

Isolation and characterization of an *AGAMOUS* homologue from cocoa

Tetty Chaidamsari^{a,1}, Samanhudi^b, Herti Sugiarti^c, Djoko Santoso^c,
Gerco C. Angenent^a, Ruud A. de Maagd^{a,*}

^aBusiness Unit Bioscience, Plant Research International B.V., P.O. Box 16, 6700 AA Wageningen, The Netherlands

^bFaculty of Agriculture, Sebelas Maret University, Jl. Ir. Sutami 36A, Surakarta 57126, Indonesia

^cIndonesian Biotechnology Research Institute for Estate Crops, Jl. Taman Kencana 1, Bogor 16151, Indonesia

Received 3 November 2005; received in revised form 4 January 2006; accepted 8 January 2006

Available online 31 January 2006

Abstract

We report the cloning of a cDNA from *TcAG*, an *AG* (*Arabidopsis thaliana* MADS-box C-type transcription factor gene *AGAMOUS*) homologue from cocoa (*Theobroma cacao* L.). *TcAG* was in the cocoa flower expressed primarily in stamens and ovaries, comparable to *AG* in *Arabidopsis*. Additionally, we found that *TcAG* is also expressed in the fruit (pod) wall and during its entire development, as well as in the fruit pulp. Ectopic expression of *TcAG* in transgenic *A. thaliana* plants resulted in a range of weak to strong *apetala2* (*ap2*) mutant-like phenotypes as well as early flowering and curly leaves, as observed in other studies of plants overexpressing a functional *AG* homologue. The severity of the phenotypes correlated positively with the *TcAG* transcript level in the transgenic plants.

© 2006 Elsevier Ireland Ltd. All rights reserved.

Keywords: Flower development; Cocoa; *AGAMOUS*; MADS-box

1. Introduction

Cocoa trees produce large numbers of flowers, at certain times of the year, from the age of 3 years and on, under good growing conditions. Flowers emerge from floral cushions (Fig. 1A), a meristematic tissue in the bark of the stem and branches at the location of old leaf axils [1], a habit called cauliflorous or truncate. The flowers are small, about 15 mm in diameter and form long pedicels (Fig. 1B). The flowers have 5 unfused sepals and petals, 10 stamens and an ovary of 5 united carpels (Fig. 1B). The petals are very narrow at the base but expand into a cup-shaped pouch and end in a broad tip or ligule. The 10 stamens are in 2 whorls, the outer whorl is consisting of 5 long non-fertile staminodes and the inner whorl of 5 fertile stamens. Cocoa trees produce flowers throughout the year, but flowering peaks in certain times of the year, when flowering progresses in waves from the lower stem flowering cushions to the branches. Flowers are open pollinated by small insects, but only 1–5% of the flowers are successfully pollinated and proceed to produce a pod [2].

In most higher eudicotyledonous flowering plants the floral organs are arranged in four different whorls, containing sepals, petals, stamens and carpels, respectively. The specification of floral organ identity is explained by the ABC model [3,4] and extensions thereof [5]. Classes A–C genes have been isolated from several eudicotyledonous model plants such as *Arabidopsis*, *Antirrhinum* and *Petunia*, and most of them belong to the family of MADS-box genes, encoding transcription factors containing a conserved 56 amino acid motif within their DNA-binding domains, named MADS after the four original members (for a review: see Ref. [5]). In *Arabidopsis thaliana* C-function, establishing identity of the reproductive organs, as well as determinacy of the floral meristem, is performed by the MADS-box gene *AGAMOUS* (*AG*) [6]. Moreover, *AG* functions in a redundant manner with the closely related genes *SHATTERPROOF* (*SHP*) 1/2 in the control of carpel development and with *SHP1/2* and *SEEDSTICK* (*STK*) in determining ovule identity [7].

In this manuscript we describe the cloning and characterization of a cDNA from flowers and fruits of the cocoa tree (*Theobroma cacao* L.), homologous to the C-type (*AG*) MADS-box gene. Results from expression analysis as well as from the phenotype caused by the ectopic expression in transgenic *Arabidopsis* plants suggest that the gene represented by this

* Corresponding author. Tel.: +31 317 475913; fax: +31 317 423110.

E-mail address: Ruud.deMaagd@wur.nl (R.A. de Maagd).

¹ Present address: Indonesian Biotechnology Research Institute for Estate Crops, Jl. Taman Kencana 1, Bogor 16151, Indonesia.

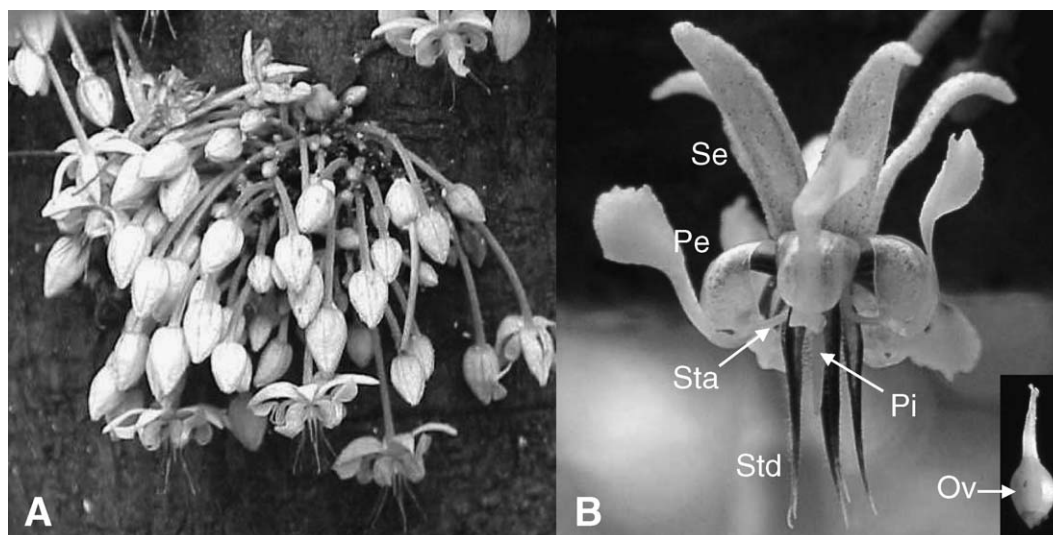


Fig. 1. (A) Flowering stem cushion of cocoa. (B) A cocoa flower with organs indicated: Se, sepal; Pe, petal; Sta, stamen; Std, staminode; Pi, pistil; Ov, ovary. The insert shows a dissected pistil.

