Plant Biotechnology 26, 121–124 (2009)

Gene Note

Isolation and expression analysis of *AGAMOUS*-like genes from *Eucalyptus grandis*

Tomohiko Kato, Takashi Hibino

Forestry Research Institute, Oji Paper Co. Ltd., Kameyama 519-0212, Japan
* E-mail: takashi-hibino@ojipaper.co.jp Tel: +81-5958-5-0122 Fax: +81-5958-5-1528

Received October 28, 2008; accepted December 19, 2008 (Edited by R. Sano)

Abstract In this study, as a first step to achieve a goal to produce genetically-engineered sterile *Eucalyptus*, we focused on the *AGAMOUS*-like genes which are especially involved in stamen formation. Three new distinct cDNA fragments (*EgAGL1*, *EgAGL2* and *EgAGL3*) and two distinct genomic DNA fragments (*EgAGL1* and *EgAGL2*) encoding MADS-box proteins were isolated from *Eucalyptus* grandis. In the present study, three genes have 62–74% homology with the *Arabidopsis AGAMOUS* (AG) gene in deduced amino acid level. *EgAGL1*, *EgAGL2* and *EgAGL3* were classified close to the *AGAMOUS* gene by phylogenic analysis. *EgAGL1* and *EgAGL2* were strongly expressed in flower buds, and consequently may regulate stamen formation in *Eucalyptus*.

Key words: *AGAMOUS*-like, *Eucalyptus*, MADS-box.

Recently, with the increasing demand for renewable energy, the importance of fiber and chemical materials from woody plants is growing rapidly. In particular, improvement of fast-grow trees such as *Eucalyptus* species, which can grow in drought, cold, or nutrient-poor environments, will be anticipated in the near future world wide (FAO 2004). *Eucalyptus* wood is important as a raw material for industrial pulp and paper making (Higgins 1984; Hillis 1984). The increased bio-mass production from cultivated trees also reduces pressure to exploit native forests. One method to improve the characteristics of cultivated plants is to artificially introduce a certain gene into the target plant using recombinant DNA techniques. However, horizontal spread of the introduced gene derived from a genetically modified (GM) plant by the spread of pollen in the environment is a serious problem (FAO 2004).

Eucalyptus has a distinct floral shape from that of many plants, although the basic floral structures are equivalent to those of most plants (Southerton et al. 1998). In the *Eucalyptus* flower, sepal primordia fuse to each other during the early stage of flower development. Petal primordia are also fused. The fused sepals and petals form a structure called operculum which covers developing reproductive structures before anthesis.

MADS-box genes encode transcription factors that control growth and development, including floral meristem and floral organ formation in plants (Rounsley et al. 1995; Alvarez-Buylla et al. 2000). These proteins are characterized by the presence of MADS-box domain which is a highly conserved DNA binding domain located in the N-terminal region of the protein (Purugganan et al. 1995). These transcription factors also contain the K-box domain which is involved in protein dimerization. The plant MADS-box genes have been divided into several classes by phylogenic analysis (Theissen et al. 2000; Parenicova L et al. 2003). The functions of the genes belonging to the same class are thought to be similar across plant species.

In this study, as a first step to achieve a goal to produce genetically-engineered sterile *Eucalyptus*, we focused on the *AGAMOUS*-like genes (Yanofsky et al. 1990; Weigel and Meyerowitz 1994) which are especially involved in reproductive structure, namely stamen and carpel, formation. Although many MADS-box genes have been identified in a number of plant species, few MADS-box genes have been reported in *Eucalyptus* (Kyozuka et al. 1997; Southerton et al. 1998) and they do not have high sequence similarity to *AGAMOUS*. To dissect a molecular basis for floral organ formation in *Eucalyptus*, we isolated *AGAMOUS*-like genes possibly involved in floral organ formation.

