, the more the amount of substrate compound required, making the reaction proceed irreversibly and step-wise. Second, 7-methylxanthine predominates over paraxanthine as the substrate with a 3-fold higher affinity and a 20-fold higher reaction velocity for both CaMXMT1 and CaMXMT2, indicating the former to be the genuine substrate. Third, CaXMT1 and CaMXMT2 may share the same reaction, being high- and low-affinity enzymes, respectively. Fourth, although CaDXMT1 is able to react with 7-methylxanthine, its affinity is less than one-sixth of that of CaXMT1 (or one-third of CaMXMT2), suggesting

that it may function only with theobromine, or if available, with paraxanthine. Fifth, the  $K_m$  for Ado-Met was around  $10 \mu\text{M}$  for CaXMT1, CaMXMT1, and CaMXMT2, consistent with values for other methyltransferases, whereas it was  $150 \mu\text{M}$  for CaDXMT1, 15-fold larger than the former. Because Ado-Met is the common methyl donor in biological methylation and, therefore, is always scarce in cells (Nakano et al., 2000), such a low affinity of CaDXMT1 may ensure that caffeine synthesis only takes place upon accumulation of sufficient substrates. It should be noted that the  $K_m$  value for each enzyme involved in caffeine biosynthesis considerably varies depending upon protein species. For example, the  $K_m$  value for 7-methylxanthine was reported to be 873 and  $171 \mu\text{M}$  for CTS1 and CTS2, respectively, and that for theobromine to be  $157 \mu\text{M}$  for CCS1 (Mizuno et al., 2003). These variations suggest that coffee plants possess multiple enzymes with distinct affinities toward substrates, thereby catalyzing caffeine biosynthesis at a broad range of available substrate concentrations.

It has been suggested that paraxanthine is the most active precursor of a coffee *N*-methyltransferase in vitro (Roberts and Waller, 1979), consistent with the present results (Table I). However, its formation from 1-methylxanthine or 7-methylxanthine is not detectable, leading to a conclusion that caffeine is not synthesized from paraxanthine in coffee plants (Roberts and Waller, 1979). Although CaDXMT1, as TCS1, showed clear potential for conversion of paraxanthine into caffeine, we could not identify a cDNA encoding a paraxanthine synthase from our coffee cDNA pool, thus leaving the question open as to whether or not paraxanthine is a substrate for caffeine synthesis in planta.

The present study also revealed the properties of some proteins previously reported to be involved in caffeine synthesis in coffee plants. First, the present CaXMT1 was found to be identical with CaMTL3, whose sequence and tissue-specific expression were described in our previous report (Ogawa et al., 2001). Although we suggested it to catalyze methylation of caffeine derivatives, no activity was assigned to any known reaction at that time, probably due to the chloroform extraction method, because xanthosines are not soluble in chloroform. Second, xanthosine 7-*N*-methyltransferase (XMT) and its corresponding gene have been claimed to be isolated from coffee leaves (Moisyadi et al., 1999). However, its sequence resembles not that of present CaXMT1, but rather a lipolytic enzyme (H. Uefuji, unpublished data), and enzymatic activity was not determined (Moisyadi et al., 1999). We also isolated a cDNA having 96% identity with XMT, but its recombinant protein did not cause any conversion of xanthosine to 7-methylxanthosine (unpublished observation). Thus, whether or not XMT is involved in xanthosine methylation is currently not clear, and further characterization is required.

Coffee plants possess multiple enzymes, or isoforms to catalyze the second step of methylation, or theobromine synthesis. Two were identified by our study (*CaMXMT1* and *CaMXMT2*), and an additional two (*CTS1* and *CTS2*) were also reported (Mizuno et al., 2003). The same might be the case for the first and the last steps, as judged from the presence of many closely related genes, and from the expression pattern of the identified genes. For example, *CaDXMT1* is almost exclusively expressed in immature fruits, whereas *CaXMT1*, *CaMXMT1*, and *CaMXMT2* are found in leaves, floral buds, and immature fruits. Because young leaves contain a high level of caffeine and DXMT activity (Ashihara et al., 1996; Mösl Waldhauser et al., 1997a), these observations suggest that another form of DXMT is present in young leaves. Thus, it is conceivable that coffee plants are equipped with multiple sets of enzymes, which may be necessary for constitutive production of caffeine in relevant tissues.

Taken together with the present findings and previous observations, it can be concluded safely that, in coffee plants, caffeine is synthesized through stepwise methylation of xanthosine at 7-, 3-, and 1-*N* positions catalyzed by corresponding specific enzymes. To draw a definitive conclusion, however, a different approach must be taken, in which expression or repression of these genes results in alteration in caffeine content.

