A genome-wide gain-of-function analysis of rice genes using the FOX-hunting system

Hidemitsu Nakamura · Makoto Hakata · Kou Amano · Akio Miyao ·
Naoko Toki · Mariko Kajikawa · Jinhuan Pang · Naokuni Higashi ·
Shigeko Ando · Seiichi Toki · Miki Fujita · Akiko Enju · Motoaki Seki ·
Miki Nakazawa · Takanari Ichikawa · Kazuo Shinozaki · Minami Matsui ·
Yoshiaki Nagamura · Hirohiko Hirochika · Hiroaki Ichikawa

Abstract The latest report has estimated the number of rice genes to be ∼ 32 000. To elucidate the functions of a large population of rice genes and to search efficiently for agriculturally useful genes, we have been taking advantage of the Full-length cDNA Over-eXpresser (FOX) gene-hunting system. This system is very useful for analyzing various gain-of-function phenotypes from large populations of transgenic plants overexpressing cDNAs of interest and others with unknown or important functions. We collected the plasmid DNAs of 13 980 independent full-length cDNA (FL-cDNA) clones to produce a FOX library by placing individual cDNAs under the control of the maize *Ubiquitin-1* promoter. The FOX library was transformed into rice by *Agrobacterium*-mediated high-speed transformation. So far, we have generated approximately 12 000 FOX-rice lines. Genomic PCR analysis indicated that the average number of FL-cDNAs introduced into individual lines was 1.04. Sequencing analysis of the PCR fragments carrying FL-cDNAs from 8615 FOX-rice lines identified FL-cDNAs in 8225 lines, and a database search classified the cDNAs into 5462 independent ones. Approximately 16.6% of FOX-rice lines examined showed altered growth or morphological characteristics. Three super-dwarf mutants overexpressed a novel *gibberellin 2-oxidase* gene, confirming the importance of this system. We also show here the other morphological alterations caused by individual FL-cDNA expression. These dominant phenotypes should be valuable indicators for gene discovery and functional analysis.

Hidemitsu Nakamura and Makoto Hakata have contributed equally.

H. Nakamura · M. Hakata · K. Amano · A. Miyao · N. Toki ·
M. Kajikawa · J. Pang · S. Ando · Y. Nagamura ·
H. Hirochika · H. Ichikawa
Division of Genome and Biodiversity Research, National
Institute of Agrobiological Sciences, 2-1-2 Kannondai, Tsukuba,
Ibaraki 305-8602, Japan

Present Address:
K. Amano
Bioresource Information Division, BioResource Center, RIKEN,
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

Present Address:
N. Toki
Institute for Advanced Biosciences, Keio University, 246-2
Mizukami, Kakuganji, Tsuruoka, Yamagata 997-0052, Japan

Present Address:
J. Pang
Institute of Microbiology, Chinese Academy of Sciences,
13 Beiyitiao, Zhongguancun, Beijing 100080, P. R. China

N. Higashi · S. Toki · H. Ichikawa (✉)
Division of Plant Sciences, National Institute of Agrobiological
Sciences, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan
e-mail: hichkw@affrc.go.jp

M. Fujita · K. Shinozaki
Gene Discovery Research Group, Plant Science Center, RIKEN
Tsukuba Institute, 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074,
Japan

A. Enju · M. Seki · M. Nakazawa · T. Ichikawa ·
K. Shinozaki · M. Matsui
Plant Functional Genomics Research Group, Plant Science
Center, RIKEN Yokohama Institute, 1-7-22 Suehiro-cho,
Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Present Address:
A. Enju
Australian Centre for Plant Functional Genomics, Plant
Genomics Centre, Hartley Grove, Urrbrae, South Australia 5064,
Australia

Abbreviations

FL-cDNA	full-length cDNA
FOX-hunting system	full-length cDNA Over-eXpressor gene-hunting system
PUbi-1	maize *Ubiquitin-1* promoter
hpt	hygromycin phosphotransferase
35S	cauliflower mosaic virus *35S*
Tnos	*nopaline synthase* terminator
RT	reverse transcription
Hyg	hygromycin B
KOME	knowledge-based Oryza Molecular biological encyclopedia
RAP-DB	the Rice Annotation Project Database
GA	gibberellin

Introduction

Rice is one of the most important crops for humans and is recognized as the best model plant of monocot crops (Khush 2005). The International Rice Genome Sequencing Project (2005) completed sequencing of the entire rice genome in 2005, and from the total number of expressed loci and unmapped rice full-length cDNA (FL-cDNA) clones the Rice Annotation Project (2007) estimated the gene number of rice to be ∼32 000. Among these genes, 28 540 are candidates for protein-coding genes. About one-fourth (7189 loci) of the potential open-reading frames (ORFs) have had functions assigned to them by BLASTX searches (Rice Annotation Project 2007). Appropriate biological experiments are needed to uncover the functions of the remaining genes.

To accelerate the identification and characterization of rice genes with so far unknown functions, various rice resources, including sets of backcross inbred lines (BILs), nearly isogenic lines (NILs), and chromosome segment substitution lines (CSSLs), have been produced and are publicly available for mapping and the subsequent map-based cloning of agronomically important genes and quantitative trait loci (QTLs) (Lin et al. 1998; Ebitani et al. 2005). Infrastructure for rice bioinformatics (reviewed in Sasaki et al. 2005), including such integrated databases as RAP-DB (Ohyanagi et al. 2006; Rice Annotation Project 2007), the TIGR Rice Genome Annotation Database (Ouyang et al. 2007), BGI-RIS (Zhao et al. 2004), and MOsDB (Karlowski et al. 2003), has been developed and is publicly available. The rice FL-cDNAs, clustered into 28 469 non-redundant clones (Rice Full-Length cDNA

Consortium 2003), and the rice proteome database (Komatsu 2005) also have great potential to promote gene discovery and the elucidation of gene function.

