ORIGINAL RESEARCH ARTICLE OPEN ACCESS

Development of Activation Tagged Mutants in Rice CV BPT 5204 and Identification of the SUMO Protease Gene Associated with Early Flowering

Reddy MR , Reddy SV, Manimaran P, Quadriya H, Mahendranath G, Barbadikar KM, Mangrauthia SK, Sundaram RM and Balachandran SM*

Biotechnology Laboratory, Crop Improvement Section, ICAR-Indian Institute of Rice Research, Rajendranagar, Hyderabad – 500030, India. *Corresponding author (email : balasena@yahoo.com)

Received: 5th March 2018, Accepted: 11th April 2018

Abstract

In present study, we developed activation tagged mutants lines in rice *cv* BPT 5204 with the help of *Cauliflower Mosaic Virus* (CaMV) 4X Enhancer which can be used as source material for functional genomics. The 4X Enhancer was flanked by *Dissociation (Ds)* transposable element along with part of *Activator* (*Ac)* element in the T-DNA of binary vector pSQ5. *Ac-Ds* tagged population was developed in rice cultivar BPT 5204 through *Agrobacterium* mediated transformation. Twelve primary transformants were confirmed by PCR. Within the plant cell, it is expected that the *Ac* element may transpose the *Ds* element along with enhancer into different loci of chromosomes from T_2 generation onwards and transposition of *Ds* element is responsible for different phenotypic variants. Among the several mutants, the plant # AD-26 showed 12 days early flowering when compared with other mutants as well as with wild type (BPT 5204). By flanking sequence analysis (TAIL PCR), we identified that the transposed *Ds* element was reintegrated at intergenic region on chromosome 3 and located 1kbp distant from the SUMO protease protein gene (LOC_Os03g29630). Relative expression studies revealed 1.8-fold up regulation of LOC_Os03g29630 gene which might be responsible for early flowering in mutant. Further characterization of the gene is required to completely unravel the function and effect of this gene on early flowering and it may be valuable in crop improvement.

Key words: Transgenic rice, Enhancer, Activator, Dissociation, TAIL PCR

Introduction

Rice is one of the most important crops, which feeds more than half of the world population. Advantages like high quality genome sequence availability (Goff *et al*., 2002) and smaller genome size (~ 400 Mb) has made rice as an appropriate model plant among monocots. Many untapped genes are yet to be discovered in post rice genome sequencing era. Functional genomics studies are required to identify the novel genes and their functions by various approaches. In order to identify and study functional genes, an efficient strategy called 'activation tagging' through maize transposon (*Ac*/*Ds*) based mutagenesis method was employed. Activation-tagging is the integration of T-DNA randomly in rice genome harbouring tetrameric (4X) enhancer element of Cauliflower mosaic virus (*CaMV*) 35S promoter. This enhancer element can enhance the expression of endogenous genes non-specifically which are located near by the T-DNA integration loci and resulting in gain-of-function phenotype (Jeong *et al*., 2002; Jeong *et al*., 2006).

Rice crop productivity depends on physiological and agronomical traits. Flowering is an important agronomical trait controlled by florigen genes, photoperiodic regulation and circadian cycle systems (Tsuji *et al*., 2011; Jiao *et al*., 2005; Izawa 2002). As flowering is a complex process, many genes are involved either directly or indirectly (Seok *et al*., 2015). Both rice and *Arabidopsis* follow similar photoperiodic regulation during flowering stage (Izawa 2002; Izawa 2007). Genes such as early heading date (*Ehd1, Ehd2 and Ehd3*), grain number, plant height and heading date (*Ghd7*) play significant role in regulation of rice flowering. Regulation of flowering genes depends upon the photoperiodic signals. Other than circadian and photoperiodic regulation, flowering is also affected by environmental factors like abiotic stresses. Plants follow adaptive mechanisms to complete their life cycle in timely manner under stress conditions. Early flowering, one of the adaptive mechanisms, is observed in plants under stress condition but the exact molecular mechanism is poorly understood. Small ubiquitin-like modifier (SUMO) proteases are reported to be involved in de-sumoylation and these proteases convert the precursor of SUMO proteins into mature protein (Hermkes *et al*., 2011). In particular, SUMO proteases affect many developmental processes in response to both biotic and abiotic stresses. Various

