Plant Biotechnology 25, 391–395 (2008)

Transgenic Note

Gene trap strategy, an effective tool for identification of novel genes expressed in anther tissues in *Arabidopsis thaliana*

Tohru Ariizumi[1], Takahiro Kawanabe[2], Shusei Sato[3], Tomohiko Kato[3,4], Satoshi Tabata[3], Kinya Toriyama[2,*]

[1] Department of Crop and Soil Science, Washington State University, Pullman, WA 99164-6420, USA; [2] Laboratory of Environmental Biotechnology, Graduate School of Agricultural Science, Tohoku University, Sendai, Miyagi 981-8555 Japan; [3] Department of Plant Genome Research, Kazusa DNA Research Institute, Kisarazu, Chiba 292-0818 Japan; [4] Forestry Research Institute, Oji Paper Company Co. Ltd., Kameyama, Mie 519-0212 Japan
* E-mail: torikin@bios.tohoku.ac.jp Tel: +81-22-717-8830 Fax: +81-22-717-8834

Received April 21, 2008; accepted May 8, 2008 (Edited by Y. Watanabe)

Abstract Analysis of GUS expression patterns was previously carried out for 20,000 gene trap lines of *Arabidopsis thaliana* generated by the Kazusa DNA Research Institute. This analysis indicated that 51 lines showed GUS expression within stamens. In this paper, we further focused on 42 of these lines in order to examine the detailed expression patterns at different anther developmental stages. Seven lines showed GUS expression exclusively in the tapetum during the early stage of anther development, thirty lines in pollen grains at the tricellular pollen stage and five lines in filaments at the tricellular pollen stage. We identified four trapped-genes by sequencing the adjacent fragments to T-DNA insertions. RT-PCR analysis confirmed that the mRNA expression patterns in these four genes were consistent with the expression pattern observed by GUS activity. Our results demonstrated that the gene trap system is an efficient strategy for identifying stamen specific genes in *Arabidopsis*.

Key words: Anther, *Arabidopsis thaliana*, gene trap system.

A gene trapping system has been developed and used as one of the strategies to monitor gene expression profiles by random T-DNA insertion with reporter genes such as luciferase, GFP and β-glucuronidase (GUS) into the genome. The advantages of this strategy include classification of genes based on expression analysis, promoter hunting and isolation of a flanking sequence adjacent to the T-DNA insertion through PCR amplification such as tail-PCR and suppression PCR techniques (Liu and Whittier 1995; Miyao et al. 1998). We previously generated a population of gene trap lines, referred to as KG lines, of *Arabidopsis thaliana* consisting of 51,000 lines, and examined the patterns of GUS expression in 20,000 lines. The illustrations of the GUS expression patterns of these lines are available at http://www.kazusa.or.jp/ja/plant/GUS. In this paper, we focused on the lines showing GUS expression exclusively in stamen. Stamens consist of an anther and filament, and their development appears to be controlled by a complicated and sophisticated mechanism (Scott et al. 2004). Pollen grains, which include male sperms, are produced in anther, while filament elongation enables transmission of pollen grains from the anther to the pistils in *Arabidopsis*. Anther tissue consists of four distinct walls, the epidermis, endothecium, middle layer and the tapetum. In particular, it has been demonstrated that the tapetum plays important roles in pollen development and maturation. Since tapetum plays a vital role in pollen development, research for identification of tapetum-specific genes and promoters has been widely carried out. Examples of tapetum-specific promoters in *Arabidopsis* include *A9*, *LTP12* and *XTH3* promoters, which induce gene expression from the meiosis stage, the uninucleate microspores stage and the bicellular pollen stages until tapetum degradation (Paul et al. 1992; Ariizumi et al. 2002). Collection of such tapetum-specific promoters showing gene induction at the distinct developmental stages would be useful not only to explore the role of tapetum for microspore development but also to develop technology in producing male sterile plants for hybrid seed production (Perez-Prat and van Lookeren Campagne 2002; Kawanabe et al. 2006). Gene trap KG lines were generated by the Kazusa DNA Research Institute using the T-DNA vector pGTAC-LUS, which contains a splice donor and acceptor sequences between the right border and the reporter gene (Kato et al. 2007). This allows synthesis of fused transcripts not only when T-DNA is inserted within exons but also within introns.

This article can be found at http://www.jspcmb.jp/

It also allows production of heterologous transcript species, and some of these are expected to be fused in frame, ensuring that some portions of the various transcripts are always translated as fusion proteins. From about 20,000 gene trap lines, we first screened the 51 lines which showed GUS expression exclusively within stamens (Kato et al. 2007).