Eucalyptus grandis trees were grown in the experimental field at the Forestry Research Institute (Oji Paper Co. Ltd.). RNA was extracted from about 1 g of flower buds from 7-year-old plants using the Plant RNA Isolation Reagent (Invitrogen, USA), according to the supplier's protocol. The SMART kit (Clonetech, USA)

Abbreviations: PCR, polymerase chain reaction
This article can be found at http://www.jspcmb.jp/

was then used to generate flower buds-specific cDNA library. *AGAMOUS*-like genes were screened from the cDNA library. A DNA fragment from the *Arabidopsis AGAMOUS* (*AG*) gene was amplified with PCR primers (AG-f: 5′-ATCAAGCGGATCGAGAACAC-3′ and AGm1-r: 5′-GCTTTCTTGAGCAAACCACT-3′) (Yanofsky et al. 1990) and used as a probe for plaque hybridization. Three cDNA sequences with high homology to the *Arabidopsis AG* cDNA fragment were isolated. These genes were designated *EgAGL1*, *EgAGL2* and *EgAGL3*. The gene accession numbers are AB465729, AB465730, and AB465731, respectively.

Figure 1 shows the alignment of the predicted protein sequences of the three *EgAGL* cDNAs with three eucalyptus MADS-boxes genes (egm) genes of *E. grandis* (Southernton et al. 1998) and with *Arabidopsis AG* using the CLUSTALW program. A predicted protein of 222 amino acids is encoded by *EgAGL3*, which shares the strongest homology (74% identity over 185 amino acids) with the AG protein. *EgAGL2* encodes a predicted protein of 231 amino acids and shares 69% homology with the AG protein over 232 amino acids. *EgAGL1* encodes a predicted protein of 251 amino acids, and shares 62% homology with the AG protein over 156 amino acids. In contrast, the predicted proteins encoded by the three egm genes share lower amino acid homology (*egm1*, 47% identity over 174 amino acids; *egm2*, 41% identity over 150 amino acids; *egm3*, 48% identity over 179 amino acids) with the AG protein than the EgAGL proteins.

