2011 International Conference on Bioscience, Biochemistry and Bioinformatics
IPCBEE vol.5 (2011) © (2011) IACSIT Press, Singapore

Isolation of an Agamous Homologue from *Jatropha Curcas*

Poonyaporn Borivetanan
Department of Genetics
Kasetsart University
Bangkok, Thailand
e-mail: mydear_dragon@hotmail.com

Sompid Samipak
Department of Genetics
Kasetsart University
Bangkok, Thailand
e-mail: ssamipak@gmail.com

***Abstract*— *Jatropha curcas* is a drought-tolerant biofuel plant whose seed oil can be used directly for making biodiesel in several countries. One common problem for *J. curcas* growers is its low yield mainly due to lack of female flowers. AGAMOUS (AG) is a floral organ identity gene in C class of ABC model which specifies stamen and carpel development. To facilitate future plant improvement, in this study, we cloned *JcAg* which is an AGAMOUS-like gene of this physic nut and reported its nucleotide sequence with a predicted amino acid sequence. Confirming by protein database search (Protein Blast), *JcAg* sequence has 87% identity with *TcAG* (AGAMOUS-like protein of cocoa, *Theobroma cacao*) and possesses the MADS-box domain and K-box domain characteristic of this gene. According to phylogenetic analysis, *JcAg* is firmly situated in AGAMOUS subclade closest to *TcAG* and is clearly separated from SHATTERPROOF (*SHP*) and SEEDSTICK (*STK*) subclades.**

Keywords-Agamous, Jatropha curcas, flower determination

I. INTRODUCTION

Jatropha curcas is a perennial, monoecious small tree that can grow up to 6 meter high and well adapted to arid and semi-arid conditions. It is a high potential biofuel plant whose seeds can be mechanically squeezed to yield oil that can be mixed in several engine types. One disadvantage of this plant is low ratio of female flower to male flower (approximate 10 flowers: 1 fruit) resulting in low seed yield per plant. Flower organ is crucial for plant reproduction and has long been studied. The genetic control of floral organ identity is explained by the ABC model with *AGAMOUS* being a gene in C class that specifies stamen and carpel development [1,2]. AGAMOUS's functions are related to the closest *AGAMOUS* paralog genes, *SHATTERPROOF* (*SHP*) 1 and 2, in control of carpel development and *SEEDSTICK* (*STK*) in determining ovule identity [3]. In this study we describe steps in cloning and characterization of an *AGAMOUS*-like gene involved in the control of flower development in *J. curcas.*

II. MATERIALS AND METHODS

A. Plant materials

All plant materials were obtained from the *Jatropha curcas* cv Anna 16: small young flower buds (1-3 mm) are excised from *J. curcas* tree and were immediately frozen in liquid nitrogen and stored at -20 °C until further use.

B. RNA extraction and cDNA synthesis

Total RNA was extracted from flower buds using the PureLink Plant RNA Extraction Reagent (Invitrogen, USA). Then mRNA was isolated from total RNA using Dynabeads® mRNA DIRECT™ Micro Kit (Invitrogen, USA). cDNA synthesis was performed using SuperScript™ III First Strand Kits (Invitrogen, USA). To design the degenerate primers, 4 published AGAMOUS-like protein sequences from selected plant species (*Cucumis sativus*, *Petunia x hybrida*, *Zea mays* and *Rosa rugosa*) were retrieved from GenBank and aligned with the Crustal W program. A pair of degenerate primers (forward primer JcAGF1 and reverse primer JcAGR1) was designed to amplify an internal fragment spanning part of the MADS-box. The parameters used during the PCR reaction were as followed: 94 °C, 5 min; then 6 cycles of 94 °C, 1 min; 35 °C, 2 min and 72 °C 1 min, followed by 25 cycles of 94 °C, 1 min; 45 °C, 2 min and 72 °C 1 min, with a final extension of 10 min at 72 °C. All primer sequences are listed in table I.