In the current study, we determined the detailed pattern of GUS expression during anther development from the tetrad stage to anthesis in these KG lines. We found that the pattern of GUS expression could be classified into three groups based on the timing and localization of GUS activity. These groups consisted of lines showing the highest GUS expression level at the early anther developmental stages (Group I, Figure 1), lines showing the highest GUS expression level at the late anther developmental stages (Group II, Figure 2) and lines showing GUS expression in filament tissue after the tricellular pollen stage (Group III, Figure 3). The anther

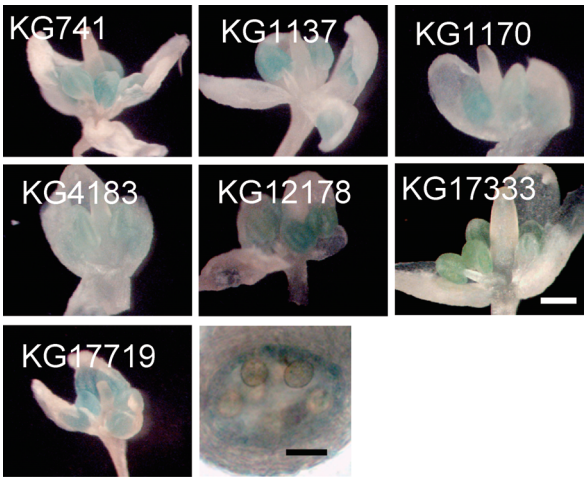

Figure 1. Flowers of the gene trap lines (Group I) showing GUS expression at the uninucleate microspore stage. GUS assay was performed according to Ariizumi et al., (2002). After the flower buds samples were washed with ethanol, they were mounted on glass slides and photographed using a light microscope. A photograph of a cross section from KG17719 at the uninucleate microspore stage is also shown in the last panel. White bar=600 μm, black bar=20 μm.

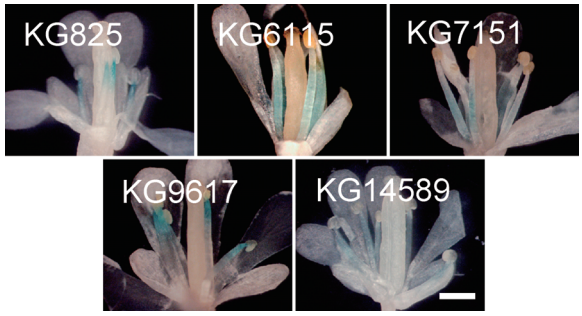

Figure 3. Flowers of the gene trap lines (Group III) showing GUS expression in the filament at the open flower stage. GUS assay was performed as described in Figure 1. White bar=600 μm.