Figure 2 shows a phylogenetic tree from amino acid sequence of the *Arabidopsis* and rice *AG* genes (*AG* and

```
AG              TAYQSELGGDSSPLRKSGRGKIEIKRIENTTNRQVTFCKRRNGLLKKAYELSVLCDAEVA  60
EgAGL1          MEFPSEFSEASSQKRIGGRGKIEIKRIENTTNRQVTFCKRRNGLLKKAYELSVLCDAEVA  60
EgAGL2          MVFPTQATPEESPQRKMGRGKIEIKRIENTTNRQVTFCKRRNGLLKKAYELSVLCEAEVA  60
EgAGL3          ----------------MGRGKIEIKRIENTTNRQVTFCKRRNGLLKKAYELSVLCDAEVA  44
egm1            ----------------MGRGRVELKRIENKINRQVTFAKRRNGLLKKAYELSVLCDAEVG  44
egm2            ----------------MGRGKIEIKRIENSNNRQVTYSKRRNGLIKKAKEISVLCDAQVS  44
egm3            ----------------MGRGKVELKRIENKINRQVTFAKRRNGLLKKAYELSVLCDAEVA  44
consensus                        ***::*:*****. *****:.******:*** *:****:*:*.

AG              LIVFSSRGRLYEYSNNS-VKGTIERYKKAISDNSNT-GSVAEINAQYYQQESAKLRQQII 118
EgAGL1          LIVFSSRGRLYEYANNS-VRGTIERYKKASSDSSHP-QSVSEVNTQFYQQEASKLRRQIR 118
EgAGL2          LIVFSSRGRLYEYANDS-VKATIERYKKACSDSSSS-GSVSEANVQFYQQESAKLQQQIN 118
EgAGL3          LIVFSSRGRLYEYSNNS-IRSTIERYKKANSDSSNT-STVTEINAQYYQQESAKLRQQIQ 102
egm1            LIIFSNRGKLYEFCSSSSMLKTLERYQKCNYGALEPNVSARESLELSCQQEYLRLKARYE 104
egm2            VIIFGSSGKMHEYCSSN--TSLVDILDQYHTQCGKR---LWDAKQENLSNELDRIKKEND  99
egm3            LIIFSNRGKLYEFCSSSSMMKTIEKYQKCSYGSLETNCSINEMQNS--YQDYLKLKARVE 102
consensus       :*:*.. *:::*:...     ::  .:            :       ::  ::: .

AG              SIQNSN-RQLMGETIGSMSPKELRNLEGRLERSITRIRSKKNELLFSEIDYMQKR--EVD 175
EgAGL1          EIQVSN-RHILGEGISDLSFKDLKNLESKLEKSISRVRSKKNEMLFAEIEYMQKR--EIE 175
EgAGL2          NMQNNN-RQLVGDSIAGMNMKDMKTTEQKLEKAIAKIRAKKNELLFAEIEYMQKRLEEID 177
EgAGL3          MLQNSN-RHLMGDSLSSLSVKELKQLENRLERGITRIRSKKHEMLLTEIEYLQKK--EIE 159
egm1            GLQRTQ-RNLLGEELGQLCSKELESLERQLDGSLKQIRSRRTQYMLDQVTDLQHR--EQV 161
egm2            NLQIQL-RHLKGEDITSLNHRELIILEDTLENGVGCVRDQKDEVLMTHRR-NQKQ----- 152
egm3            VLQRSQ-RNPPWEELGPLNSKELEQLEHQLENSLKQIRSAKTQFMFDQLXHLQHK--EQM 159
consensus        :*   *    : :  :  :::   *  *:.:  :*  : : :  .    *::

AG              LHNDNQILRAKIAENER-NNPSIS------LMPGGSNYEQLMPPPQ-TQSQPFDSRNYFQ 227
EgAGL1          LQNDNMYLRAKIAENEGAQQQQQQGSDHHFNMPGSSSVYEALP------SQPAYDRNFLQ 229
EgAGL2          LHNNNQVLRAKIAESERTQHADMN------LMPGGTNYDFMQP----SSSQPFDSRNYFQ 227
EgAGL3          LENESVFLRTKIAEVDRIQQGNMV---------AGPQVNVMEA---------LASRNFFP 201
egm1            VHEANRTLNQRLMEG----------------------YQVMRSS---------------- 183
egm2            LEEDAKELHFYAQQKDMMMAETGR--------AGNDGYHQRMK-------ADFGPSTYHV 197
egm3            LVEANRELWKKLEESNTRIPLRLGWEA---EDHNNISYSRLPTQSQGLIFQPLGGNPTLQ 216
consensus       : :    *     :

AG              VAALQ-PNNHHYSSAGRQDQTALQLV--- 252
EgAGL1          VNVLE-PNHQSYS---RFDHTALQLV--- 251
EgAGL2          VNVL--------------------------- 231
EgAGL3          SNMVEGGTAYSHS-----DKKVLHLG--- 222
egm1            ------------------------------
egm2            QPIQPNLQRRF---------------------- 208
egm3            IGYNPAGSNELNVSAADQHPNGFIPGWML 245
consensus
```

Figure 1. A comparison of the deduced amino acid sequences of six eucalyptus MADS-box genes and the *Arabidopsis* AG. The deduced amino acid sequences of *EgAGL1*, *EgAGL2*, *EgAGL3*, *egm1*, *egm2*, *egm3*, and *AG* were aligned using the CLUSTALW program. The MADS-box domain is underlined in the consensus line. Asterisk symbols indicate amino acids that are identical in all sequences, colons indicate conserved substitutions, and periods indicate semi-conserved substitutions.