For the large-scale identification of rice gene functions, several systems and tools to generate loss-of-function mutants have been developed (Hirochika et al. 2004; An et al. 2005a). Databases compiled for transferred DNA (T-DNA) insertion lines (An et al. 2003; Chen et al. 2003; Ryu et al. 2004; Sallaud et al. 2004) and for populations insertion-tagged by the retrotransposon *Tos17* (Miyao et al. 2003, 2007) or the transposable element *Ds* (Upadhyaya et al. 2002; Kolesnik et al. 2004; van Enckevort et al. 2005; Park et al. 2007) have been constructed. These loss-of-function resources and databases are definitely of enormous importance for the tagging and hunting of non-redundant genes in the rice genome. However, 29% of rice genes are estimated to exist as clustered and redundant gene families (IRGSP 2005). Therefore, it is difficult to find loss-of-function phenotypes for these duplicated genes. Moreover, when knockouts of genes confer embryonic lethality or severe developmental defects upon the plants, it is impossible to find out the mutant phenotypes corresponding to the mutated genes.

To solve such problems, the activation tagging, a gain-of-function strategy, has been adopted. This technique involves the introduction of a transcriptional enhancer at random into a plant genome to activate transcription of the gene(s) adjacent to the enhancer (Jeong et al. 2002, 2006; Nakazawa et al. 2003; Hsing et al. 2007; Mori et al. 2007). Although this system is effective in changing the transcription level of the gene(s), the transcriptional enhancer often affects the expression of multiple genes, resulting in complex phenotypes. Moreover, T-DNA insertions preferentially occur in gene-rich regions and are not distributed uniformly throughout the entire rice genome (An et al. 2003; Chen et al. 2003; Sallaud et al. 2004).

Taking advantage of the availability of a large collection of Arabidopsis FL-cDNA clones (Seki et al. 2002), Ichikawa et al. (2006) recently developed an alternative gain-of-function approach for the systematic elucidation of gene functions in Arabidopsis—the so-called Full-length cDNA Over-eXpressor gene-hunting system (FOX-hunting system). This system uses ectopic overexpression of a single, or limited numbers of, FL-cDNA(s) in individual transgenic plants. It generates large numbers of dominant mutations, enabling the comprehensive characterization of novel and important traits and identification of the causal genes.

The FOX-hunting technology is in principle applicable not only to Arabidopsis but to any other plants in which transformation technology aided by *Agrobacterium* is available. A large collection of FL-cDNA clones from Arabidopsis, rice and whatever origin can be ectopically

overexpressed to cause changes in phenotypes. As one of the most important and basically studied crops, rice is the most desirable plant for application of the FOX-hunting technique, as a recently developed rapid and highly efficient transformation system for rice (Toki et al. 2006), a large population of rice FL-cDNA clones (Rice Full-Length cDNA Consortium 2003), and the genome sequence information (IRGSP 2005) are available.

We have developed a FOX-hunting system in rice for the systematic functional analysis of rice genes. We have generated approximately 12 000 FOX-rice lines. By genomic PCR analysis, 81.4% of the FOX-rice lines so far analyzed have shown amplification of single fragments containing FL-cDNAs. Sequence analysis of the cDNA-containing fragments revealed that 5462 independent FL-cDNAs were inserted in 8225 FOX-rice lines carrying identified FL-cDNAs. Phenotypic observation of the FOX-rice plants revealed that approximately 16.6% of the tested lines showed altered phenotypes such as in cell and tissue proliferation, organ morphology, plant height, growth habit, heading date, and seed fertility. Here, we present information on extremely dwarfed mutants produced by overexpressing a novel *gibberellin 2-oxidase* gene. We also give details of several morphological mutants and the genes responsible. Our collection of FOX-rice lines should be a valuable tool for analyzing rice gene functions at the genome-wide scale and for efficiently identifying agronomically important genes.

Materials and methods

Plant materials, growth conditions, and rice transformation

Rice plants (*Oryza sativa* L. 'Nipponbare') were grown in growth chambers under white fluorescent light [a cycle of 14-h light (28°C)/10-h dark (25°C)] at 28°C and 60% relative humidity. Rice transformation was performed as described (Toki et al. 2006).

Plasmid construction

To make a double-stranded *Sfi*I linker with *Xba*I-cohesive ends on both sides, 1 pmol each of phosphorylated *Sfi*I-linker A (5′-CTAGACCTGCAGG<u>GGCC</u>AAATC<u>GGCC</u>GAGCTCGAATTCGTCGACCTCGAG<u>GGCC</u>ATAAG<u>GGCC</u>-3′, *Sfi*I sites underlined) and phosphorylated *Sfi*I-linker B (5′-CTAGA<u>GGCCC</u>TTAT<u>GGCC</u>CTCGAGGTCGACGAATTCGAGCTC<u>GGCC</u>GATTT<u>GGCC</u>CTGCAGGT-3′, *Sfi*I sites underlined) were mixed and annealed at room temperature for 1 h. The resultant linker was then

inserted into the *Xba*I site in the pBI-based plant transformation vector pBIG2113-S (Fujita et al. 2007) to generate pBIG2113-SfiIL. The maize *Ubiquitin-1* promoter (PUbi-1) was excised by using *Hin*dIII and *Sbf*I from pUbi2.0/HSb; it carried a 2.0-kb *Hin*dIII-*Sbf*I fragment derived from the 5′-upstream region of the maize *Ubiquitin-1* gene (Christensen et al. 1992). The promoter was then inserted into the same restriction sites of pBIG2113-SfiIL to construct pBIGS-PUbi-SfiL. A 35S-*hpt*-Tnos [a *hygromycin phosphotransferase* (*hpt*) gene governed by cauliflower mosaic virus *35S* (35S) promoter and the *nopaline synthase* terminator (Tnos)] fragment was excised with *Spe*I and *Kpn*I from pSKHyg [a derivative of pBluescript SK(+) bearing a 2.2-kb *Hin*dIII–*Eco*RI cassette that consisted of 35S-*hpt*-Tnos] and ligated at compatible *Xba*I and *Kpn*I sites of pLITMUS28 (New England Biolabs, Beverly, MA, USA) to generate pLIT-35SHyg. Then the 35S-*hpt* fragment was excised from pLIT-35SHyg with *Hin*dIII, *Mfe*I, and *Sca*I, and the resulting two fragments (a 0.5-kb *Hin*dIII–*Mfe*I and a 1.3-kb *Mfe*I–*Sca*I fragments) were simultaneously ligated at the *Hin*dIII and *Sca*I sites of pBIGS-PUbi-SfiL to construct the pRiceFOX vector.