cDNA has similar functions in regulating cocoa flower organ specification and the ability to induce flowering when overexpressed, as its homologue in *Arabidopsis*.

2. Materials and methods

2.1. Plant materials

The plant material was obtained from the cacao tree clone DR2, which was grown in fields of Rajamandala, PTPN VIII, Bandung, Indonesia. Fruits and flowers were separated into the different tissues or organs, immediately frozen in liquid nitrogen, and stored at -80°C until further use. Leaves were harvested from plants grown in the laboratory.

2.2. RNA extraction

The method used for RNA extraction from cocoa is a modification of a method described earlier for banana tissues [8]. After the extraction of RNA according to this method, the pellet was dissolved in MilliQ grade water and was extracted with phenol, phenol:chloroform:isoamyl alcohol (25:24:1) and chloroform:isoamyl alcohol (24:1), respectively, each time followed by centrifugation at $12,000 \times g$ for 15 min at 4°C . RNA in the supernatant was precipitated with 0.1 volumes of 3 M sodium acetate pH 5.8 and 3 volumes of 100% ethanol at -70°C for 4 h to overnight. The RNA was recovered by centrifugation at $17,000 \times g$ at 4°C for 30 min. The pellet was washed with an equal volume of 70% cold ethanol and after drying the purified RNA pellet was resuspended in water.

2.3. Isolation of pod wall-specific cDNA's by suppression subtractive hybridization (SSH)

mRNAs isolated from total RNA using oligo (dT)₂₅ Dynabeads (GenoVision, Philadelphia, PA, USA). Pod wall-specific cDNA-fragments were isolated using the PCR-Select

cDNA subtraction kit (CLONTECH, BD Biosciences, Alphen a/d Rijn). Starting material consisted of $1 \mu\text{g}$ inner pod wall mRNA as tester and $1 \mu\text{g}$ of leaf and bean mRNA (mixed equal amounts) as a driver. All SSH products were ligated into the pGEM-T Easy vector (Promega Benelux BV, Leiden) and transformed into JM109 super competent cells.

2.4. 5' RACE and amplification of the coding region of the cocoa homologue of AGAMOUS

For the selected SSH-fragment, corresponding 5' cDNA-fragments were amplified using the SMART RACE-kit from CLONTECH. 5'-Ready cDNA pools were produced from cocoa flower mRNA. The cDNA-specific primer TcAG-R, complementary to the sequence beyond the stop codon of SSH contig 24 was used in the 5' cDNA end amplification (Table 1). The PCR product was used as template in a nested PCR with the TcAG-R and NUP primers. Amplified cDNA-end fragments were purified using a QIAquick PCR purification kit (Qiagen, Benelux B.V., Venlo) and ligated into pGEM-T Easy.

To amplify a cDNA-fragment covering the full coding sequence of *TcAG*, $2 \mu\text{L}$ of the first strand cDNA from cocoa flower buds were used as a template for PCR using the kit HSRT-20 (Sigma Bioscience) with the primers TcAG-F (covering the presumed start codon and creating an *NcoI*-site) and TcAG-R (Table 1). The parameters used during the PCR reaction were as follows: 95°C , 2 min; 36 cycles: 95°C , 25 s;

Table 1
Oligonucleotide primers used in this study

Name	Nucleotide sequence (5' to 3')
TcAG-R	GGTGACCGTAGCACTTACTCCACCAGA
TcAG-F	CCATGGAGTACCAAAGTGAATCC
TcAG-F396	CGCATTGCCTATGAAGGATC
TC250F	CACCCTCGAGGTGGAAAGCTC
TC250R	CCACCACGGAGTCGCAACAAC

50 °C, 30 s; 70 °C, 2 min. An additional 5 min at 70 °C was added in the last step. The obtained PCR-products were and cloned as described above.

2.5. Sequence analysis

The sequencing reaction was performed on PCR-amplified fragments using Bigdye reaction mix (Applied Biosystems, Nieuwerkerk a/d IJssel, The Netherlands) and was analyzed in a ABI Prisma 3100 Genetic Analyzer (Perkin-Elmer, The Netherlands). Sequences were analyzed and assembled into contigs using the Lasergene SeqMan II software (DNASTAR Inc., Madison, WI, USA).

2.6. Expression analysis by reverse transcription polymerase chain reaction (RT-PCR)

First stand cDNA was synthesized from total RNA of cocoa tissues, using oligo (dT)₂₃ primer and Superscript Reverse Transcriptase (Invitrogen BV, Leek, The Netherlands). Two microlitres each of the first strand cDNA's were used as a template for PCR with gene specific primers. The primers for *TcAG* were *TcAG-F396* and *TcAG-R* (Table 1). As control for the equal efficiency of reverse transcription, a reaction was performed in parallel using primers (*TC250for* and *TC250rev*, Table 1) amplifying a 420 bp fragment of cocoa contig *TC250* (TIGR Cocoa Gene Index, http://www.tigr.org/tigr-scripts/tgi/T_reports.cgi?species=cocoa), encoding a putative polyubiquitin. The parameters used during PCR reactions were as follows: 94 °C, 4 min; 39 cycles: 95 °C, 30 s; 50 °C, 30 s; 70 °C, 1 min. An additional 5 min at 70 °C was added in the last step.