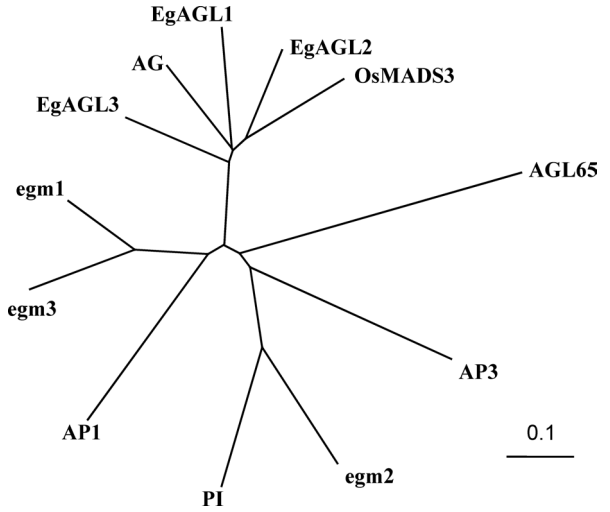

Figure 2. Phylogenic tree of the *Arabidopsis* and rice *AGAMOUS* proteins (AG and OsMADS3), *Arabidopsis* four MADS-box proteins (AP1, AP3, PI and AGL65) and six *Eucalyptus* MADS-box proteins Amino acid sequences were aligned using the CLUSTALW program and the tree was constructed with the TreeView software.

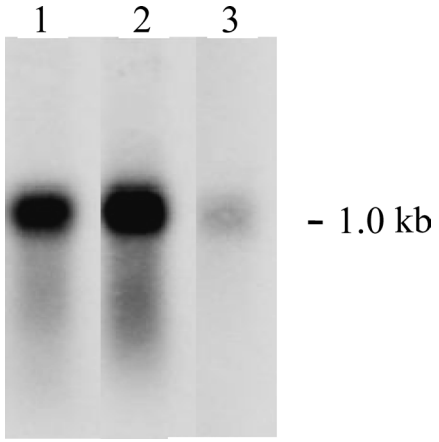

Figure 3. Expression analysis of *EgAGL* genes in *E. grandis*. Flower bud mRNA was purified using the PolyATtract mRNA purification system III kit (Promega, USA) according to the supplier's protocol. Purified flower bud mRNA (400 ng) was separated by agarose gel electrophoresis, and then was transferred to a nylon membrane. Blotting and hybridization were performed with Koyama's procedures (Koyama et al, 2006). Expression of *EgAGL1*, *EgAGL2*, and *EgAGL3* was detected using the 3′ untranslated region of each gene as a probe. Lane 1, *EgAGL1*; lane 2, *EgAGL2*; lane 3, *EgAGL3*.

OsMADS3), *Arabidopsis* four MADS-box genes (*AP1*, *AP3 PI* and *AGL65*) and six *Eucalyptus* MADS-box genes using CLUSTALW and the TreeView software. *EgAGL1*, *EgAGL2* and *EgAGL3* were clearly classified close to the *AGAMOUS* gene. Expression of these three genes in flower buds was analyzed by Northern blot analysis (Figure 3). *EgAGL1* and *EgAGL2* mRNAs were strongly expressed in flower buds. Expression of *EgAGL3* in flower buds was weaker than *EgAGL1* and *EgAGL2*. These results indicate that the three *EgAGL* genes are candidate genes for *Eucalyptus AGAMOUS* orthologs.

We examined *EgAGL1* and *EgAGL2* as potential targets for commercial use since the expression of these genes in the flower buds was extremely high. Full-length *EgAGL1* and *EgAGL2* cDNAs were used to screen an *E. camaldulensis* genomic library (Koyama et al. 2006). Two partial genomic DNA sequences with high homology to the *EgAGL1* cDNA and *EgAGL2* cDNA were isolated.