TABLE I. OLIGONUCLEOTIDE PRIMERS USED IN THIS STUDY

Name	Type	Nucleotide sequence (5' to 3')
JcAGF1	Forward primer	GCGCGAATTCGGNAARATHG ARATHAA
JcAGR1	Reverse primer	GCGCAAGCTTTTYTGCATRT ANTCDAT
nJcAgF	Forward primer	AATCTCTCGGAGGCATTAAT
AG2degR	Reverse primer	AACAAGTTGARRARMHRKTT GATC
NotI12T	Reverse primer	GAGCGGCCGCTTTTTTTTTTT T
nAgR5	Reverse primer	GCATACTCATAGAGGCGACC GCGGC
JCAGA2	Reverse primer	AGAGCAACCTCAGCATCACA T

C. 3' end fragment of the AGAMOUS-like gene in J. curcas

To obtain 3' end fragment, nJcAgF primer (designed from obtained sequence) and AG2degR (degenerate primer) were designed from 10 published AGAMOUS-like protein sequences (*Theobroma cacao*, *Vitis vinifera*, *Citrus unshiu*, *Populus balsamifera*, *Coryus avellana*, *Prunus mume*, *Glycine max*, *Hydrangea macrophylla*, *Betula pendula* and

Gossypium hirsutum). The parameters used during the PCR reaction were as followed: 94 °C, 5 min; then 30 cycles of 94 °C, 30 s; 50 °C, 30 s and 72 °C 1 min, with a final extension of 10 min at 72 °C. The other reverse primer NotI12T was used with forward primer nJcAgF to perform PCR reaction with the AG2degR. PCR products were purified and sequenced with forward primer nJcAgF.

D. 5' RACE to obtain 5' end fragment of the AGAMOUS like gene in J. curcas

To obtain 5' end fragment, RACE strategy was done using SMARTer™ RACE cDNA Amplification Kit (Clonetech, USA) with nAgR5 primer. The PCR product was used as template to perform a nested PCR with a new reverse primer, JCAGA2. This product was ligated into the pGEM®-T easy vector (Promega, USA) and transformed into *Escherichia coli* DH5α. The plasmids from selected clones were isolated and sequenced.

E. Sequence and phylogenetic analysis

All obtained nucleotide sequences were used to assemble together to obtain a contig and predicted amino acid sequence. The predicted amino acid sequence was aligned with 8 AGAMOUS-like and 6 relative protein sequences from other plant species (Table II) published in GenBank by using the Clustal W program. Sequences were imported into MEGA 4.0 software for phylogenetic analysis and the phylogenetic tree was constructed using Neighbor-Joining method combined with JTT Matrix (Jones-Taylor-Thornton) model. Tree nodes were evaluated by bootstrap analysis for 5000 replicates.

III. RESULTS, DISCUSSION AND CONCLUSION

To obtain *AGAMOUS*-like gene involved in flower development of *J. curcas*, degenerate primers were designed using conserved regions of AGAMOUS-like proteins in 4 plant species. These primers were used in PCR reaction with cDNA of *J. curcas* cv Anna 16 isolated from young flower bud. The obtained sequence was approximately 500 bp in length and highly similar to the Agamous-like proteins from various plant species in database (BLAST X) search results.

Based on the sequence obtained, primers were designed to amplify 3' end and 5' end fragment respectively. For 3' end fragment, degenerate reverse primer (AG2degR) was designed using 10 reported AGAMOUS-like proteins sequences and an approximate 300 bp fragment was obtained. For 3' UTR fragment, reverse primer (NotI12T) was used in PCR reaction and an approximate 600 bp fragment was obtained. For the 5' end, an approximate 500 bp fragment upstream of the sequence was obtained from 5' RACE reaction.

A contig from assembled sequences, here after called *JcAg*, was 1,171 bp in length with 726 bp open reading frame corresponding to 241 deduced amino acid residues (Fig. 2). Protein database search (Protein Blast) revealed that the deduced protein sequence showed 87%, 86% and 80% identity with AGAMOUS-like protein of *Theobroma cacao*, *Mangifera indica* and *Prunus mume* respectively. Sequence analysis of *JcAg* revealed that MADS-box domain and K-box domain were present (Fig. 3).

TABLE II. AGAMOUS-LIKE AND RELATIVE PROTEINS USED IN THIS STUDY

Organism	Name
Arabidopsis thaliana	AG, SHP1 (AGL1), SHP2 (AGL5), STK (AGL11)
Dendrobium thyrsiflorum	AG1, SEEDSTICK-like protein (STKL)
Nymphaea sp.	AG
Populus balsamifera subsp. trichocarpa	PTAG1, PTAG2
Prunus persica	AG, SEEDSTICK-like protein (STKL), SHATTERPROOF-like protein (SHPL)
Solanum lycopersicum	TAG1
Theobroma cacao	AG
Zea mays	AG