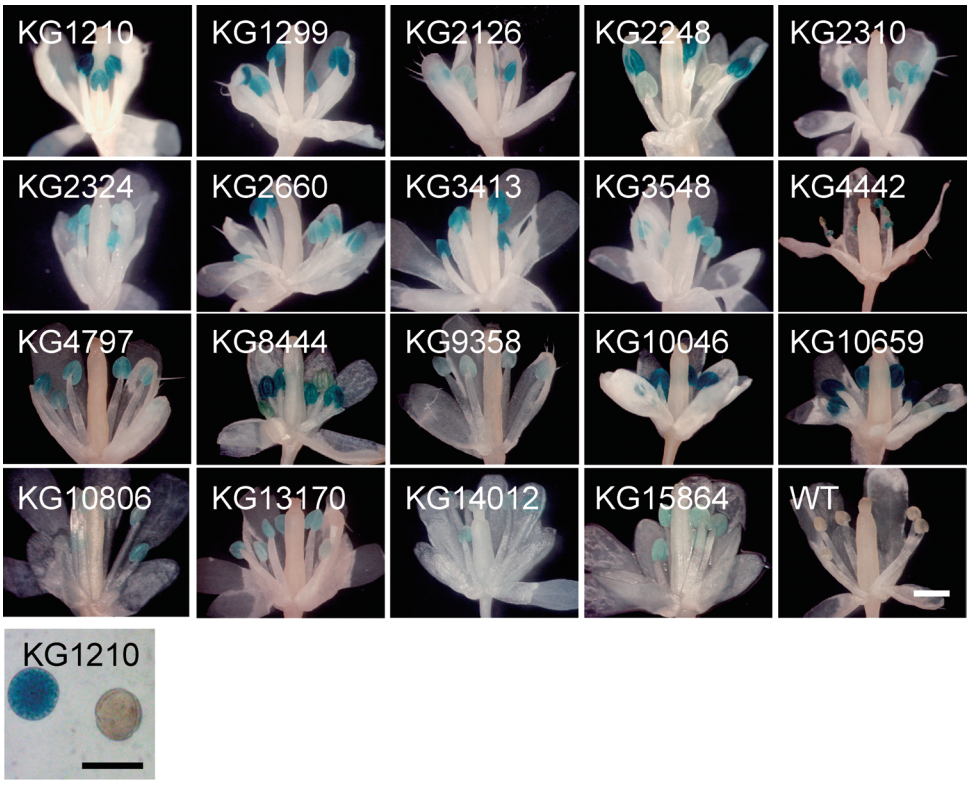

Figure 2. Flowers of the gene trap lines (Group II) showing GUS expression in anthers at the tricellular pollen stage. GUS assay was performed as described in Figure 1. A photograph of pollen grains from KG1210 separated from anthers is also shown in the last panel. White bar= 600 μm, black bar=20 μm.

developmental stages were previously defined based on the number of nuclei in the microspores using DAPI staining (Okada et al. 2000; Ariizumi et al. 2002).

Group I includes 7 gene trap lines (KG741, KG1137, KG1170, KG4183, KG12178, KG17333, KG17719), and their GUS expression patterns are shown in Figure 1 and Supplementary Table S1. Although these lines did not show GUS expression in flower buds whose stages were before the meiotic stage, KG741 and KG17719 lines showed GUS expression from the tetrad stages, while KG1137, KG1170, KG4183, KG12178 and KG17333 lines showed GUS expression from the uninucleate microspore stage (Supplementary Table S1). However, the GUS expression in these lines disappeared after the bicellular pollen stage. Since degradation of tapetum takes place after the bicellular pollen stage, we expected that the GUS expression might be observed exclusively in the tapetum. To examine this possibility, localization of the GUS activities in these lines was observed using a cross section of anther at the bicellular pollen stages (a typical example is shown in Figure 1, data not shown for others). As expected, the GUS activities were exclusively observed in the tapetum, and not in other anther walls (etc., middle layer, endothecium, epidermis). These results indicated that these seven gene trap lines showed GUS expression exclusively in the tapetum in anther tissue. Thirty lines were classified into Group II, and these gene trap lines showed the maximum GUS activity at the tricellular pollen stage (Figure 2; Supplementary Table S2). In this group, nine lines (KG1299, KG4442, KG4797, KG9358, KG10805, KG13170, KG13624, KG14012, KG14382) exhibited GUS expression solely at the tricellular pollen stage, while two lines (KG1210 and KG4702) began to show GUS activities from the tetrad or the uninucleate microspore stages. On the other hand, thirteen lines (KG801, KG2126, KG2248, KG2324, KG8444, KG10046, KG10659, KG14420, KG18257, KG18549, KG18566, KG18633, KG19729) began to show GUS activities from the bicellular pollen stage. At the tricellular pollen stage, the tapetum no longer exists due to programmed cell death, and pollen grains maturation proceeds (Scott et al. 2004). To examine the notion of whether GUS activity was found in the anther walls and/or pollen grains, pollen grains at the tricellular pollen stage were separated from anther tissues by smashing in the centrifuged tubes using pestles, and the grains were stained by GUS solution. This experiment demonstrated that GUS expression was exclusively restricted within pollen grains, and no GUS expression was observed in anther wall tissues (a typical example is shown in Figure 2, data not shown for others), indicating that GUS expression at the tricellular pollen stage was exclusively within pollen grains. We examined the number of pollen grains showing GUS positive and negative in the lines of Group II. The segregation ratio and expected numbers of the integrated T-DNA locus are presented in Supplementary Table S4. It appeared that the proportion of segregation ratio in GUS positive to GUS negative were 1 : 1 (significant at the 5% level) in fourteen lines (Supplementary Table S4). As a result, it was expected that the transgene would be integrated in one locus in these lines. The segregation patterns were 3:1 (significant at the 5% level) and integration of the transgene was expected to be at two loci in eight lines. Insertion of transgene was expected to be integrated in more than three loci in the other six lines. Taken together, these segregations of GUS-positive and GUS-negative pollen grains indicated that the GUS activity was gametophytically found in pollen. Group III includes five lines that showed GUS expression in the filament after the tricellular pollen stage, and these lines did not show GUS expression in other anther tissues (Figure 3, Supplementary Table S3). Interestingly, line KG825, KG9617 and KG14589 showed GUS expression at the upper region of the filament, while KG6116 and KG7151 lines showed GUS expression at the lower region of the filament.