studies suggested that SUMO proteases are involved in the regulation of plant development, cell growth and flowering time. In *Arabidopsis*, SUMO proteases located at cytosol (ULP1a) and also in nucleus (ULP1b, ULP 1c, ULP 1d, ESD4, ULP2a and ULP 2b) are involved in dwarfism, early flowering and salt stress induced early flowering (Park *et al*., 2011). Other SUMO family proteases like ESD4 (Murtas *et al*., 2003; Reeves *et al*., 2002; Xu *et al*., 2007) and *AtUPL1* (Hermkes *et al*., 2011) gene repression were responsible for early flowering. AtSUM3 protein was positively regulated and thus early flowering in *Arabidopsis* (Van den Burg *et al*., 2010).

In this study, we developed activation-tagged rice lines in the background of an elite indica *cv*. BPT5204 (Samba Mahsuri) with pSQ5 binary vector (Qu *et al*., 2008) through *Agrobacterium*-mediated transformation. From the confirmed transgenic plants, different phenotypic variants were observed among tagged lines and one such phenotype, early flowering was identified among the T_2 generation plant. The plant AD-26 is unstable because of presence of the both *Ac* and *Ds* elements. Further, flanking sequence analysis of AD-26 plant revealed that the *Ds* element with 4Xenhancer was re-inserted near LOC_Os03g29630 gene (ulp1 protease family protein) which belong to ulp1 (SUMO) protease gene family. Quantitative Real time PCR (qRT-PCR) analysis confirmed the elevated expression of LOC_Os03g29630 in AD-26 plant.

Materials and Methods

Gene construct and rice transformation

The gene construct pSQ5 used in this study was kindly provided by Prof. P.B. Kirti, University of Hyderabad, India. Preparation and details of pSQ5 construct was described earlier (Qu *et al*., 2008). For rice transformation, pSQ5 construct was mobilized into *Agrobacterium* strain EHA105 by tri-parental method. Twenty one days old embryogenic calli of indica rice *cv* BPT 5204 were transformed with *Agrobacterium* EHA105 which harbor pSQ5 vector, followed by co-cultivation and washing of transformed calli as described previously (Manimaran *et al*., 2013). The washed calli were maintained in selection medium containing MS basal salts (Murashige and Skoog 1962), 2 mg/L 2,4-D, 0.5 mg/L kinetin, 500 mg/L L-proline, 500 mg/L casein hydrolysate, 30 g/L maltose, solidified with 0.3% phytagel and supplemented with 50 mg/L hygromycin (Hi-Media, India) for 15 days in dark. After three cycles of selection in hygromycin containing medium, resistant calli were transferred to regeneration medium supplemented with MS basal salt, 2 mg/L kinetin, 0.3 mg/L NAA, 30 g/L sucrose, 30 g/L D-sorbitol and 0.4% phytagel. The regenerated plantlets were transferred to rooting medium (1/2 MS basal salt + 15 g/L sucrose + 0.4% phytagel) and then complete plants were transferred to hardening medium (Yoshida *et al*., 1976). The hardened plants were transferred to earthen pots and maintained under controlled conditions in a biosafety glass house.

Molecular confirmation and inheritance of transgene

Total genomic DNA extraction, PCR reactions of transgenic plants and WT (non-transformed control BPT 5204) were followed as described previously (Manimaran *et al*., 2013). Multiplex PCR was performed to confirm the presence of *Ac-Ds* elements in transgenic plants, primer pairs were designed from *hpt* and *rfp* genes to confirm the presence of *Ac* and *Ds* elements respectively (HPT-FWD: 5'-TAT TTC TTT GCC CTC GGA CGA G -3', HPT-REV: 5'-ATG AAA AAG CCT GAA CTC ACC G-3' and RFP-FWD: 5'-GAA GCT GAA GGT GAC CAA GG -3', NOS-REV: 5'-CGC TAT ATT TTG TTT TCT ATC GCG T -3'). For inheritance study of the transgene, T_1 seeds were collected and germinated on ½MS medium supplemented with 50 mg/L hygromycin and without hygromycin medium.