TCATCTTGCTTCCATTTTCTGCATCTCTCCTACTCAGATTT
GTAGAAACAAAGAACTGAGAAACCCACAACCCAAAAAG
GCCTTTGTTTCTCTCCTCAATTAGCATCTCTACTTCTCCTT
TCCTCTTCACTTTGTTTATTTCCTTACCAAGCTAGAAACAG
CTGCC

```
167 atggcataccccagcgattcccgggagacttcaccgcagaggaga
    M  A  Y  P  S  D  S  R  E  T  S  P  Q  R  R
212 atgggtaggggaaagatcgagatcaagcggatcgaaaacaccaca
    M  G  R  G  K  I  E  I  K  R  I  E  N  T  T
257 aatcgccaagtcactttctgcaaaagaagaaatggtttgctcaag
    N  R  Q  V  T  F  C  K  R  R  N  G  L  L  K
302 aaagcctatgaattatctgttttatgtgatgctgaggttgctctc
    K  A  Y  E  L  S  V  L  C  D  A  E  V  A  L
347 atcgtattctctagccgcggtcgcctctatgagtatgctaataat
    I  V  F  S  S  R  G  R  L  Y  E  Y  A  N  N
392 agtgttaaatctacaattgagaggtacaagaaagcatgtgcagat
    S  V  K  S  T  I  E  R  Y  K  K  A  C  A  D
437 tcatcaaatactggatctgtttctgaagctaatgcccagttctat
    S  S  N  T  G  S  V  S  E  A  N  A  Q  F  Y
482 cagcaacaagctgccaagctgcgtgatcaaattagcggcttgcag
    Q  Q  Q  A  A  K  L  R  D  Q  I  S  G  L  Q
527 aaatcaatcaggaacatgctgggtgaatctctcggaggcattaat
    K  S  I  R  N  M  L  G  E  S  L  G  G  I  N
572 cccaaggaccttaggggcttggagagcaggctagagaaaggaatt
    P  K  D  L  R  G  L  E  S  R  L  E  K  G  I
617 agtagaattcggtccaaaaagaatgagctgttgtttgcagagatc
    S  R  I  R  S  K  K  N  E  L  L  F  A  E  I
662 gagtacatgcaaaaagggaaattgatttgcacaataataaccag
    E  Y  M  Q  K  R  E  I  D  L  H  N  N  N  Q
707 cttctccgagcgaagattgctgagaatgaaaggaagcaacagaac
    L  L  R  A  K  I  A  E  N  E  R  K  Q  Q  N
752 atgaatctgatgccaggaggaggtaactatgagataattcagtct
    M  N  L  M  P  G  G  G  N  Y  E  I  I  Q  S
797 cagccatttgacaatcgaaactattttcaagtcaatgcattacaa
    Q  P  F  D  N  R  N  Y  F  Q  V  N  A  L  Q
842 cccaccaatcattatccacaacaagatcaaatggccctccagtta
    P  T  N  H  Y  P  Q  Q  D  Q  M  A  L  Q  L
887 gtttaa 892
    V  *
```

TAAGCTTTGAGAGTGAGCATCAACTTCTTGCCCTCTATTA
GTCTCTGTAGATCAACCTCAGGTATTTCACCTTCTGAAAG
CTGCAAGTATATATCTATATATACCAAAAATTTCTTGCAG
AAATCAAGTTTGCCTGAGAACTGGCTGCGCTAGAAATGT
GCTATCCACCATGTGCTAGAAATGTGTAATCCAACTTAAG
ACATTTAAGTATTGTTGTAAGAGAAGGATTTCTTTTATGG
ATGTTTAAAACTACTATGCCATCATCACTAGCTTCATCTA

Figure 2. Nucleotide and deduced amino acid sequence of the JcAg cDNA. The putative 5' and 3' untranslated regions are shown before and after of the open reading frame.

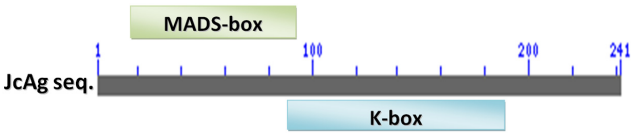

Figure 3. MADS-box and K-box locations in the deduced amino acid sequence *(JcAg)*

A. thaliana AG, together with *STK*, *SHP1* and *SHP2* genes are part of subfamily of MIKC-type genes, which likely represent a monophyletic clade, and whose members have partially redundant functions [4]. To place a more accurate position of *JcAg* in this subfamily, phylogeny reconstruction was performed with the predicted JcAg amino acid sequence, along with AG STK and SHP homologue proteins from selected species (Table 2). The result confirms that *JcAg* is situated in AGAMOUS subclade closest to *T. cacao* AGAMOUS-like protein (TcAG) and two *P. balsamifera* AGAMOUS homologue proteins (PTAG1 and 2) which clearly separates from the SHP and STK subclades (Fig. 4). Moreover, this phylogenetic tree shows a monophyletic of AGAMOUS-like protein from monocotyledon plants, separated from dicotyledon branch containing AGAMOUS-like protein of *Nymphaea sp.* This result indicates that *Nymphaea sp.* is closely related to monocotyledon plant and may reveal the connection between mono- and dicotyledon class. However, more data is needed to confirm this hypothesis.