The suppression PCR experiments were performed in order to isolate adjacent regions of T-DNA insertions. Among the gene trap lines, we succeeded in identifying the adjacent regions from 4 lines (KG1299, KG7151, KG10659 and KG17719). The T-DNA insertions in these lines were found in open reading frames (Figure 4). The AGI codes of these genes were At5g23160, At4g35640, At2g31985 and At5g17110, which were identified from KG1299, KG7151, KG10659 and KG17719 lines, respectively. It appeared that the At4g35640 encodes a protein annotated to the serine acetyltransferase. The serine acetyltransferase gene family encodes enzyme protein which produces *O*-acetylserine known to play an important role in the control of sulphate reduction and cysteine biosynthesis in plants (Howarth et al. 2003). However, the proteins encoded by At5g23160, At2g31985 and At5g17110 were annotated to unknown proteins, and no sequence similarities were found by BLAST search.

To confirm whether the pattern of GUS expression observed in gene trap lines was consistent with the mRNA transcription pattern, RT-PCR experiments were performed (Figure 5). In this experiment, we extracted mRNA from flower buds at the early stage, middle stage and late stage. The early stage corresponded to anther developmental stages before the tetrad stage, and the middle stage corresponded to the developmental stages between the uninucleate microspore stage to bicellular pollen stage, and the late stage and open flower stage corresponded to the tricellular pollen stage as described previously (Ariizumi et al. 2004). The mRNA accumulation of the gene encoded by the At5g23160 was

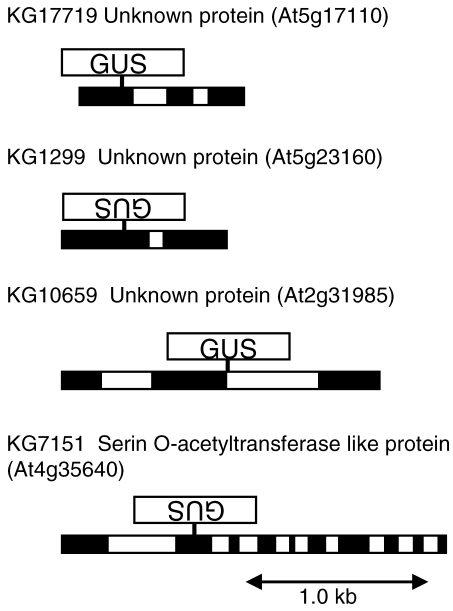

Figure 4. Flanking gene structure of T-DNA insertion identified by suppression-PCR. The locations of T-DNA insertions are shown. The directions of the GUS gene are shown. Exon is indicated as filled box and intron is indicated as opened box. Suppression PCR was performed according to Miyao et al. (1998). The nested primers adjacent to RB and LB of the T-DNA were as follows: AD1 (5′-NGTCGA-(G/C)(A/T)GANA(A/T)GAA-3′), AD2 (5′-GTNCGA(G/C)(A/T)CA-NA(A/T)GTT-3′), AD3 (5′-(A/T)GTGNAG(A/T)ANCANAGA-3′).