Comparison of the cDNA and genomic sequences of both the *EgAGL1* and *EgAGL2* genes have five introns within the coding region. We isolated upstream sequences (around 1 kbp) from the *EgAGL1* and *EgAGL2* genes, both of which included a TATA-box in a possible promoter region, respectively. Gene accession numbers for the *EgAGL1* and *EgAGL2* promoters are AB465732 and AB465733. These regions will be important components of a transgene to prevent flowering in transgenic *Eucalyptus* trees grown for commercial use.

In the present study, three MADS-box genes were newly identified from flower buds of *Eucalyptus*. These genes have 62–74% homology with the Arabidopsis AG gene in deduced amino acid level. *EgAGL1* and *EgAGL2* were strongly expressed in flower buds, and consequently may regulate stamen and carpel formation in *Eucalyptus*.

Acknowledgements

We are grateful to Ms. Tomoe Shimamura of the Forestry Research Institute of the Oji Paper Co. Ltd for technical assistance with plaque hybridization and sequencing. This work was supported by the New Energy and Industrial Technology Development (NEDO) program, which is part of the "Development of Fundamental Technologies for Controlling the Material Production Process of Plants" project.

References

Alvarez-Buylla ER, Liljegren SJ, Pelaz S, Gold SE, Burgeff C, Ditta GS, Vergara-Silva F, Yanofsky MF (2000) MADS-box gene evolution beyond flowers: expression in pollen, endosperm, guard cells, roots and trichomes. *Plant J* 24: 457–466

FAO (2004) *Preliminary review of biotechnology in forestry, including genetic modification. Forest Genetic Resources Working Paper FGR/59E*. Forest Resources Development Service, Forest Resources Division, Rome, pp 1–118

Higgins HG (1984) Pulp and Paper. In: Hillis WE and Brown AG (eds) *Eucalyptus for wood production*. Academic press Inc, London, pp 290–316

Hillins WE (1984) Wood quality and utilization. In: Hillis WE and Brown AG (eds) *Eucalyptus for wood production*. Academic press inc, London, pp 259–289

Koyama T, Kato N, Hibino T, Kawazu T, Kimura T, Sakka K (2006) Isolation and expression analysis of phosphate transporter genes from *Eucalyptus camaldulensis*. *Plant Biotechnol* 23: 215–218

Kyozuka J, Harcourt R, Peacock WJ, Dennis ES (1997) *Eucalyptus* has functional equivalents of the Arabidopsis AP1 gene. *Plant Mol Biol* 35: 573–584

Parenicova L, de Folter S, Kieffer M, Horner DS, Favalli C, Busscher J, Cook HE, Ingram RM, Kater MM, Davies B, Angenent GC, Colombo L (2003) Molecular and Phylogenetic Analyses of the Complete MADS-Box Transcription Factor Family in Arabidopsis: New Openings to the MADS World. *Plant Cell* 15: 1538–1551

Purugganan MD, Rounsley SD, Schmidt RJ, Yanofsky MF (1995) Molecular evolution of flower development: diversification of the plant MADS-box regulatory gene family. *Genetics* 140: 345–356

Rounsley SD, Ditta GS, Yanofsky MF (1995) Diverse Roles for MADS Box Genes in Arabidopsis Development. *Plant Cell* 7: 1259–1269

Southerton SG, Marshall H, Mouradov A, Teasdale RD (1998) Eucalypt MADS-box genes expressed in developing flowers. *Plant Physiol* 118: 365–372

Theissen G, Becker A, Di Rosa A, Kanno A, Kim JT, Munster T, Winter K-U, Saedler H (2000) A short history of MADS-box genes in plants. *Plant Mol Biol* 42: 115–149

Weigel D, Meyerowitz EM (1994) The ABCs of floral homeotic genes. *Cell* 78: 203–209

Yanofsky MF, Ma H, Bowman JL, Drews GN, Feldmann KA, Meyerowitz EM (1990) The protein encoded by the Arabidopsis homeotic gene *AGAMOUS* resembles transcription factors. *Nature* 346: 35–39