2.7. Plant vectors and transformation

The fragment containing the *TcAG* full open reading frame encoding cDNA was isolated by *NcoI* and *BstEII* digestion from pGEM-T Easy and ligated into the *NcoI*–*BstEII* sites of pCAMBIA2301, giving plasmid pTC22.

The plasmids pTC22 and pCAMBIA2301 (as negative control) were used for transformation into *Agrobacterium tumefaciens* strain Agl0 by electroporation, and then introduced into *A. thaliana* ecotype Columbia by the floral dip method [9]. Resulting *Arabidopsis* seeds were plated on 50% MS medium containing 50 mg/L kanamycin and 200 mg/L cefotaxim, for selection of transformed seedlings. After 3 days at 4 °C the plates were transferred to a growth chamber with long day conditions (16 h light, 24 °C). Kanamycin-resistant plants were transferred to soil and the plants were grown in the growth chamber for analysis of the flowering time and phenotype.

2.8. DIG-labeling of RNA and Northern blot analysis

To produce a template for in vitro transcription, 145 bp at the 5' end of the original 660 bp *TcAG* SSH clone insert were removed by digesting the pGEM-T Easy clone with *NdeI*, followed by re-ligation. The resulting 415 bp cDNA insert was amplified using the M13 Reverse and the T7 Promoter Primer

for pGEM-T Easy. The amplified fragment was then used as template for T7 RNA polymerase, with the DIG RNA in vitro transcription labeling kit (Roche Diagnostics GmbH, Mannheim, Germany).

For Northern blots, 1.5 µg of denatured total RNA from entire *Arabidopsis* plants was separated by electrophoresis in a 1% agarose gel containing 6% formaldehyde, blotted onto positively charged nylon membranes (Roche) by capillary transfer and hybridized with RNA probes overnight at 68 °C in DIG-easy Hyb buffer (Roche), washed and visualized according to the manufacturer's instructions.

3. Results

3.1. Isolation and sequence analysis of *TcAG* cDNA from cocoa

In an experiment to identify pod wall-specific cDNA's we produced, by suppression subtractive hybridization, a cDNA-fragment library resulting from a subtraction of pod wall cDNA as tester and mixed leaf + bean cDNA as driver. Subtracted cDNA was ligated into pGEM-T Easy and subjected to colony PCR, revealing inserts corresponding to cDNA-fragments of 250–1000 bp, of which 96 random fragments were sequenced.

Of the 96 cDNA-fragments sequenced initially, four constituted a contig of approximately 660 bp, encoding a peptide that in a database (BlastX) search showed considerable homology to *AGAMOUS*-like MADS-box proteins from various plant species. The cDNA-fragment corresponded to the part of the gene of its closest homologue, *Populus balsamifera* subsp. *trichocarpa* *PTAG1* encoding the C-terminal 2/3 of the protein, as well as 145 bp 3' of the putative open reading frame. In order to amplify a cDNA covering the entire open reading frame of the cocoa *AG* homologue, we used 5' RACE on flower cDNA using a 3' primer complementary to the region including and following the stop codon, as a cDNA-specific primer. The sequences of the cloned 5' cDNA-fragment and of the original SSH contig were indeed overlapping and could be assembled into a cDNA sequence representing an mRNA of 939 nucleotides (deposited in GenBank with accession number. [DQ157163](http://www.ncbi.nlm.nih.gov/Genbank/lookup/DQ157163)). The sequence coded for a protein, hence called *TcAG*, of 241 amino acids, of which the C-terminal 158 amino acids were identical to those encoded by the earlier identified SSH-fragments. The resulting protein has as its closest homologue in GenBank cotton, *Gossypium hirsutum* *AG*-like protein GhMADS-3 (86.3% amino acid identity), followed by the two black cottonwood, *Populus balsamifera* subsp. *trichocarpa* *AG* proteins, *PTAG1* and *PTAG2* (84.6% amino acid identity, for both), hazel nut, *Corylus avellana* CaMADS1 and peach, *Prunus persicae* PpMADS4. An amino acid alignment of *TcAG* with its closest homologues as well as with the more distant *Arabidopsis* *AG* and tomato *TAG1* proteins is shown in Fig. 2. Clearly recognizable and indicated are the strongly conserved MADS-box and somewhat lesser conserved K-box, while the C-terminus is the least conserved portion of the *AG* homologues. Both cotton and cocoa *AG* homologues have two in-frame putative start codons preceding the MADS-box, of which only the first is conserved in other *AG*