Finally, it is worthy to mention that because caffeine inhibits pest feeding and is pesticidal at concentrations known to occur in coffee and tea plants (Nathanson, 1984; Hollingsworth et al., 2002), transgenic plants producing caffeine would be expected to be resistant to herbivorous insects. Identification of all three genes for *N*-methyltransferases makes conceivable the construction of such plants for practical use.

## MATERIALS AND METHODS

### Plant Materials

Coffee (*Coffea arabica*) plants were grown to maturity in a greenhouse, and young leaves (<2.5 cm long), old leaves (>10 cm long), floral buds, immature (green, stage 6–7) fruits, and mature (red, stage 10) fruits were collected and preserved at  $-80^\circ\text{C}$  until used for total RNA extraction. Developmental stage of fruits was identified as previously described (Keller et al., 1972).

### cDNA Isolation

Total RNA was extracted as previously described (Chang et al., 1993), and mRNAs were purified using PolyAtract mRNA Isolation System III (Promega, Madison, WI) and converted into double-stranded cDNAs using a ZAP-cDNA Synthesis Kit (Stratagene, La Jolla, CA). Two oligonucleotide primers, 5'-ATGGAGCTCCAAGAAGTCCT-3' and 5'-CTTTTACACGCT-GACTTCTCTG-3', containing the start and stop codon regions of *CaMXMT1*, respectively, were synthesized. PCR was performed using *Pyrobest* DNA polymerase (Takara, Otsu, Japan) under the condition of a 36-cycle denaturation at  $98^\circ\text{C}$  for 10 s, annealing at  $55^\circ\text{C}$  for 30 s, and extension at  $72^\circ\text{C}$  for 1.5 min. Amplified cDNAs containing complete coding sequences were subcloned into pBluescript II KS- (Stratagene), and *Escherichia coli* (DH5 $\alpha$ )

was transformed with the resulting plasmids. Fifty-six of the cDNA clones were randomly selected and nucleotide sequences were determined using a DNA sequencing kit (PE-Applied Biosystems, Foster City, CA) and a DNA sequencer (ABI Prism 3100 Genetic Analyzer; PE-Applied Biosystems). Nucleotide and deduced amino acid sequences were analyzed with the GeneWorks 2.5.1 program (Oxford Molecular Group, Oxford).

### Production of Recombinant Proteins

The cDNA clones (*CaXMT1*, *CaMXMT1*, *CaMXMT2*, and *CaDXMT1*) were inserted into the GST fusion vector pGEX-4T-2 (Amersham-Pharmacia Biotech, Uppsala), and *E. coli* (BL21) was transformed with the resulting plasmids. Expression of GST fusion proteins was induced by 1 mM isopropyl- $\beta$ -D-thiogalactoside at 20°C for 16 h, and recombinant proteins were purified essentially according to the manufacturer's manual (Amersham-Pharmacia Biotech). Samples in 50 mM Tris-HCl (pH 8.0) containing 1 mM EDTA and 5 mM dithiothreitol were sonicated and centrifuged, and supernatants were collected as the crude recombinant proteins. Ten micrograms of the recombinant protein sample was subjected to SDS-PAGE on a 9% (w/v) gel and visualized by Coomassie Brilliant Blue staining.

### N-Methyltransferase Activity Assay

A 25- $\mu$ L reaction mixture containing 50 mM Tris-HCl (pH 8.0), 500  $\mu$ M substrate, 16  $\mu$ M S-adenosyl-L-[methyl- $^{14}$ C] Met (2.2 GBq  $\text{mmol}^{-1}$ ; Amersham-Pharmacia Biotech), 200  $\mu$ M  $\text{MgCl}_2$ , and recombinant protein (5  $\mu$ g of purified sample or 50  $\mu$ g of crude sample) was incubated at 27°C for 16 h. After filtration using centrifugal filter units (Ultrafree-MC, Millipore, Bedford, MA) and mixing with an equal volume of methanol, 4- $\mu$ L aliquots were then subjected to TLC analysis using TLC aluminum sheets (Silica gel 60 F254, Merck, Rahway, NJ) with a solvent consisting of *n*-butyl alcohol:acetic acid:water (4:1:2 [v/v]). The chromatogram sheets were sprayed with  $\text{En}^3\text{Hance}$  (NEN Life Science Products, Boston), and reaction products were identified by autoradiography with exposure to x-ray film (Bio Max MS, Eastman-Kodak, Rochester, NY) for 1 week at  $-80^\circ\text{C}$ . Control samples as the standards were subjected to the same procedure and detected by UV illumination. The Rib-removal reaction was performed with [methyl- $^{14}$ C] 7-methylxanthosine substrate, which was prepared with purified *CaXMT1* as described above. The reaction mixture was then supplemented with non-transformed *E. coli* crude extract containing 50  $\mu$ g of proteins, further incubated for 16 h, and assayed for the product by HPLC.