The pRiceFOX and FL-cDNA-inserted pRiceFOX plasmids were transferred into *Agrobacterium tumefaciens* strain EHA105 (Hood et al. 1993) by electroporation.

Construction of rice-FOX *Agrobacterium* library

A rice FL-cDNA expression library was made according to the method of Ichikawa et al. (2006). In brief, a mixture of 13 980 rice FL-cDNAs that were cloned in *Sfi*I sites of the lambda-FLC-1-B vector (Rice Full-Length cDNA Consortium 2003) was digested with *Sfi*I and ligated, in the forward orientation, into the compatible *Sfi*I sites on a binary vector, pRiceFOX (see Results section for details on the vector). The ligated mixture was transformed into *E. coli* to produce a rice FL-cDNA library. Then, *Agrobacterium* EHA105 was transformed with a plasmid mixture extracted from the cDNA *E. coli* library. The resulting transformants were collected to construct the rice FL-cDNA *Agrobacterium* library to be introduced into rice.

PCR analysis of transgenes

A MagAttract 96 DNA Plant Kit (QIAGEN, Hilden, Germany) was used according to the protocol provided by the manufacturer to extract genomic DNA from the leaf blades of each transgenic plant grown for 2 weeks after transfer to soil. Transgenes inserted into the genome of FOX-rice lines were amplified by PCR by using one of the three thermostable DNA polymerases, TaKaRa Ex Taq polymerase

(Takara Bio, Ohtsu, Japan), TaKaRa LA Taq with GC Buffer (Takara Bio), or PrimeSTAR HS DNA Polymerase (Takara Bio); a forward primer from one of the two as follows (located on Pubi-1): FOX5-1 (5′-CTTTGGG GAATCCTGGGGATGGCTCTAGCCGTTCCGCAGACG GGA-3′) or FOX5-2 (5′-AGCCCTGCCTTCATACGC TATTTATTTGCTTGGTACTGTTTC-3′); and a reverse primer (located on Tnos), FOX3 (5′-GAAACTTTAT TGCCAAATGTTTGAACGATCGGGGAAATTCGAG-3′) (Fig. 1A). The conditions for PCR with TaKaRa Ex Taq were 30 s at 94°C for denaturation, followed by 33 cycles of 10 s at 98°C for denaturation, 30 s at 58 to 62°C for annealing, and 5 min at 72°C for elongation. Those for PCR with TaKaRa LA Taq or PrimeSTAR HS were 1 min at 94°C for denaturation, followed by 33 cycles of 10 s at 98°C for denaturation and 5 min at 68°C for annealing and elongation. The PCR products were checked for size(s) and numbers by electrophoresis and subjected to DNA sequence analysis by the direct-sequencing method with FOX5-2 and/ or FOX3 primer(s) in an ABI3100 or ABI3730LX sequencer (Applied Biosystems, Foster City, CA, USA). Sequenced FL-cDNAs were identified by using the Knowledge-based Oryza Molecular biological Encyclope-dia (KOME, http://www.cdna01.dna.affrc.go.jp/cDNA/; Rice Full-Length cDNA Consortium 2003) and the Rice Annotation Project Database (RAP-DB, http://www.rapdb. dna.affrc.go.jp/; Ohyanagi et al. 2006; Rice Annotation Project 2007).

The Tnos fragment was similarly amplified by using the forward primer N5 (5′-GAGCTCGAATTTCCCCGATC GTTCAAAC-3′) and the reverse primer N3 (5′-CCCGA TCTAGTAACATAGATGACACCGC-3′) (Fig. 1A). The conditions for PCR with PrimeSTAR HS were 1 min at 94 °C for denaturation, followed by 30 cycles of 10 s at 98°C for denaturation, 5 s at 60°C for annealing, and 30 s at 72°C for elongation. The presence and size of each PCR product were checked by electrophoresis.
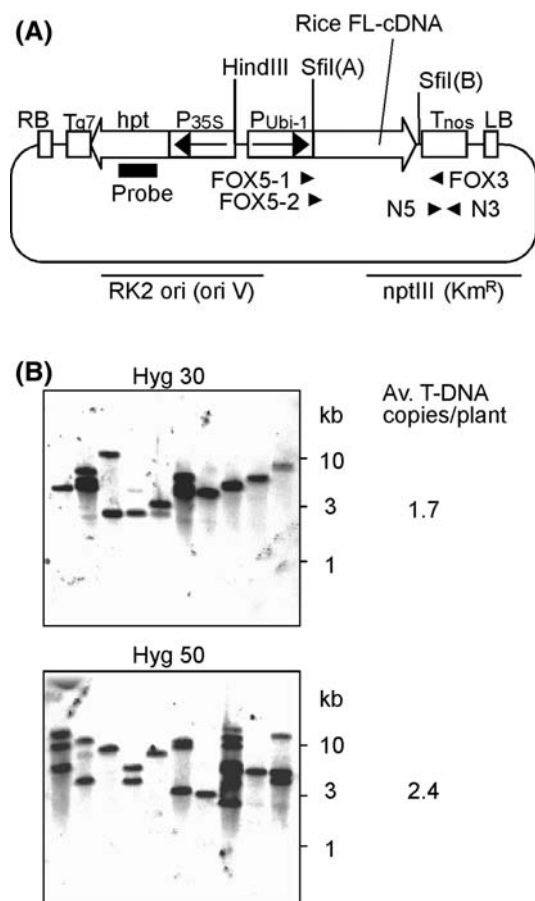
Southern blot analysis