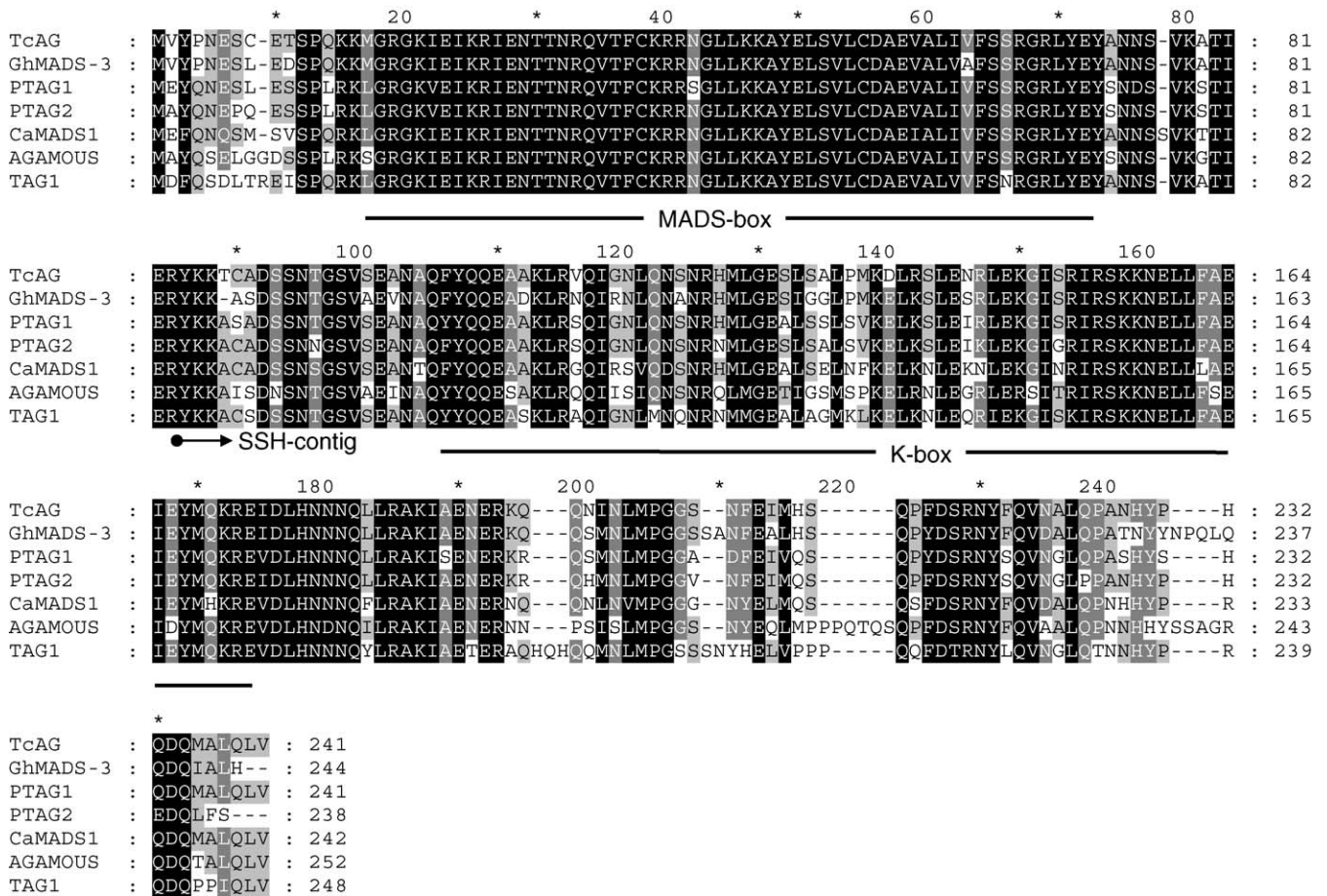


Fig. 2. Amino acid alignment of TcAG and its closest homologues (accession numbers in parentheses). GhMADS-3, *Gossypium hirsutum* (AAL92522); PTAG1 and PTAG2, *Populus balsamifera* subsp. *trichocarpa* (black cottonwood) (AAC06237 and AAC06238); CaMADS1, *Corylus avellana* (hazel nut) CaMADS1 (AAD03486); AGAMOUS, *Arabidopsis thaliana* AG (P17839); TAG1, *Lycopersicon esculentum* (tomato) (Q40168). Alignments were made with the ClustalW program and edited with the GeneDoc program (<http://www.psc.edu/biomed/genedoc>). The conserved MADS-box and K-box are underlined, and the start of the original SSH contig sequence is indicated.

homologues. If the first start codon is predominantly used in vivo, TcAG also contains an extension N-terminal to the MADS-box that is typical for AG and closely related proteins.

A. thaliana AG, together with STK and the SHP1/2 genes is part of a subfamily of MIKC-type genes, which likely represent a monophyletic clade, and whose members have partially redundant functions [10]. In order to more accurately position TcAG in this subfamily, we created a phylogenetic tree of a selection of AG, STK and SHP protein homologues from selected species (*Arabidopsis*, tomato, cotton and poplar) (Fig. 3). As can be seen in this figure, TcAG is firmly embedded in the AG subclade, closest to a cotton AG-like protein GhMADS-3, and to the two AG proteins from poplar, while clearly distinct from the SHP and STK subclades.

3.2. TcAG expression pattern in cocoa tissues

To gain insight into the developmental role of TcAG, its expression pattern was analyzed by RT-PCR, using primers designed to amplify the less-conserved 3' region, on various vegetative and reproductive tissues of cocoa trees: leaf, shoot-

producing stem cushions, passive and actively flowering cushion tissue, flower buds, floral organs (sepals, petals, staminodes, stamens and pistils) and developing (young, middle and mature) cocoa pod wall, mature pod pulp and beans. The results (Fig. 4) revealed that the transcript was not detectable in vegetative tissues (leaf and vegetative shoot-producing stem cushion), nor in passive or actively flowering cushion tissue, but only first detectable in flower buds. Within open flowers, transcript levels were highest in the organs of the inner two whorls (stamens and pistil) while lower in staminodes and barely detectable in sepals and petals. TcAG transcript was also clearly detectable in all stages of the developing pod wall, as well as in the pulp of a ripe fruit. This expression pattern suggested that this gene is involved in the development of reproductive floral organs since it is expressed in pistils and stamens, while expression is almost undetectable in sepals and petals of the cocoa flower.