### Kinetic Analysis

For determination of substrate specificity, a 25- $\mu$ L reaction mixture containing 500  $\mu$ M methyl group acceptor, 50  $\mu$ M [methyl- $^{14}$ C] Ado-Met, and 200 ng  $\mu\text{L}^{-1}$  purified protein was incubated at 27°C for 1 h. The  $K_m$  was determined by varying the concentration of methyl acceptor between 1 and 500  $\mu$ M or the concentration of Ado-Met between 1 and 50  $\mu$ M [methyl- $^{14}$ C] Ado-Met. Reaction mixtures containing a methyl donor and acceptor, and 2, 20, or 200 ng  $\mu\text{L}^{-1}$  of purified protein were incubated for 1 h. The linearity of the reaction velocity in terms of elapsed time and enzyme amount was confirmed by plotting the initial velocity against four different time points. The methylated products separated by TLC were detected and quantified with a bio-imaging analyzer system (BAS2500, Fuji). Kinetic parameters were calculated from the Michaelis-Menten equation with the Anemona program (Hernandez and Ruiz, 1998).

### Combination Reactions of N-Methyltransferases

A 100- $\mu$ L reaction mixture containing 50 mM Tris-HCl (pH 8.0), 500  $\mu$ M xanthosine, 1.5 mM Ado-Met, 200  $\mu$ M  $\text{MgCl}_2$ , and combinations of proteins (20  $\mu$ g of purified or 200  $\mu$ g of crude samples) was incubated at 27°C for 16 h and subjected to extraction with 1 mL of chloroform. After extracts were dried at 60°C and resolved in 200  $\mu$ L of a solvent containing 50 mM Na-phosphate (pH 6.0):methanol (4:1 [v/v]), 20- $\mu$ L aliquots were subjected to HPLC analysis using a Puresil C18 column (Waters, Milford, MA). The column was developed at a flow rate of 1 mL  $\text{min}^{-1}$  with the solvent by multisolvent delivery system (Waters) and reaction products were densito-

metrically monitored at 270 nm with a tunable absorbance detector (Waters).

### RT-PCR Analysis

Total RNAs were prepared from various tissues using an RNeasy Plant Mini Kit (Qiagen USA, Valencia, CA) and subjected to RT-PCR with five sets of gene-specific primers: *CaXMT1*, 5'-ATCAACTGGTTCTCGCAAG-3' and 5'-CTGCTCTAACGGAAGATGCA-3'; *CaMXMT1*, 5'-TCCTACAATCTGCTCTTGC-3' and 5'-TGCTTTAATTTGTCATGGGATC-3'; *CaMXMT2*, 5'-CCTACAATCTGGCTCTTGC-3' and 5'-TTCATCGCCGTATACATTGGA-3'; *CaDXMT1*, 5'-TCATTTCTACAATCTGTTTCTCATCAG-3' and 5'-TATGGAATTCGGGTTCTCGA-3'; and *CaMTL1*, 5'-ATTCATCCTCAACTCAACTGGT-3' and 5'-TTCTACTGAAGCTGTATAGATTGGAAC-3', and an RNA PCR Kit (AMV) version 2.1 (Takara). PCR was carried out under the condition of a 28-cycle of denaturation at 94°C for 30 s, annealing at 58°C for 30 s, and extension at 72°C for 1 min. After fractionation on agarose gel electrophoresis, products were identified by visualization with ethidium bromide staining.

### ACKNOWLEDGMENTS

We would like to thank Syo Kurokawa (The Botanic Gardens of Toyama, Japan) for a generous gift of plant materials and Dr. Malcolm Moore (Internal, Nagoya, Japan) for critical reading of the manuscript.

Received December 23, 2002; returned for revision January 29, 2003; accepted March 2, 2003.