TAIL-PCR and flanking sequence analysis

Thermal Asymmetric Interlaced Polymerase Chain Reaction (TAIL-PCR, Liu *et al*., 1995) was performed using Ex-Taq DNA polymerase (Takara, Japan) in a Thermal Cycler (Eppendorf, Germany). Genomic DNA isolated from transgenic rice plants was used as a template for TAIL-PCR reaction. Three nested primers specific to 5' *Ds* element and one non-specific arbitrary degenerate primer (AD: 5' NTC GAS TWT SGW GTT 3') were used in primary, secondary and tertiary PCR reactions for tracking of *Ds* element. The tertiary PCR product was run on 1.2% agarose gel and purified through wizard ® SV Gel and PCR Clean-Up system (Promega, USA) and sequenced through Sanger sequencing. The emanating sequences were checked using the BLAST tool against the rice annotation database sites at MSU (http://rice.plantbiology.msu.edu) and RAP-DB (http://rapdb.dna.affrc.go.jp/).

RNA isolation and Real Time PCR

Total RNA was isolated at flowering stage from 100 mg of rice flag leaf according to the manufacturer's instructions (Nucleospin RNA Plant kit, Macherey Nagel, Germany). RNA was quantified by Nanodrop® ND-1000 Spectrophotometer. Optical density (OD) of RNA samples with 260/280 ratio between 1.9 and 2.0 and 260/230 ratio on or above 2.0 was used for real-time PCR analysis. One

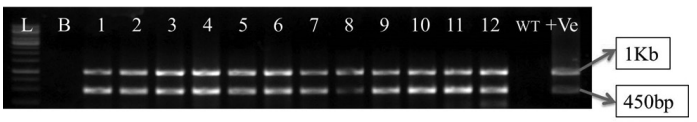
µg of total RNA was taken for first strand cDNA synthesis using oligo d (T) primers (PrimeScript First Strand cDNA Synthesis Kit, Takara, Japan) according to manufacturer's instructions. The cDNA quantity was normalized in all the samples. The cDNA was mixed with 10 µl of 2X SYBR premix (Takara, Japan), 2 µM each of LOC_Os03g29630 and LOC_Os03g29600 gene specific primers (Ulp F: 5' GGG TGG GAT TGT GGT ATG TT 3'; Ulp R: 5' CTC AGT ATC TCC TTC GCT GTT C 3') and (Tpn F: 5' ATT CCT CTA GCG GCA ACA C; 3' Tpn R: 5' CCC GCG TAC ATA GCC ATT 3') in a final volume of 20 µl. The *OsActin1* primers (*OsActin1* F: 5' CCG GTG GAT CTT CAT GCT TAC CTG G 3'; R: 5' CGA CGA GTC TTC TGG CGA AAC TGC 3') were used as internal control. The real-time PCR was performed in Roche Light Cycler. The PCR profile for qRT-PCR was as follows: 10 min at 95ºC, and 35 cycles of 15 seconds at 95ºC, 15seconds at 58ºC and 10 seconds at 72ºC in 96-well optical reaction plates (Roche Light cycler). Each sample was analyzed in two biological and three technical replicates. The quantification cycle (Cq) method was used to quantify the relative expression levels in real-time PCR. ΔCt was calculated by the difference between Ct target and Ct reference. Further, ΔΔCt values were calculated using the formula ΔΔCt= ΔCt of AD-26 – ΔCt of WT, and then fold difference was calculated from $2^{-\Delta\Delta Ct}$ (Schmittgen *et al*., 2008). *OsActin1* was used for normalizing gene expression level.

Results

Development of Ac-Ds transgenic plants and molecular confirmation

About 8000 embryogenic calli were infected with *Agrobacterium* strain EHA105 harbouring pSQ5 binary vector. After three cycles of selection in 50 mg/ml hygromycin medium, 1250 calli were selected from which 450 plants were regenerated. These plants were screened by multiplex PCR using primers targeting *hpt* and *rfp* genes to confirm presence of *Ac* and *Ds* elements respectively. Twelve plants were confirmed positive which showed 1kbp amplicon with *hpt* primers and 450 bp amplicon with *rfp* primers (Figure 1).

Figure 1: Multiplex PCR confirmation of T_0Trasgenic plants by using *hpt* and *rfp* genes specific primers.