3.3. Ectopic expression of TcAG in Arabidopsis

The function of TcAG was further investigated by ectopic expression in *Arabidopsis* in order to determine whether the

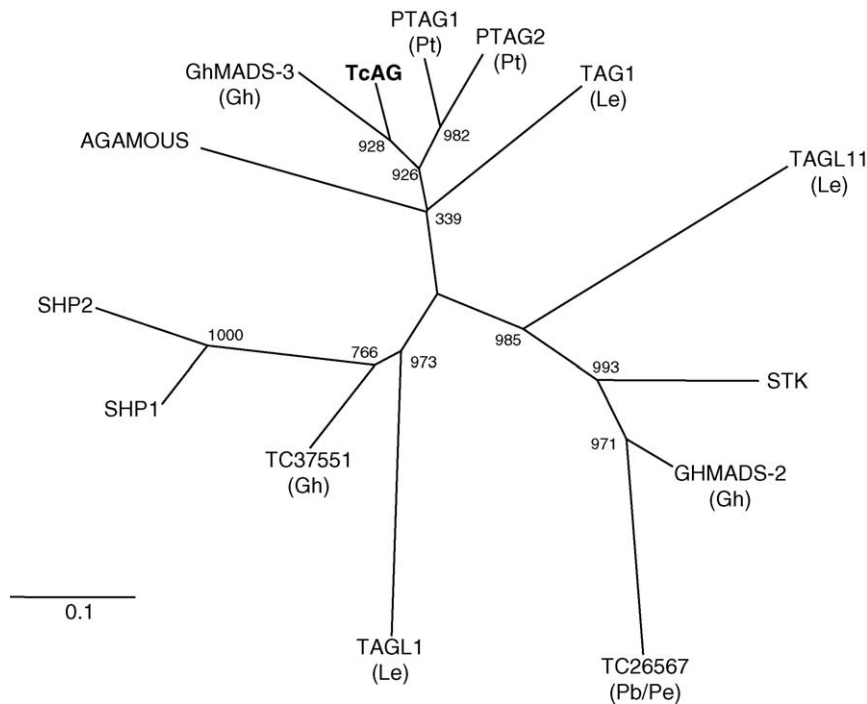


Fig. 3. Phylogenetic tree of TcAG and related MADS-box proteins, comprising the subclades represented by AG, STK and SHP from *Arabidopsis*. Also included in the analysis were, as far as available, the closest homologues of the *A. thaliana* SHP and STK proteins in cotton and poplar, respectively. Protein sequences were taken from GenBank, unless these were not found or when translated tentative contigs from the TIGR gene index (labeled TC) displayed higher homology. Phylogenetic trees were produced with the ClustalX package [26]. Phylogenetic reconstruction was obtained by the neighbor-joining method [27] together with bootstrap analysis using 1000 replicates. Kimura correction for multiple substitutions was applied [28]. The tree was visualized using the TreeView package [29]. All non-*Arabidopsis* sequences, with the exception of TcAG, have the species indicated between parentheses. The accession numbers of the additional protein sequences: *Arabidopsis thaliana* SHP1/AGL1 (P29381); SHP2/AGL5 (P29385); STK/AGL11 (Q38836); *Lycopersicon esculentum* (Le) TAGL1 (AAM33101); TAGL11 (AAM33102); *Gossypium hirsutum* (Gh) GHMADS-2 (AAN15183); *Populus balsamifera* subsp. *trichocarpa* (Pt); *Prunus persicae* PpMADS4 (AAU29513). Additionally, amino acid translations were used for two tentative contig sequences from the TIGR poplar (Pb/Pe, TC26567) and cotton (Gh, TC37551) gene indices, encoding AGL11 and SHP homologues, respectively.

sequence and expression pattern similarities between *TcAG* and *AG* also indicate a functional relationship. A binary vector containing 35S::*TcAG*, as well as a kanamycin-resistance gene was used in *A. tumefaciens*-mediated transformation of *Arabidopsis* plants by the floral dip method. Following transformation, transgenic seeds were selected on kanamycin and surviving plants transferred to soil and observed.

Out of 62 independent kanamycin-resistant T1-plants, 25 exhibited visible phenotypic alterations. According to the ABC model of flower development, transgenic plants overexpressing *AG* are expected to show homeotic modifications in the first and second whorl of the flower, similar to an *ap2*-mutant phenotype, due to the negative effects of C-function on A-function (*AP2*) in these whorls [11]. Plants showing homeotic changes were classified as having weak or strong phenotypes. Plants with

strong phenotypes were smaller and had small and curled leaves (compare Fig. 5A and B). Plants with weak phenotypes had flowers with reduced size organs in the outer two whorls (sepals and petals) as well as curly cauline leaves (compare Fig. 5C and D). Flowers in plants with strong phenotypes lacked discernable petals and the sepals were reduced in size and had carpel-like features, such as stigmatic papillae at the apex and ovule-like structures at the margins (Fig. 5E). The latter structures were also occasionally observed on modified cauline leaves (Fig. 5F), which has never been observed in other studies with *Arabidopsis* plants overexpressing *AG* homologues. Many inflorescences terminated early in carpelloid structures (Fig. 5E and F). *TcAG* expressing plants, including some with no visible phenotypic alterations in flowers, appeared to flower earlier than plants transformed with the empty vector (results not shown).

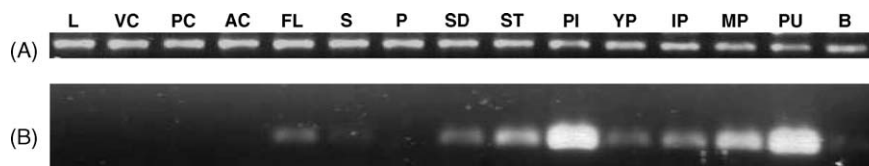


Fig. 4. RT-PCR products from cocoa RNA. (A) Products produced with polyubiquitin primers, used as control. (B) Products of RT-PCR with *TcAG*-specific primers. Vegetative tissues (L, leaf; VC, vegetative shoot-producing stem cushion; PC, passive flowering cushions) and reproductive tissues (AC, actively flowering cushions; FL, flower buds), floral organs (S, sepals; P, petals; SD, staminodes; ST, stamen; PI, pistils), three stages of the developing inner pod wall (YP, young pod; IP, intermediary pod; MP, mature pod) and PU, pulp of mature pod; B, bean.