Southern blot hybridization was performed according to standard protocols (Sambrook et al. 1989). Genomic DNA was extracted from the leaf blades of each transgenic plant with a DNeasy Plant Mini Kit (QIAGEN) according to the protocol provided by the manufacturer. Ten micrograms of rice genomic DNA was digested with HindIII and frac-tionated in a 0.8% agarose gel. An hpt gene fragment was PCR-amplified with 5′ primer (5′-CCGATTCCGGAAG TGCTTGAC-3′) and 3′ primer (5′-TGGGAATCCCC GAACATCGCC-3′) and pRiceFOX as a template. The hpt fragment was labeled by the ECL Direct Nucleic Acid Labeling and Detecion System (GE Healthcare Bio-Sciences, Piscataway, NJ, USA) according to the procedure recommended by the manufacturer.

Fig. 1 Transformation of rice with pRiceFOX, a binary Ti plasmid vector for overexpressing FL-cDNAs in plants. (A) Schematic structure of the pRiceFOX plasmid. The expression vector pRiceFOX allows unidirectional cloning of individual rice FL-cDNAs from RIKEN by using two different SfiI sites [SfiI(A) and SfiI(B)]. P35S, CaMV 35S promoter; PUbi-1, maize Ubiquitin-1 promoter; Tg7 and Tnos, polyadenylation signals from gene 7 and nopaline synthase (nos) in the T-DNA, respectively; hpt, hygromycin resistance gene; LB, left border; RB, right border. Arrows: directions of transcription; arrowheads: positions of the primers for genomic PCR. (B) Genomic Southern blot analysis of transgenic rice plants (T$_0$ generation) transformed with the empty pRiceFOX vector. Trans-genic plants were cultured and regenerated in the presence of hygromycin at 30 mg/L (Hyg 30) or 50 mg/L (Hyg 50). HindIII-digested DNAs were separated by agarose gel electrophoresis and hybridized with the probe specific for the hpt gene, indicated in (A)

RNA preparation and RT-PCR

An RNeasy Plant Mini Kit (QIAGEN) was used according to the manufacturer's instructions to extract total RNA from the leaf blades of each transgenic plant grown for 2 weeks after transfer to soil. First-strand cDNAs were synthesized from each RNA preparation (1-μg/reaction)

with an oligo(dT) primer by using an ExScript RT reagent Kit (Takara-Bio) in a total volume of 20-μL, in accordance with the manufacturer's instructions.

The specific sequences of each of the primer pairs used in semi-quantitative and quantitative reverse transcription PCR (RT-PCR) are listed in Table S1. Primer pairs amplifying cDNAs for *Actin1* (AK100267; 5′-CTTCA TAGGAATGGAAGCTGCGGGTA-3′ and 5′-TTCCTGT GCACAATGGATGG-3′) and *UBQ5* (AK061988; 5′-AC CACTTCGACCGCCACTACT-3′ and 5′-ACGCCTAAG CCTGCTGGTT-3′) were used as loading controls for semi-quantitative RT-PCR. The *Actin1* primers were also used as internal controls for normalization of the quantitative RT-PCR reaction (Jain et al. 2006).

Semi-quantitative RT-PCR was performed with 1 μL of template cDNA per 50-μL reaction with TaKaRa Ex Taq for 30 s at 94°C, followed by 28 cycles of 10 s at 98°C, 30 s at 60°C, and 30 s at 72°C. Quantitative RT-PCR was performed with a Thermal Cycler Dice Real Time System (Takara-Bio) using SYBR Premix Ex Taq (Takara-Bio) and 1 μL of template cDNA per 25-μL reaction, in accordance with the manufacturer's instructions. The threshold cycle (Ct) was auto-calculated by the analysis software that accompanied the system. The expression level normalized to that of the endogenous control gene (*Actin1*) was calculated by relating the measured Ct to a standard curve obtained by diluting PCR-amplified DNA for which the exact DNA concentration was known. The resulting RT-PCR products were resolved by agarose gel electrophoresis followed by staining with ethidium bromide.

Results

Construction of a binary vector for overexpression of rice FL-cDNAs in rice

For the Arabidopsis FOX-hunting system, Ichikawa et al. (2006) constructed the binary plasmid pBIG2113SF from the pBI-based plasmid pBIG2113N. By the insertion of individual FL-cDNAs in forward orientation between two *Sfi*I sites producing different cohesive ends in the pBIG2113SF vector, each FL-cDNA was placed under the transcriptional control of the 35S promoter. For FOX-hunting in rice, we constructed the binary vector pRice-FOX (Fig. 1A) from pBIG2113-S (Fujita et al. 2007), a derivative of pBIG2113N. Because PUbi-1 is stronger than the 35S promoter in rice cells and is highly expressed in various rice tissues (Cornejo et al. 1993), we exchanged the 35S promoter in pBIG2113N with PUbi-1. We also replaced the *nopaline synthase* promoter of the marker gene, *hpt*, with 35S for more efficient selection of transformed calluses (Dekeyser et al. 1989).

pRiceFOX was transformed into rice according to the method of Toki et al. (2006) by using two selection media, containing hygromycin B (Hyg) at 30 or 50 mg/L. The transformation efficiency was 38.9% in the former medium (Hyg30 selection group) and 26.8% in the latter (Hyg50). Ten plants were randomly chosen from each group and analyzed by Southern blotting with the *hpt* fragment as a probe (Fig. 1A). Single fragments were detected in 5 out 10 plants on Hyg30 (average T-DNA copy number per plant: 1.7) and 4 on Hyg50 (2.4) (Fig. 1B). These results imply that the use of higher concentrations of Hyg may lead to the generation of smaller numbers of transgenic rice plants with greater T-DNA copy numbers in their genomes. On the basis of these data, and with a view to economy, we thereafter conducted the rice transformation with 30 mg/L of Hyg.