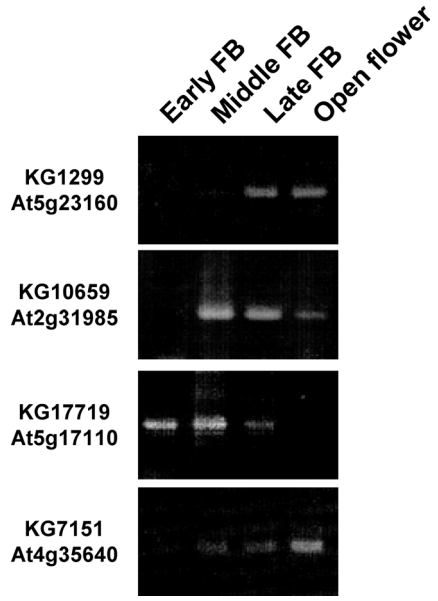

Figure 5. mRNA expression pattern of genes flanked by T-DNA insertion. mRNA accumulation of genes encoded by At5g23160, At2g31985, At5g17110 and At4g35640 was determined by RT-PCR. Early FB, early buds that corresponded to bud 3 of the inflorescence (anthers prior to the tetrad stage); Middle FB, the middle stage of flower buds corresponded to buds 4–5 of the inflorescence (uninucleate microspore stage and bicellular pollen stage); Late FB, the late stage of flower buds corresponded to buds 1 and 2 (tricellular pollen stage); Open flowers (tricellular pollen stage). mRNA was isolated and then reverse-transcribed to synthesize first strand cDNA using a First-Strand cDNA Synthesis Kit (GE healthcare, Sweden). cDNA was diluted and then used as a template for PCR amplification with primer KG1299-F (CGCGGGTCTCTAAACTTGAC), KG1299-R (CGAAACTCT-GTGTTGCCGAA), KG4183-F (CCACGATCATCATCATTCCG), KG4183-R (ACCACGAAGATGAAGTAGCG), KG7151-F (TGAC-ATTAGGAGGAACTGGG), KG7151-R (AGACTCACCATGCTC-CATTG), KG10659-F (ATGATGGCTGTTACACGCGA), KG10659-R (AACTCAGACAAAGACGCTGC), KG17719-F (CATGGCAGAG CTTGCTAATG) and KG17719-R (CGAGCCATAAAAGTG) primers. *Actin2* was amplified as a loading control using actin2-F (CACCTAGCAGCATGAAGATC) and actin2-R (GTGAACAAT-CGATGGACCTG) primers. In our system the control *Actin2* gene showed very reproducible similar expression levels among flower buds from the early, middle and late developmental stages. PCR was carried out with Takara Taq polymerase for 30 cycles of denaturation for 1 min at 94°C, annealing for 1 min at 58°C, and extension for 1 min at 72°C, followed by final extension for 5 min at 72°C.

observed at the tricellular pollen stage. The mRNA accumulation of gene encoded by the At2g31985 was detectable from the middle stage to the open flower stage. The mRNA accumulation of the gene encoded by the At5g17110 was observed from the early stage of flower buds, and significantly decreased at the tricellular pollen stage. The mRNA accumulation of the gene encoded by the At4g35640 was observed from the middle stage of flower buds, and the highest expression was observed at the open flower stage. These results indicated that the expression patterns observed by the RT-PCR were consistent with the GUS expression pattern observed in the gene trap lines (Figures 1 to 3; Supplementary Tables S1 to S3).

Since tapetum is a tissue well known to act as a supplier of nutrients and lipid components required for microspore development and pollen wall formation, the genes responsible for the GUS activity in Group I might be involved in these processes. It has been demonstrated that many of the genes expressed in pollen grains induce gene expression under the gametophytic control (Ariizumi et al. 2002). Consistent with this, the GUS activity in the examined lines of Group II all showed their expression under the gametophytic control (Supplementary Table S2). This suggested that the process which followed pollen maturation (i.e., pollen tube elongation) might be gametophytically controlled. Anther filament elongation is a critical process in the transmission of male sperm to female stigma cells in angiosperm. It is suggested that successful filament elongation requires regulation of water transport from the upper portion to the lower portion of filament (Ishiguro et al. 2001). Therefore it is possible that the genes regulated by Group III may be associated with a filament elongation process.

Generally, it is suggested that the *cis*-acting regulatory elements within promoters are responsible for spatial and temporal regulation of gene expression. To confirm this notion, we compared the promoter sequences of At2g31985 and At5g23160 with those of other pollen-specific genes, since the sequences of two pollen-specific

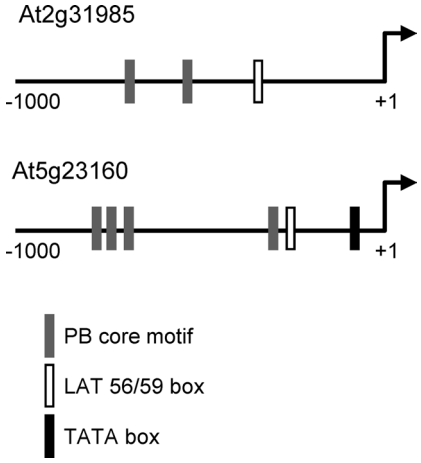

Figure 6. Localization of the pollen-specific *cis*-regulatory elements in the promoters of At2g31985 and At5g23160. Nucleotides are numbered with the first nucleotide of the translation initiation codon marked as +1. The box indicates sequences significantly similar to that of the PB core motif, LAT56/59 box and the TATA box.