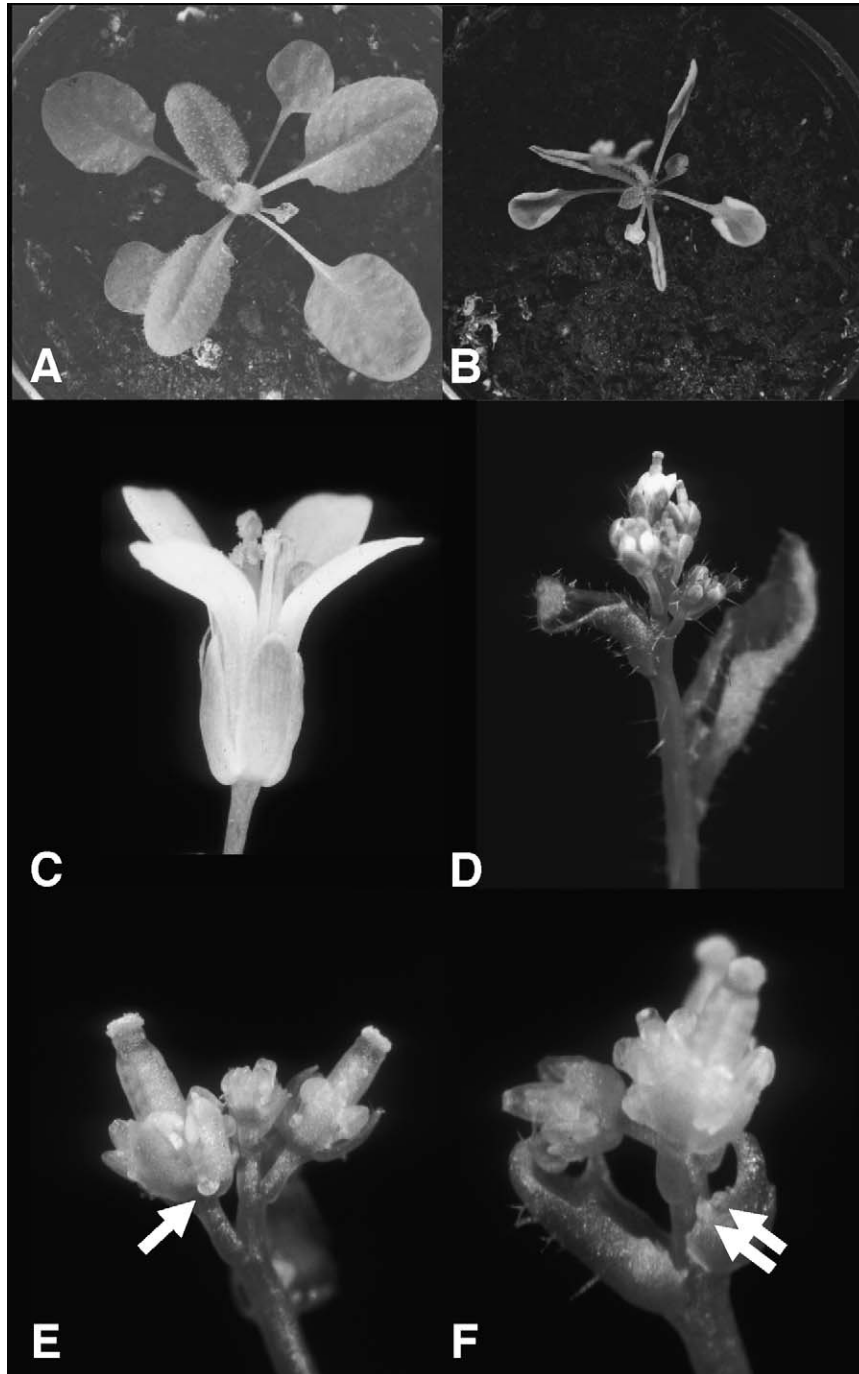


Fig. 5. Vegetative and floral morphology of *Arabidopsis thaliana*. (A) Wild type plant showing rosette leaves. (B) Strong *TcAG* overexpression phenotype showing decreased size and curled rosette leaves. (C) Wild type flower. (D) Weak *TcAG* overexpression flower phenotype, showing reduced sepal and petal size. (E) Strong overexpression flower phenotype with absent petals and carpel-like sepals. Arrow indicates ovule-like structure on the sepal edge. The inflorescence contains only two recognizable flowers and ends in a group of carpelloid structures. (F) Strong *TcAG* overexpression phenotype showing a carpel-like cauline leaf with ovule-like structures (arrows).

Northern blot analysis of the level of *TcAG* expression in the transgenic *Arabidopsis* plants showed an overall positive correlation between *TcAG* mRNA level and severity of the *ap2*-like phenotype. A negative control plant and a transformed plant without visible phenotype had no detectable *TcAG* mRNA (Fig. 6, lanes 1 and 2), while plants with weak phenotype (Fig. 6, lanes 3–6) and strong phenotype (Fig. 6, lanes 7–10) had increasing amount of *TcAG* transcripts.