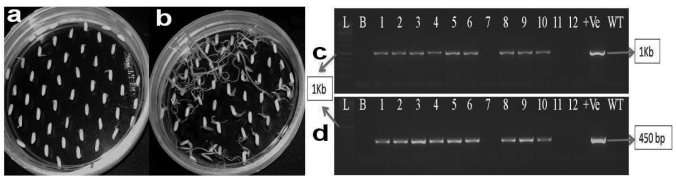


L: 1Kb ladder, B: Blank, 1- 12: T_0Transgenic plants, WT: Non transformed BPT 5204, +Ve: pSQ5 plasmid

Advancement of generation

From the twelve PCR confirmed plants, 50 seeds from each plant were forwarded to T_1 and T_2 generations for checking segregation pattern / inheritance (Figure 2a & 2b). Genomic DNA was isolated from fresh leaf samples of all the twelve lines (~800 plants) and analysed through PCR by using *hpt* and *rfp* primers. The results revealed that eight out of twelve lines followed the Mendelian inheritance (3:1 ratio) (Figure 2c&2d). These results indicated that these eight lines might harbour single copy T-DNA insertion whereas other lines may have two or more copies.

Figure 2: Segregation analysis of T_1 seeds through antibiotic selection and PCR.



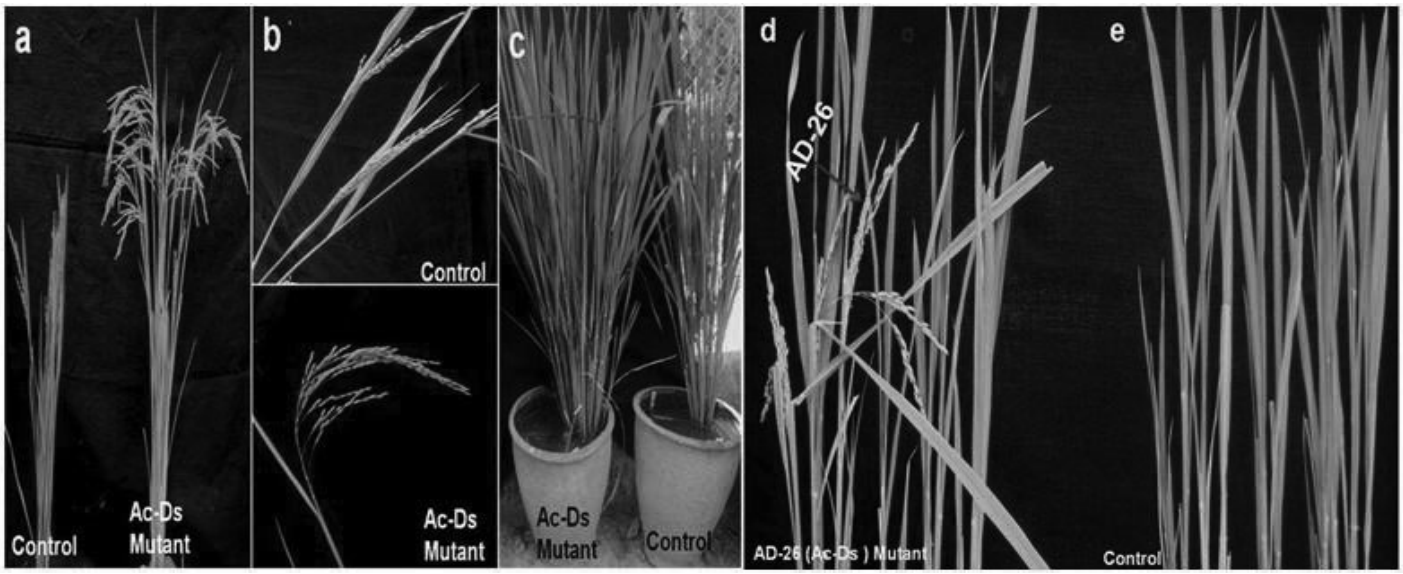
(a) & (b): WT and transgenic T_1 seeds were selected on hygromycin (50mg/ml) supplemented medium respectively. (c) & (d): PCR confirmation of T_1 plants by using *hpt* and *rfp* genes specific primers respectively. L: 1Kb ladder, B: Blank, 1- 12: T_0 Transgenic plants, WT: Non transformed BPT 5204, +Ve: pSQ5 plasmid