Construction of a FL-cDNA expression library in *Agrobacterium* containing 13 980 rice FL-cDNAs

Rice FL-cDNA clones were originally collected and sequenced by two groups, the Laboratory of Genome Sequencing and Analysis Group of the Foundation for Advancement of International Science (FAIS) in Tsukuba, and the Laboratory for Genome Exploration Research Group, the Genomic Sciences Center, Institute of Physical and Chemical Research (RIKEN) in Yokohama. They used different vector systems for the construction of cDNA libraries (Rice Full-Length cDNA Consortium 2003).

We aliquoted 13 980 independent rice FL-cDNAs from RIKEN in nearly equal amounts to generate a normalized rice FL-cDNA mixture. After construction of the rice FL-cDNA overexpression library in agrobacteria (rice-FOX *Agrobacterium* library), plasmid DNAs were isolated from randomly chosen *Agrobacterium* colonies. Digestion of plasmids from 100 colonies with *Sfi*I followed by assessment by agarose gel electrophoresis revealed that 96 colonies contained cDNA fragments and only four colonies contained no cDNA inserts (data not shown). Three out of the 96 colonies contained two different cDNA fragments, and seven colonies contained three different cDNAs. Since *Sfi*I-digestion of the pRiceFOX vector and the RIKEN-cDNA clones generated different, but intercompatible, cohesive ends, it is probable that three different *Sfi*I-digested cDNA fragments were inserted into the *Sfi*I-digested pRiceFOX vector simultaneously.

Generation and PCR analysis of FOX-rice lines

FOX-rice lines were generated by transforming rice with the rice-FOX *Agrobacterium* library. To date we have

produced about 12 000 independent FOX-rice plants (T_0 generation) with Hyg resistance and, after they have differentiated roots, transplanted them into soil.

To identify integrated cDNAs in these lines, we have analyzed 10 219 FOX-rice lines by PCR, using their genomic DNAs as templates and T-DNA-specific primers (FOX5-1/FOX5-2 and FOX3, as shown in "Materials and Methods" and Fig. 1A). Figure 2A is a typical example of the PCR products from 21 plants from randomly selected FOX-rice lines. The sizes of the introduced rice FL-cDNAs were variable, ranging 1.1 to 3.4 kb. Among the 10 219 lines analyzed by PCR, 332 possessed a fragment of the same size as that from lines transformed with the empty vector ("Empty vector" in Fig. 2B, C). The PCR products from 8322 lines (81.4% of the FOX-rice lines analyzed) showed single bands after electrophoresis ("Single fragment" in Fig. 2B, C), and those from 293 lines showed multiple fragments ("Multiple fragments" in Fig. 2B, C). The remainder, 1272 lines, gave no fragment amplification, suggesting that only the *hpt* marker gene was integrated into their genomes, without the accompanying FL-cDNA(s). To confirm the above possibility, we performed PCR with a pair of specific primers for the Tnos sequence (N5 and N3 in Fig. 1A). The Tnos fragment ("No amplification/Tnos+" in Fig. 2B, C) was amplified from 254 of the 1272 lines. This indicates some difficulty in amplifying the cDNA fragments in these lines. Rice exon sequences exhibit higher GC contents (54.2%) (IRGSP 2005) than those from dicot species such as Arabidopsis (44.1%) (Arabidopsis Genome Initiative 2000). This tendency could have caused the difficulty in amplification of target fragments, even with the use of PCR conditions for templates with high GC contents (see "Materials and Methods" section for details). In contrast, no Tnos fragment was amplified from 1018 lines ("No amplification/ Tnos–" in Fig. 2B, C). This result suggested that only the *hpt* gene was integrated into the rice genome, whereas the FL-cDNA to be located at the LB (left border) side of the T-DNA was not. Thus, the subtotal of FOX-rice plants without cDNA integration was 1350 (13.2%) ("Empty vector" plus "No amplification/Tnos–" in Fig. 2B, C).

Among the 8615 FOX lines showing PCR-amplification of a single or multiple cDNA fragment(s) (Fig. 2C), amplification of one, two, three, four, and five fragments was detected in 8322, 277, 14, 1, and 1 plants, respectively. On average, 1.04 FL-cDNAs were integrated into each diploid genome. This is smaller than the number of T-DNA copies inserted into pRiceFOX-transformed rice (Fig. 1B). Fifteen FOX-rice lines amplifying single cDNA fragments were chosen randomly and subjected to DNA gel-blot analysis. Approximately 2.1 copies of T-DNA inserts per plant were observed (Fig. S1). The results indicate the frequent occurrence of multiple insertions of single T-DNAs
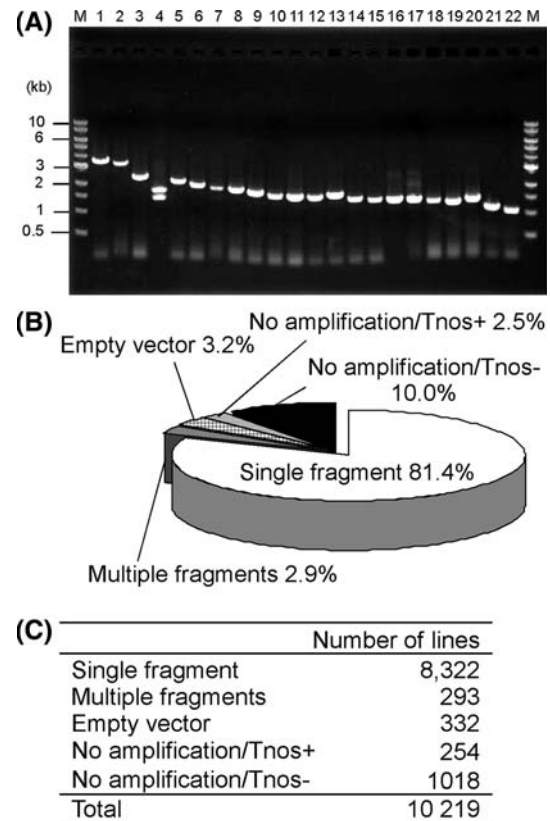