4. Discussion

In this manuscript, we describe for the first time the cloning and characterization of a MADS-box type transcription factor gene involved in flower development of a tropical tree species, cocoa (*Theobroma cocoa* L.). In several aspects regulation of the onset of flowering as well as the floral architecture in cocoa are substantially different from other plant species, in which the

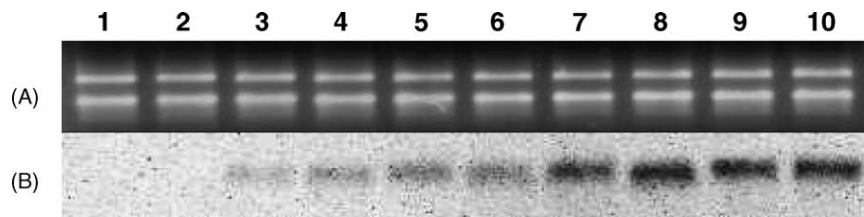


Fig. 6. Northern blot analysis of *TcAG* expression in *Arabidopsis*. (A) Gel electrophoresis of total RNA. (B) Northern blot of total RNA probed with a *TcAG* DIG-labeled RNA probe. Lane 1, negative control, plant transformed with pCAMBIA2302. Lanes 2–10, plants transformed with pTC22 (*TcAG*); lane 2, transformed plant with no visible phenotype. Lanes 3–6, plants with weak phenotype; lanes 7–10, plants with strong phenotype.

role of MADS-box genes during flowering was studied so far. Therefore, the study of the role of MADS-box genes during flowering of cocoa may yield important insights in the mechanisms leading to the evolutionary diversity of floral architecture and the regulation of the onset of flowering.

Based on the sequence of a cDNA-fragment from a library of subtracted pod wall-specific cDNA's, we amplified and cloned a full-length cDNA representing an *AG* homologue (*TcAG*) from flowers of cocoa. Phylogenetic analysis of this and other *AG* protein sequences (Fig. 3) placed *TcAG* clearly in the *AG* subclade rather than in one of the two closely related subclades represented by the *Arabidopsis* STK and SHP proteins, respectively. Identification of *TcAG* as a true functional homologue of *AG* was further supported by its expression pattern in the flower. *TcAG* was typically expressed in the so-called C-domain of the flower: the inner whorls containing the stamens and the pistil. The occurrence of highly specialized staminodes is an interesting aspect of cocoa flower architecture. Staminodes are defined as stamen having lost the function of pollen production in evolution and are, although relatively uncommon, widely distributed taxonomically [12]. Staminodes may occur simply as reduced stamen or stamen remnants, but often they have acquired new functions during evolution, such as in attracting pollinators or in regulating pollination [12]. In cocoa they are likely to function in the prevention of self-pollination of the flower by typical dipteran pollinators (small midges). After opening the flower has a male phase, in which the staminodes surround the style and protect it against contact with the pollinators. During the following female phase the staminodes bend outwards and the midges can land on them, contacting the style [13]. The staminodes of the cocoa flower are likely to have developed from the outer of two whorls of stamens. The expression of *TcAG* in the staminodes as well as in the stamens supports their common ontology.

Expression of *TcAG* was also found in the pod wall throughout its development, as well as, to a high level, in the pulp surrounding the beans. The pulp is probably derived from the seed testa, although the published evidence for this is scarce [14]. *AG* in *Arabidopsis* as well as *TAG1* in tomato are expressed in ovule integuments [15,16], but little is known of their expression in later stages. Expression of *TAG1* in developing tomato fruit occurs particularly in the turning and red ripe stages [17], but no differentiation was made between seed and pericarp, whereas in the initial stages after fertilization *TAG1* expression in carpels diminished [18]. *AG* in *Arabidopsis* is expressed during the early stages of silique

development but drops off in the last stages [19]. In grapevine, an *AG* homologue is also expressed in the skin of the berry during development up to the ripening stage, when it drops off [20]. Thus it is likely that *AG* orthologues are expressed in the pericarp during some stages of fruit development in *Arabidopsis* and tomato, as well as in cocoa and other fruits. The precise tissue localization of this expression and its role in fruit development will be interesting to study.

The importance of the MADS-box genes in determining floral organ identity is reflected by their strong evolutionary conservation of function and of interactions with other transcription factors of the MADS family [21]. Ectopic expression of orthologues from other species in *Arabidopsis* often resulted in producing a similar phenotype as ectopic expression of the *Arabidopsis AG* gene would [11]: severely affected flowers have carpelloid first whorl organs and missing or staminoid second whorl organs, while less affected flowers have small petals [22–25]. Other phenotypes displayed are curled leaves and early termination of the inflorescence in carpelloid structures. Our observation of very similar phenotypes in *TcAG*-expressing transgenic *Arabidopsis* plants, together with the pattern of *TcAG* expression in cocoa floral organs and the phylogenetic study of *AG* homologues supports our hypothesis that *TcAG* is a true orthologue of *AG*.

Acknowledgements

T. Chaidamsari, D. Santoso, H. Sugiarti and Samanhudi were supported by the BIORIN Programme (Scientific Programme Indonesia–The Netherlands) of the Royal Netherlands Academy of Arts and sciences (KNAW), and by an Indonesian International Joint Research Grant (RUTI) of the Ministry of Research and Technology of the Republic of Indonesia. We thank Stefan de Folter (PRI) for assistance with photography and helpful discussions, and both him as well as Richard Immink for critically reading the manuscript.

References

- [1] R. Lent, The origin of cauliflorous inflorescence of *Theobroma cacao*, Turrialba 16 (1966) 352–358.
- [2] H. Toxopeus, Botany, types and populations, in: G. Wood, R. Lass (Eds.), *Cacao*, vol. 4, Longman Scientific and Technical, 1987, pp. 11–37.
- [3] Z. Schwarz-Sommer, P. Huijser, W. Nacken, H. Saedler, H. Sommer, Genetic control of flower development: homeotic genes of *Antirrhinum majus*, Science 250 (1990) 931–936.