Early flowering phenotype in AD-26 and flanking sequence analysis

Among twelve transgenic lines, different economically important phenotypic variations were observed in T_2 generation like enhanced plant height (Figure 3a), complete panicle emergence (Figure 3b), increased number of tillers and yield (Figure 3c). Interestingly, one of the T_2 transgenic plants (AD-26) showed early flowering when compared to remaining mutants and wild type (WT). The early flowering T_2 plant was named as AD-26. The term 'flowering' was used even for inflorescence initiation, panicle extrusion (heading), development of the flower structure and flower opening (anthesis). Here panicle extrusion (Heading) was considered as flowering. In AD-26 mutant flowering (heading) attained after 91 days and WT plant attained after 103 days. Nearly 12 days early flowering was recorded in AD-26 mutant when compared with WT (Figure 3d & 3e) and the image was captured on 97th day. This early flowering in AD-26 mutant might be due to transposition effect of *Ds* element carrying CaMV35 4Xenhancer.

The genomic DNA of AD-26 mutant was isolated to perform the PCR with *hpt* and *rfp* gene specific primers. The plant showed amplifications of both 1 kb and 450 bp

Figure 3: Economically important phenotypic variations observed in activation tagged mutants at T_2 generation



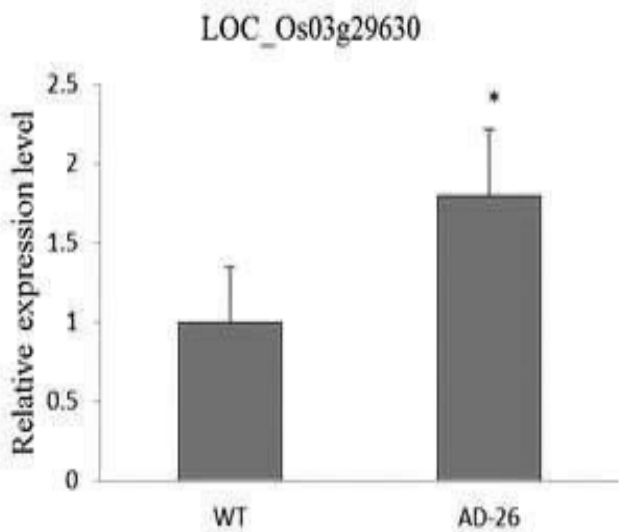
(a): Variations in plant height (b): Complete panicle emergence and grain yield (c): Variation in number of tillers (d): 12 days early flowering observed in AD-26 plant in T_2 generation (e): BPT 5204 control plant.

fragments respectively. This indicated that AD-26 plant contained both the *Ac* and *Ds* elements. TAIL-PCR was performed for the AD-26 targeting the 5' *Ds* element to identify the transposition. Through flanking sequence analysis, we identified that the transposed *Ds* element was reintegrated in non-coding region on chromosome number 3. Nearby genes from integrated locus were identified as LOC_Os03g29630 (ulp1 protease family protein) and LOC_Os03g29600 (transposon protein (CACTA, En/Spm sub-class). Expression analysis of these genes was performed to identify the enhanced tagged gene.

Enhanced expression of LOC_Os03g29630

Relative expression analysis of LOC_Os03g29630 and LOC_Os03g29600 was performed which reside within 20 kb region of *Ds* element integration. The results revealed that the gene LOC_Os03g29630 (~1 Kb downstream from enhancer) showed 1.8 fold up-regulation as compared to WT (Figure 4). However, there was no significant change in expression level of LOC_Os03g29600 gene which resides ~16.5 Kb distance from the enhancer.

Figure 4: Analysis of transcript levels of LOC_Os03g29630 gene in activation tagged AD-26 line in comparison with WT-BPT 5204



OsActin1 was used as a reference gene for quantitative RT-PCR. Means of three independent samples and standard errors are presented. * indicates significant difference at $P < 0.05$.