cis elements (PB core motif and LAT56/59 box) are available. These two elements have been reported to be enhancer elements identified in tomato pollen-specific genes (Twell et al. 1991). The sequence motif, in which six out of seven nucleotides matched those of the PB core motif (TGTGGTT), was found at −690 and −544 for the promoter of At2g31985, while it was found at −783, −757, −713 and −297 for the promoter of At5g23160 when translational start was indicated as +1 (Figure 6). The sequence motif, in which eight out of ten nucleotides matched those of the LAT56/59 (GAAATTGTGA), was found at −325 for the promoter of At2g31985, and at −297 for the promoter of At5g23160. Also, the putative TATA box sequence, which is the core promoter sequence involved in the initiation of transcription by RNA polymerases, was found at −55 for the promoter of At5g23160. These results suggested that GUS activity of these promoters are partly controlled by the pollen-specific *cis*-regulatory elements.

We also identified T-DNA homozygous lines in the gene trap lines KG1299, KG10659, KG7151 and KG17719. However, anther development and fertility were equivalent to that of wild type plants (T. Ariizumi and K. Toriyama, Unpublished data). This could be partly due to the gene redundancy. Or it is also possible that they did not play important roles in anther and pollen development, although we could not exclude a possibility that they might have a slight effect on the development. Further analysis is still needed to determine the biological role of these genes during the regulation of anther development.

Acknowledgements

This work was supported by a grant-in-aid (No. 18075002 and No. 16380002) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

Ariizumi T, Amagai M, Shibata D, Hatakeyama K, Watanabe M, Toriyama K (2002) Comparative study of promoter activity of three anther-specific genes encoding lipid transfer protein, xyloglucan endotransglucosylase/hydrolase and polygalacturonase in transgenic *Arabidopsis thaliana*. *Plant Cell Rep* 21: 90–96

Ariizumi T, Hatakeyama K, Hinata K, Inatsugi R, Nishida I, Sato S, Kato T, Tabata S, Toriyama K (2004) Disruption of the novel plant protein NEF1 affects lipid accumulation in the plastids of the tapetum and exine formation of pollen, resulting in male sterility in *Arabidopsis thaliana*. *Plant J* 39: 170–181

Howarth JR, Dominguez-Solis JR, Gutierrez-Alcala G, Wray JL, Romero LC, Gotor C (2003) The serine acetyltransferase gene family in Arabidopsis thaliana and the regulation of its expression by cadmium. *Plant Mol Biol* 51: 589–598

Ishiguro S, Kawai-Oda A, Ueda J, Nishida I, Okada K (2001) The DEFECTIVE IN ANTHER DEHISCIENCE gene encodes a novel phospholipase A1 catalyzing the initial step of jasmonic acid biosynthesis, which synchronizes pollen maturation, anther dehiscence, and flower opening in *Arabidopsis*. *Plant Cell* 13: 2191–2209

Kato T, Tabata S, Sato S (2007) Expression analysis of gene trap lines and mapping of donor loci for *Dissociation* transposition in *Arabidopsis*. *Plant Biotechnol* 24: 467–479

Kawanabe T, Ariizumi T, Kawai-Yamada M, Uchimiya H, Toriyama K (2006) Abolition of the tapetum suicide program ruins microsporogenesis. *Plant Cell Physiol* 47: 784–787

Liu YG, Whittier RF (1995) Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics* 25: 674–681

Miyao A, Yamazaki M, Hirochika H (1998) Systematic screeining of mutants of rice by sequencing retrotransposon-insertion sites. *Plant Biotechnol* 15: 253–256

Okada T, Sasaki Y, Ohta R, Onozuka N, Toriyama K (2000) Expression of Bra r 1 gene in transgenic tobacco and Bra r 1 promoter activity in pollen of various plant species. *Plant Cell Physiol* 41: 757–766

Paul W, Hodge R, Smartt S, Draper J, Scott R (1992) The isolation and characterisation of the tapetum-specific *Arabidopsis thaliana* A9 gene. *Plant Mol Biol* 19: 611–622

Perez-Prat E, van Lookeren Campagne MM (2002) Hybrid seed production and the challenge of propagating male-sterile plants. *Trends Plant Sci* 7: 199–203

Scott RJ, Spielman M, Dickinson HG (2004) Stamen structure and function. *Plant Cell* 16: S46–S60

Twell D, Yamaguchi J, Wing RA, Ushiba J, McCormick S (1991) Promoter analysis of genes that are coordinately expressed during pollen development reveals pollen-specific enhancer sequences and shared regulatory elements. *Genes Dev* 5: 496–507