In conclusion, we have isolated an *AGAMOUS* homologue in *J. curcas*, *JcAg*, and predicted an amino acid sequence. From Protein Blast, it showed 87% identity to AGAMOUS-like protein of *T. cacao*. It posses the two protein domains, MADS- box and K-box which are special characters of an AGAMOUS protein [5]. Phylogeny study showed the obtained sequence to be placed firmly in the dicotyledon AGAMOUS group. This *JcAg* sequence information can be used in future genetic manipulation to improve quality and yield of this
biofuel plant.

ACKNOWLEDGMENT

We thank Subhisa Kharuwannapat for her technical help. We are grateful for Dr. Vipha Hongtrakul and Ms Siriwan Burikam for their great advices. This work was supported by funding from Kasetsart University Research and Development Institute, Faculty of Science at Kasetsart University, and partially from Thailand Research Fund.

REFERENCES

[1] J. L. Bowman, D. R. Smyth and E. M. Meyerowitz, "Genetic interactions among floral homeotic genes of *Arabidopsis*," Development, vol. 112, 1991, pp. 1-20.

[2] T. Jack, "New members of the floral organ identity *AGAMOUS* pathway," TRENDS in Plant Science, vol. 7, 2002, pp. 286-287.

[3] R. L. Hong, L. hamaguchi, M.A. Busch and D. Weigel. "Regulatory elements of the floral homeotic gene *AGAMOUS* identified by phylogenetic footprinting and shadowing," The Plant Cell, vol. 15, 2003, pp. 1296-1309.

[4] A. Becker and G. Theissen. "The major clades of MADS-box genes and their role in the development and evolution of flowering plants," Mol. Phylogen. Evol**,** vol. 29, 2003, pp. 464-489.

[5] D. Thakare, W. Tang, K. Hill and S.E. Perry. "The MADS-domain transcriptional regulator *AGAMOUS-LIKE15* promotes somatic embryo development in *Arabidopsis* and soybea," Plant Physiol, vol. 146, 2008, pp. 1663-1672.

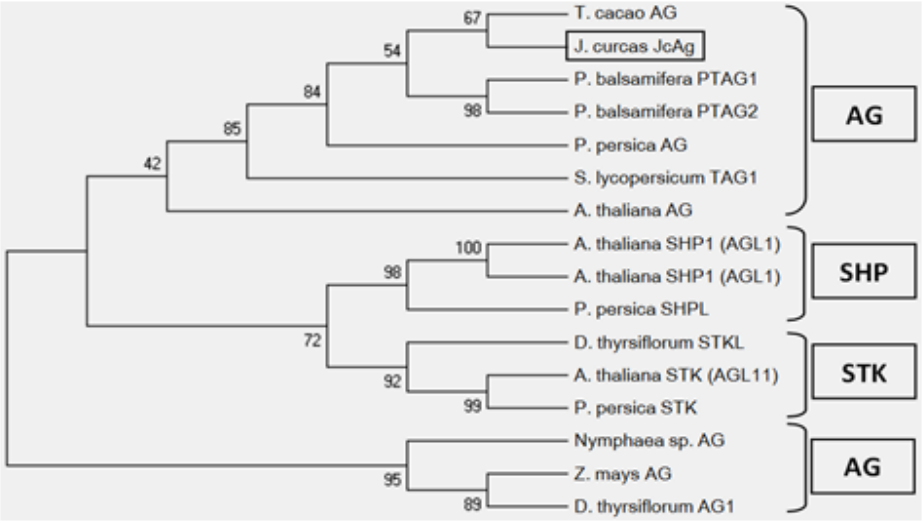

Figure 3. Phylogenetic reconstruction (amino acid sequence data). Neighbor-Joining tree derived from predicted JcAg (J. curcas) and 8 AGAMOUS-like protein sequences addition with 6 relative protein sequences (STK and SHP) from other plant species (Table II) reported in GenBank together with bootstrap analysis using 5000 replicates.