- [4] E.S. Coen, E.M. Meyerowitz, The war of the whorls: genetic interactions controlling flower development, *Nature* 353 (1991) 31–37.
- [5] S. Ferrario, R.G. Immink, G.C. Angenent, Conservation and diversity in flower land, *Curr. Opin. Plant Biol.* 7 (2004) 84–91.
- [6] M.F. Yanofsky, H. Ma, J.L. Bowman, G.N. Drews, K.A. Feldmann, E.M. Meyerowitz, The protein encoded by the *Arabidopsis* homeotic gene *agamous* resembles transcription factors, *Nature* 346 (1990) 35–39.
- [7] A. Pinyopich, G.S. Ditta, B. Savidge, S.J. Liljegren, E. Baumann, E. Wisman, M.F. Yanofsky, Assessing the redundancy of MADS-box genes during carpel and ovule development, *Nature* 424 (2003) 85–88.
- [8] M.H. Asif, P. Dhawan, P. Nath, A simple procedure for the isolation of high quality RNA from banana fruit tissue, *Plant Mol. Biol. Rep.* 18 (2000) 1–7.
- [9] S.J. Clough, A.F. Bent, Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*, *Plant J.* 16 (1998) 735–743.
- [10] A. Becker, G. Theissen, The major clades of MADS-box genes and their role in the development and evolution of flowering plants, *Mol. Phylog. Evol.* 29 (2003) 464–489.
- [11] Y. Mizukami, H. Ma, Ectopic expression of the floral homeotic gene *AGAMOUS* in transgenic *Arabidopsis* plants alters floral organ identity, *Cell* 71 (1992) 119–131.
- [12] J. Walker-Larsen, L.D. Harder, The evolution of staminodes in angiosperms: patterns of stamen reduction, loss, and functional re-invention, *Am. J. Bot.* 87 (2000) 1367–1384.
- [13] A.M. Young, E.H. Erickson Jr., M.A. Strand, B.J. Erickson, Pollination biology of *Theobroma* and *Herrania* (Sterculiaceae). I. Floral biology, *Insect Sci. Appl.* 8 (1987) 151–164.
- [14] R. Nichols, Studies of fruit development of cacao (*Theobroma cacao*) in relation to cherrille wilt. I. Development of pericarp, *Ann. Bot.* 28 (1964) 619–634.
- [15] J.L. Bowman, G.N. Drews, E.M. Meyerowitz, Expression of the *Arabidopsis* floral homeotic gene *AGAMOUS* is restricted to specific cell types late in flower development, *Plant Cell* 3 (1991) 749–758.
- [16] L. Pnueli, D. Hareven, S.D. Rounsley, M.F. Yanofsky, E. Lifschitz, Isolation of the tomato *AGAMOUS* gene *TAG1* and analysis of its homeotic role in transgenic plants, *Plant Cell* 6 (1994) 163–173.
- [17] G.E. Bartley, B.K. Ishida, Digital fruit ripening: data mining in the TIGR tomato gene index, *Plant Mol. Biol. Rep.* 20 (2002) 115–130.
- [18] M. Busi, C. Bustamante, C. D'Angelo, M. Hidalgo-Cuevas, S. Boggio, E. Valle, E. Zabaleta, MADS-box genes expressed during tomato seed and fruit development, *Plant Mol. Biol.* 52 (2003) 801–815.
- [19] S. de Folter, J. Busscher, L. Colombo, A. Losa, G.C. Angenent, Transcript profiling of transcription factor genes during silique development in *Arabidopsis*, *Plant Mol. Biol.* 56 (2004) 351–366.
- [20] P.K. Boss, M. Vivier, S. Matsumoto, I.B. Dry, M.R. Thomas, A cDNA from grapevine (*Vitis vinifera* L.), which shows homology to *AGAMOUS* and *SHATTERPROOF*, is not only expressed in flowers but also throughout berry development, *Plant Mol. Biol.* 45 (2001) 541–553.
- [21] R.G.H. Immink, G.C. Angenent, Transcription factors do it together: the hows and whys of studying protein–protein interactions, *Trends Plant Sci.* 7 (2002) 531–534.
- [22] V.A. Benedito, P.B. Visser, J.M. van Tuyl, G.C. Angenent, S.C. de Vries, F.A. Krens, Ectopic expression of *LLAG1*, an *AGAMOUS* homologue from lily (*Lilium longiflorum* Thunb.) causes floral homeotic modifications in *Arabidopsis*, *J. Exp. Bot.* 55 (2004) 1391–1399.
- [23] K. Kitahara, Y. Hibino, R. Aida, S. Matsumoto, Ectopic expression of the rose *AGAMOUS*-like MADS-box genes 'MASAKO C1 and D1' causes similar homeotic transformation of sepal and petal in *Arabidopsis* and sepal in *Torenia*, *Plant Sci.* 166 (2004) 1245–1252.
- [24] J. Lemmetyinen, M. Hassinen, A. Elo, I. Porali, K. Keinonen, H. Makela, T. Sopanen, Functional characterization of *SEPALLATA3* and *AGAMOUS* orthologues in silver birch, *Physiol. Plant.* 121 (2004) 149–162.
- [25] R. Rutledge, S. Regan, O. Nicolas, P. Fobert, C. Cote, W. Bosnich, C. Kauffeldt, G. Sunohara, A. Seguin, D. Stewart, Characterization of an *AGAMOUS* homologue from the conifer black spruce (*Picea mariana*) that produces floral homeotic conversions when expressed in *Arabidopsis*, *Plant J.* 15 (1998) 625–634.
- [26] J.D. Thompson, T.J. Gibson, F. Plewniak, F. Jeanmougin, D.G. Higgins, The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, *Nucleic Acids Res.* 25 (1997) 4876–4882.
- [27] N. Saitou, M. Nei, The neighbor-joining method: a new method for reconstructing phylogenetic trees, *Mol. Biol. Evol.* 4 (1987) 406–425.
- [28] M. Kimura, *The Neutral Theory of Molecular Evolution*, Cambridge, England, 1983.
- [29] R.D. Page, TreeView: an application to display phylogenetic trees on personal computers, *Comp. Appl. Biosci.* 12 (1996) 357–358.