### LITERATURE CITED

- Ashihara H, Crozier A (2001) Caffeine: a well known but little mentioned compound in plant science. *Trends Plant Sci* 6: 407–413
- Ashihara H, Monterio AM, Gillies FM, Crozier A (1996) Biosynthesis of caffeine in leaves of coffee. *Plant Physiol* 111: 747–753
- Baumann TW, Koetz R, Morath P (1983) N-methyltransferase activities in suspension cultures of *Coffea arabica* L. *Plant Cell Rep* 2: 33–35
- Bernays EA, Oppenheim S, Chapman RF, Kwon H, Gould F (2000) Taste sensitivity of insect herbivores to deterrents is greater in specialists than in generalists: a behavioral test of the hypothesis with two closely related caterpillars. *J Chem Ecol* 26: 547–563
- Chang S, Puryear J, Cairney J (1993) A simple and efficient method for isolating RNA from pine trees. *Plant Mol Biol Rep* 11: 113–116
- Hernandez A, Ruiz MT (1998) An EXCEL template for calculation of enzyme kinetic parameters by non-linear regression. *Bioinformatics* 14: 227–228
- Hollingsworth RG, Armstrong JW, Campbell E (2002) Caffeine as a repellent for slugs and snails: at high concentrations this stimulant becomes a lethal neurotoxin to garden pests. *Nature* 417: 915–916
- Kato M, Kanehara T, Shimizu H, Suzuki T, Gillies FM, Crozier A, Ashihara H (1996) Caffeine biosynthesis in young leaves of *Camellia sinensis*: in vitro studies on N-methyltransferase activity involved in the conversion of xanthosine to caffeine. *Physiol Plant* 98: 629–636
- Kato M, Mizuno K, Crozier A, Fujimura T, Ashihara H (2000) Caffeine synthase gene from tea leaves. *Nature* 406: 956–957
- Keller H, Wanner H, Baumann TW (1972) Caffeine synthesis in fruits and tissue cultures of *Coffea arabica*. *Planta* 108: 339–350
- Mazzafera P, Wingsle G, Olsson O, Sandberg G (1994) S-adenosyl-L-methionine: theobromine 1-N-methyltransferase, an enzyme catalyzing the synthesis of caffeine in coffee. *Phytochemistry* 37: 1577–1584
- Mizuno K, Okuda A, Kato M, Yoneyama N, Tanaka H, Ashihara H, Fujimura T (2003) Isolation of a new dual-functional caffeine synthase gene encoding an enzyme for the conversion of 7-methylxanthine to caffeine from coffee (*Coffea arabica* L.). *FEBS Lett* 534: 75–81
- Moisyadi S, Neupane KR, Stiles JI (1998) Cloning and characterization of a cDNA encoding xanthosine-N<sup>7</sup>-methyltransferase from coffee (*Coffea arabica*). *Acta Hort* 461: 367–377
- Moisyadi S, Neupane KR, Stiles JI (1999) Cloning and characterization of xanthosine-N<sup>7</sup>-methyltransferase, the first enzyme of the caffeine biosyn-

- thetic pathway. Proceedings of the 18th International Scientific Colloquium on Coffee. Association Scientifique International du Café, Paris, pp 327–331
- Möslí Waldhauser SS, Gillies FM, Crozier A, Baumann TW (1997a)** *N*-methyltransferase activities in caffeine biosynthesis: biochemical characterization and time course during leaf development of *Coffea arabica*. *Phytochemistry* **44**: 853–859
- Möslí Waldhauser SS, Gillies FM, Crozier A, Baumann TW (1997b)** Separation of *N*-7 methyltransferase, the key enzyme in caffeine biosynthesis. *Phytochemistry* **45**: 1407–1414
- Nakano Y, Koizumi N, Kusano T, Sano H (2000)** Isolation and properties of an *S*-adenosyl-*L*-methionine binding protein from the green alga, *Chlamydomonas reinhardtii*. *J Plant Physiol* **157**: 707–711
- Nathanson JA (1984)** Caffeine and related methylxanthines: possible naturally occurring pesticides. *Science* **226**: 184–187
- Ogawa M, Herai Y, Koizumi N, Kusano T, Sano H (2001)** 7-Methylxanthine methyltransferase of coffee plants. *J Biol Chem* **276**: 8213–8218
- Page RD (1996)** TREEVIEW: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci* **12**: 357–358
- Roberts MF, Waller GR (1979)** *N*-Methyltransferases and 7-methyl-*N*<sup>9</sup>-nucleoside hydrolase activity in *Coffea arabica* and the biosynthesis of caffeine. *Phytochemistry* **18**: 451–455
- Schulthess BH, Morath P, Baumann TW (1996)** Caffeine biosynthesis starts with the metabolically channeled formation of 7-methyl-XMP: a new hypothesis. *Phytochemistry* **41**: 169–175
- Smyth DA (1992)** Effect of methylxanthine treatment on rice seedling growth. *J Plant Growth Regul* **11**: 125–128
- Suzuki T, Waller GR (1984)** Biosynthesis and biodegradation of caffeine, theobromine and theophylline in *Coffea arabica* L. fruits. *J Agric Food Chem* **32**: 845–848
- Thompson JD, Higgins DG, Gibson TJ (1994)** CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673–4680
- Waller GR (1989)** Biochemical frontiers of allelopathy. *Biol Plant* **31**: 418–447