Fig. 2 PCR analysis of transgenes in FOX-rice lines. (A) Example of size distribution of the rice FL-cDNAs integrated in FOX-rice lines transformed with the rice-FOX *Agrobacterium* library. PCR-amplified fragments including FL-cDNA(s) were electrophoresed. In lane 4, two bands were amplified. Lane M, 1-kb DNA size markers. (B, C) Summaries of PCR analysis of transgenes. Genomic PCR data from 10 219 lines were compiled to show the patterns and numbers of cDNA-containing fragment(s) in the transgenic population. See text (Results section) for further details

into the host genome. In other words, approximately two copies of a T-DNA were inserted into each diploid genome in our FOX-rice plants (T_0 generation). A similar tendency was also observed in the Arabidopsis FOX-hunting system (Ichikawa et al. 2006).

Size distribution of FL-cDNAs inserted into the genomes of FOX-rice lines

To further examine the feature of FL-cDNAs integrated into our FOX-rice lines, we randomly selected 238 FOX-rice lines, which were generated from the early experiments and showed a single cDNA integration in each genome, and checked whether or not a bias in their size distribution of the FL-cDNAs is observed. We could not find any notable difference in the range of size distributions (Fig. 3) and the average sizes of FL-cDNAs between the
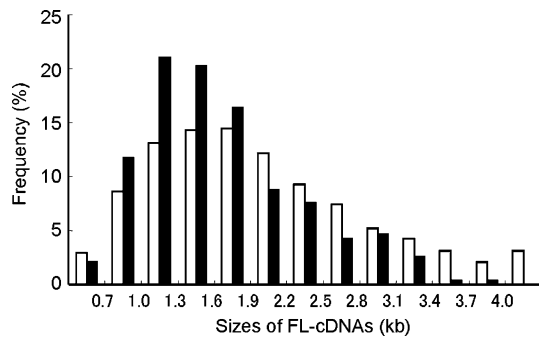
Fig. 3 Size distribution of the rice FL-cDNAs integrated into 238 FOX-rice plants (black bars), compared with 16 897 RIKEN rice FL-cDNAs cloned in the lambda-FLC vector (white bars)

238 FOX-rice lines (average insert size: 1.66 kb) and the original RIKEN rice FL-cDNA library comprised of 16 897 clones (1.99 kb).

Sequence analysis of FL-cDNAs integrated into FOX-rice lines

All the amplified PCR fragments bearing the integrated FL-cDNA(s) from the 8615 FOX-rice lines (Fig. 2C) were subjected to sequence analysis. Thus far, cDNA sequences from 8225 of the 8615 FOX lines have been successfully read, and database searches of the sequence data using KOME and RAP-DB identified 5462 independent FL-cDNAs. Of these 5462 FL-cDNAs, 2090 appeared twice or more.

Expression levels of transgenes in FOX-rice lines

To examine whether expression of transgenes was elevated in the FOX-rice lines, 24 transgenic lines were randomly selected (line names: CO004 to CP076 in Table S1). Single FL-cDNAs were inserted in 23 of these 24 lines; two different cDNAs were inserted in the remaining line, designated CO130 (Table S1). Expression of these 25 genes was examined by semi-quantitative RT-PCR analysis (Fig. 4) with individual gene-specific primers (Table S1). Expression of 22 genes was obviously enhanced in the corresponding FOX lines. The RT-PCR products for the transgenes in lines CO100 and CO120 (*AK071882* and *AK067076*, respectively) and that for one of the two transgenes in line CO130 (*AK069185*, labeled "CO130-2" in Fig. 4) showed faint bands. We therefore performed quantitative real-time RT-PCR analysis of these three genes. The transcript level of *AK071882* integrated in line CO100 was approximately 370 times higher than that in the control line. In line CO120, *AK067076* transcripts were accumulated to approximately 400% of those in the control line. However, there was no obvious difference in the expression level of *AK069185* between CO130 and the control. These results demonstrate that most of the transgenes were overexpressed in our FOX-rice lines.

Observation of phenotypes in the FOX-rice lines

Miyao et al. (2007) adopted 53 phenotype descriptors belonging to 12 classes for the classification of visible
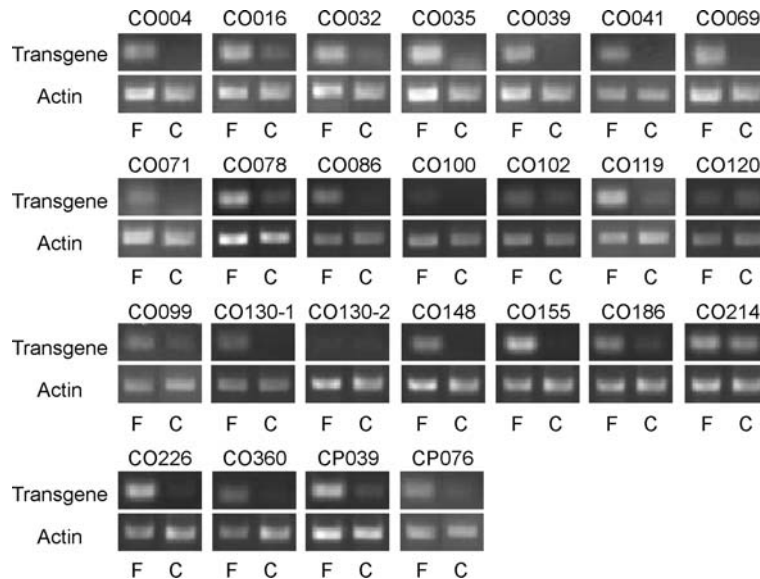