Discussion

In order to identify the genes function, an efficient strategy 'activation-tagging' through transposon-mediated mutagenesis approach was employed. In our attempt to identify the genes function, the Cauliflower Mosaic Virus (CaMV) 35S 4X enhancer was used to activate the nearby genes, located either at upstream and/or downstream regions (Jeong *et al*., 2002). For random distribution of enhancer in rice genome, maize *Ac/Ds* system was used. The enhancer was placed within *Ds* transposable element along with part of *Ac* element in T-DNA of binary vector pSQ5 (Qu *et al*., 2008). Upon transformation, 12 primary transgenic plants were confirmed through multiplex PCR by using gene specific primer pairs targeting *hpt* and *rfp* genes. Within the plant cell, it is expected that the *Ac* element may translocate the *Ds* element along with enhancer into different loci of chromosomes at T_2 or further generations. Thus, it is possible to create more number of transgenic plants with enhancer element at different positions on rice genome, in subsequent generations.

CaMV 35S 4X enhancer could enhance the gene expression non-specifically and caused different phenotypic variations (Lu *et al*., 2014). In present study, we observed various economically important phenotypic variations in plant height, panicle emergence, date of flowering, number of tillers and grain yield, in different lines from T_2 generation onwards. Transposition of *Ds* element from *Ac* element was responsible for trait variations and these variations were not stable until *Ds* element completely separated from *Ac* element. One of the T_2 plants, AD-26, showed early flowering (12 days) and also contained both *Ac* and *Ds* elements. Early flowering is an economically important agronomic trait which regulated by various molecular and environmental factors (Zhao *et al*., 2012). TAIL-PCR result of AD-26 plant (T_2 generation) showed that the T-DNA was integrated on chromosome no 3. Through BLAST analysis, LOC_Os03g29630 and LOC_Os03g29600 were identified as nearby genes at 1 Kb and ~16.5 Kb distance, respectively from *Ds* element. Quantitative RT-PCR data revealed that the gene LOC_Os03g29630 showed a 1.8 fold up-regulation as compared to WT whereas no enhancer effect on the expression level of LOC_Os03g29600 gene was detected. LOC_Os03g29630 gene encodes ulp1 (SUMO) protease family protein. These SUMO family genes are mainly involved in abiotic and biotic stress response. Their role in important agronomic traits like flowering time, cell growth and development has also been reported (Park *et al*., 2011).

This study suggests that the elevated level of LOC_Os03g29630 gene expression may be responsible for early flowering in AD-26 activation tagged mutant of BPT 5204. Further characterization of LOC_Os03g29630 by over-expression or silencing through RNA interference or genome editing will provide important information on role of this gene in controlling flowering time in rice. This can be an important candidate gene to be deployed in breeding or genetic engineering for development of early maturing rice cultivars.

Acknowledgements

The authors thank the Director, for providing facilities and support to carry the entire work at Indian Institute of Rice Research. The authors thankfully acknowledge the financial support from the Department of Biotechnology (DBT), GOI, through network project on *"Identification of candidate genes for enhanced water use efficiency in rice through activation tagging"* (Sanction File No. BT/PR-13105/AGR/02/684/2009).

Author contribution statement: Sub-cloning: MRR, PM, Transformation and molecular characterization: MRR, SVR and HQ, TAIL-PCR and gene identification: SKM, MRR and MG, Identification of mutants: SMB, SKM and RMS, manuscript writing and editing: MRR, KK and SMB, Conception, designing and overall guidance for the execution of the experiments drafting and editing of the manuscript: SMB.

References

Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M *et al*. 2002. A draft sequence of the rice genome. *Oryza sativa L*. ssp. japonica. *Science 296(5565)*: 92–100.

Hermkes R, Fu YF, Nürrenberg K, Budhiraja R, Schmelzer E, Elrouby N, Dohmen RJ, Bachmair A and Coupland G 2011. Distinct roles for *Arabidopsis* SUMO protease ESD4 and its closest homolog ELS1. *Planta 233(1)*: 63-73.

Izawa T, Oikawa T, Sugiyama N, Tanisaka T and Yano M 2002. Phytochrome mediates the external light signal to repress FT orthologs in photoperiodic flowering of rice. *Genes & Development 16(15):* 2006–2020.

Izawa T 2007. Adaptation of flowering-time by natural and artificial selection in *Arabidopsis* and rice. *Journal of Experimental Botany 58(12)*: 3091–3097.

Jeong DH, An S, Kang HG, Moon S, Han JJ, Park S, Lee HS, An K and An G 2002. T-DNA insertional mutagenesis for activation tagging in rice. *Plant Physiology 130(4):* 1636-1644.

Jeong DH, An S, Park S, Kang HG, Park GG, Kim SR, Sim J, Kim YO, Kim MK, Kim SR, Kim J, Shin M, Jung M and An G 2006. Generation of a flanking sequence-tag database for activation-tagging lines in japonica rice. *Plant Journal 45(1)*: 123-132.

Jiao Y, Ma L, Strickland E and Deng XW 2005. Conservation and divergence of light-regulated genome expression patterns during seedling development in rice and *Arabidopsis*. *Plant Cell 17*: 3239–3256.

Liu YG, Mitsukawa M, Oosumi T and Whittier RF 1995. Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. *The Plant Journal 8(3):* 457–463.

Lu G, Wang X, Liu J, Yu K, Gao Y, Liu H *et al*. 2014. Application of T-DNA activation tagging to identify glutamate receptor-like genes that enhance drought tolerance in plants. *Plant Cell Reports 33(4)*: 617-631.

Manimaran P, Ravikumar G, Raghurami Reddy M, Jain S, BhaskarRao T, Mangrauthia SK, Sundaram RM, Ravichandran S and Balachandran SM 2013. Infection of early and young callus tissues of indica rice BPT 5204 enhances regeneration and transformation efficiency. *Rice Science 20(5)*: 415–426.

Murashige Tand Skoog F 1962. A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiologia Plantarum 15*: 473–497.

Murtas G, Reeves PH, Fu Y, Bancroft I, Dean C and Coupland G 2003. A nuclear protease required for flowering-time regulation in *Arabidopsis* reduces the abundance of SMALL UBIQUITIN-RELATED MODIFIER conjugates. *The Plant Cell 15(10)*: 2308-2319.

Park HJ, Kim WY, Park HC, Lee SY, Bohnert H J and Yun DJ 2011. SUMO and SUMOylation in plants. *Molecules and Cells 32(4)*: 305-316.

Qu S, Desai A, Wing R and Sundaresan V 2008. A versatile transposon-based activation tag vector system for functional genomics in cereals and other monocot plants. *Plant Physiology 146(1)*: 189-199.

Reeves PH, Murtas G, Dash S and Coupland G 2002. Early in short days 4, a mutation in *Arabidopsis* that causes early flowering and reduces the mRNA abundance of the floral repressor FLC. *Development 129*: 5349-5361.

Schmittgen TD and Livak KJ 2008. Analyzing real-time PCR data by the comparative CT method. *Nature Protocol 3(6):* 1101-1108.

Seok Y and GynheungAn L 2015. Complex Regulatory Networks of Flowering Time in Rice. *Rice Research 03(03)*.

Tsuji H, Taoka K and Shimamoto K 2011. Regulation of flowering in rice: two florigen genes, a complex gene network, and natural variation. *Current Opinion in Plant Biology 14(1)*: 45–52.

Van den Burg HA, Kini RK, Schuurink RC and Takken FLW 2010. *Arabidopsis* small ubiquitin-like modifier paralogs have distinct functions in development and defense. *The Plant Cell 22*: 1998-2016.

Xu XM, Rose A, Muthuswamy S, Jeong SY, Venkatakrishnan S, Zhao Q and Meier I 2007. NUCLEAR PORE ANCHOR, the *Arabidopsis* Homolog of Tpr/Mlp1/Mlp2/Megator, is involved in mRNA export and SUMO homeostasis and affects diverse aspects of plant development. *The Plant Cell 19*: 1537-1548.

Yoshida S, Forno DA, Cock JH and Gomez KA1976. *Laboratory for physiological studies of rice, 3rd edition*, Manila: International Rice Research Institute.

Zhao J, Huang X, Ouyang X, Chen W, Du A, Zhu L, Wang S, Deng XW and Li S 2012. OsELF3-1, an ortholog of *Arabidopsis EARLY FLOWERING* 3, regulates rice circadian rhythm and photoperiodic flowering. *PLoS One 7(8)*: e43705.