Fig. 4 Expression of transgenes in FOX-rice lines. Total RNA preparations were isolated from leaf blades of 24 independent transgenic lines and subjected to RT-PCR analysis. Each transgene was amplified with the gene-specific primers listed in Table S1. RT-PCR products were resolved in 2.5% agarose gel. In FOX line CO130, two FL-cDNAs were integrated (indicated as CO130-1 and CO130-2). For individual transgenic lines, upper panels represent the expression levels of introduced FL-cDNAs, and lower panels indicate those of internal *Actin1* used for loading adjustment. F, FOX-rice lines; C, control lines transformed with the empty pRiceFOX vector

phenotypes in the *Tos17* insertion lines. Other than the 53-phenotype descriptors, we added 8 possible phenotypes (3 classes) related to the characteristics on callus, regenerants, and root. According to the 61-phenotype descriptors (15 classes), we observed and classified the visible phenotypes of the T_0 plants of 9021 FOX-rice lines. Our classification revealed that 1496 out of the 9021 lines (16.6%) had altered phenotypes (Table S2).

Several FL-cDNAs conferred interesting and remarkable phenotypes upon the corresponding FOX-rice lines (Fig. 5 and Table 1). To confirm the reproducibility of the altered phenotypes, those cDNAs were individually subcloned into the pRiceFOX vector and then reintroduced into rice. The phenotypes of these lines (retransformants) are summarized in Table 1. All the reintroduced lines showed the same phenotypes as those of the original FOX lines, except for

line BL276, which showed a partial difference in the retransformants (Table 1). Line BL276 had drooping leaves in the early and middle stages of their growth, and its plant height was greater than in control plants at the maturing stage. BL276-RE, the reintroduced lines bearing the same cDNA as that integrated in line BL276, also exhibited drooping leaves, but the retransformant plants were no higher than the wild-type plants. Furthermore, the BL276-RE lines were pale green, whereas the original BL276 line was a normal green.

Accordingly, it is likely that the position or pattern of integration of the T-DNA insert(s) into the genome of the original FOX line (BL276) affected the expression of neighboring genes that could be involved in plant growth. At least, however, we observed reproducibility of the "drooping leaf" phenotype in the BL276 and the BL276-RE lines.

Fig. 5 Representative phenotypes of FOX-rice lines. Line names are indicated in the corresponding photographs. T_0 plants were grown for 7 weeks (**A**) or 14 weeks (**B**, **C**, **F**) on soil after transfer from hormone-free medium, or for 2 weeks on hormone-free medium after transfer from regeneration medium (**E**). (**D**) T_1 seeds were sown and grown on callus-induction medium containing 2,4-dichlorophenoxyacetic acid at 2 mg/L (Toki et al. 2006). Bars = 10 cm in **A**, **C** and **F**, 1 cm in **E**, and 0.1 cm in **D**
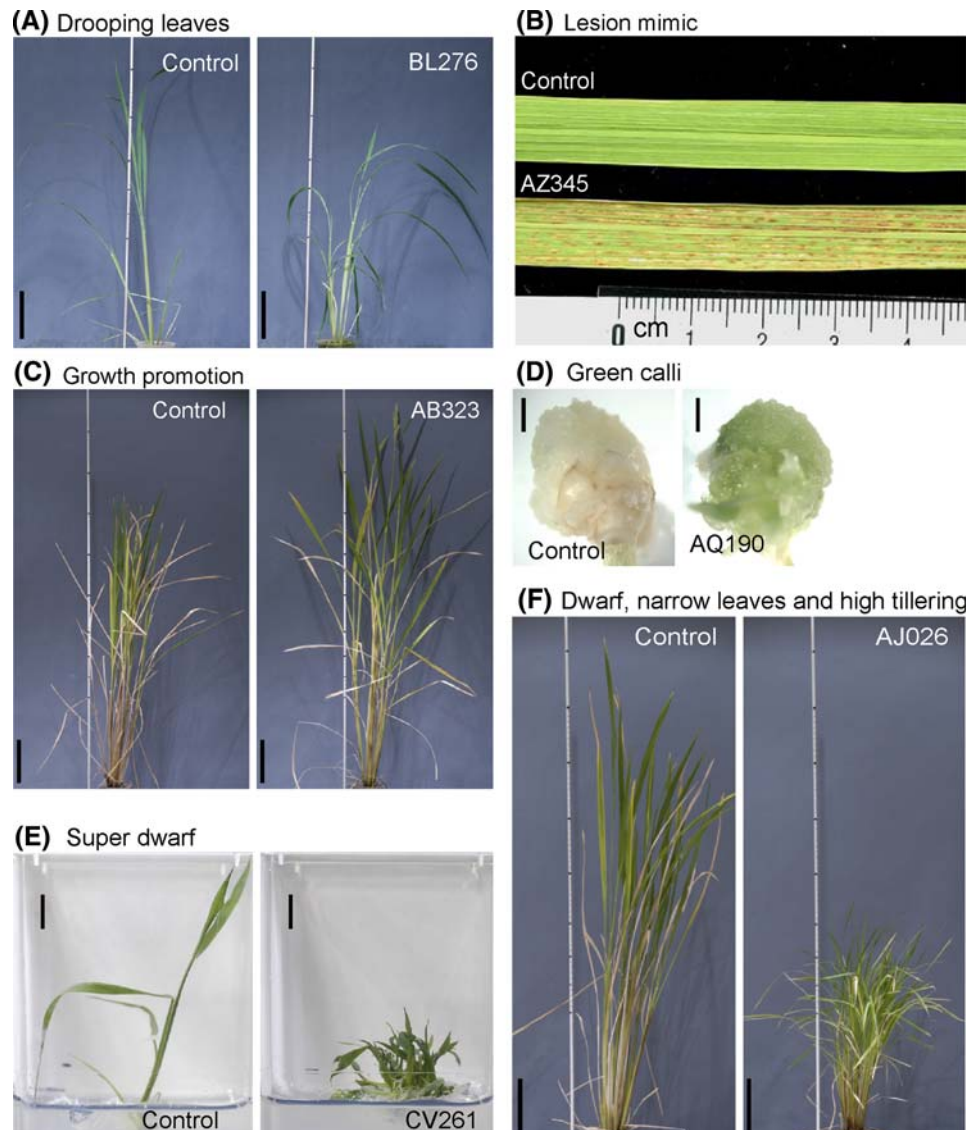
Table 1 Reproducibility by retransformation of altered phenotypes appearing in the FOX-rice lines

Line number	Putative gene function	Observed phenotypes	
		Original FOX-rice line	Retransformant
AC214 AG244 AM249	Gibberellin 2-oxidase	Super-dwarf	Super-dwarf
BD108 BB089	Phospholipid/glycerol acyltransferase	Weak and lethal	Weak and lethal
BL276	Isopenicillin N synthase family	Drooping leaf and long plant height	Drooping leaf and pale green
AZ345	Mu2 adaptin subunit (AP50) of AP2	Lesion mimic	Lesion mimic
AB323 AG163 AZ310 AZ326 AZ337	ZIM-domain–containing protein	Growth promotion	Growth promotion
AI109 AQ190 AX290 BW244	Golden2-like transcription factor (GLK1)	Green calluses	Green calluses
AJ026 AV127 CC179	Non-protein–coding transcript	Dwarf, narrow leaf, and high tillering	Dwarf, narrow leaf, and high tillering
CV261	RING-finger–containing protein	Super-dwarf	Super-dwarf
AG186	Eukaryotic translation-initiation factor 4G	Dwarf	Dwarf

Putative gene functions are from the Rice Annotation Project Database (RAP-DB; Ohyanagi et al. 2006; The Rice Annotation Project 2007) and the Database of full-length cDNA clones from *japonica* rice (KOME; Rice Full-Length cDNA Consortium 2003)

These results indicate that the various phenotypes appearing in the FOX-rice lines were brought about by the individual introduction of a variety of FL-cDNAs under the control of the maize *Ubi-1* promoter.

In the course of production of the FOX-rice lines, albino plants occasionally appeared. We randomly chose 12 albino lines and individually reintroduced their transgenes. None of the reintroduced lines was albino or pale green (data not shown). Thus, we concluded that most albino phenotypes in our FOX-rice lines were generated by other factors, such as somatic variation(s) during tissue culture, and not from overexpression of transgenes.

The weak growth and lethality found in line BD108 were reproducibly observed in the reintroduced lines. Line BD108 grew weakly and died 7 weeks after transfer to soil. A rice FL-cDNA, *AK071961*, was integrated into this line. The deduced protein encoded by *AK071961* was annotated in RAP-DB as a phospholipid/glycerol acyltransferase. Since rice growth at the early stage is especially affected by changing environmental conditions, it seemed unconvincing that the phenotypes observed in line BD108 were caused by the overexpression of *AK071961*. To confirm this, we reintroduced *AK071961* FL-cDNA into wild-type rice, obtained 13 independent *AK071961*-FOX transgenic lines, and then observed their growth after transfer to soil. All of them grew weakly compared with control plants and died 4 to 11 weeks after transfer to soil (Fig. 6A, B). Total RNA was extracted from four independent *AK071961*-FOX lines, and semi-quantitative RT-PCR analysis was performed. Although variable lifetimes were observed, the expression levels of *AK071961* in all four *AK071961*-FOX lines were similar (Fig. 6C). The final plant heights and lifetimes of the 13 *AK071961*-FOX lines were largely dependent on their initial plant heights (Fig. 6B). These data strongly suggest that overexpression of *AK071961* caused the growth defect and death in the transgenic rice plants.
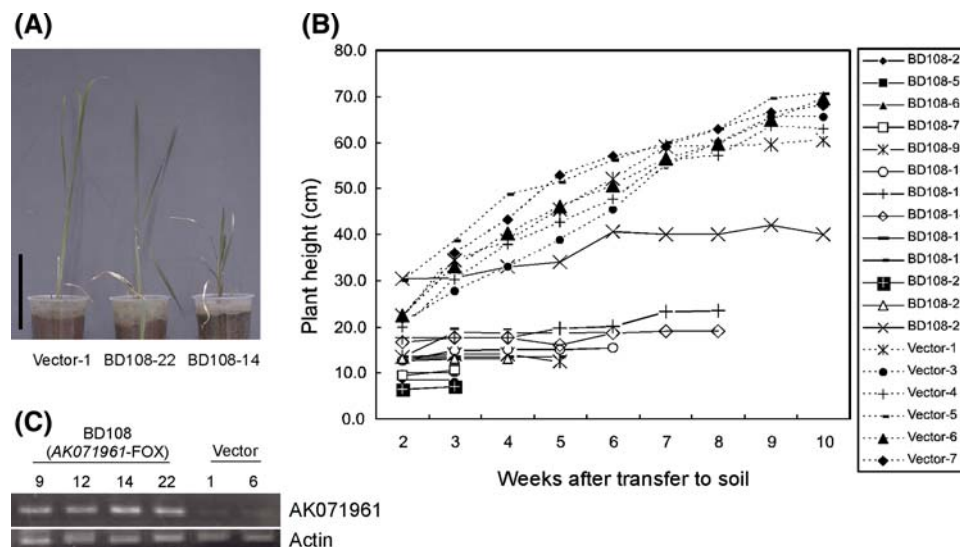
Fig. 6 Altered phenotypes related to the *AK071961*-FOX retransformed lines. (**A**) Growth of T_0 plants of the *AK071961*-FOX retransformed lines grown for 4 weeks on soil after transfer from hormone-free medium. Bar = 10 cm. (**B**) Growth curves of T_0 plants of the *AK071961*-FOX lines (shown as "BD108-#") and control lines transformed with empty pRiceFOX vector (shown as "Vector-#").

BD108-22, the most vigorous plant among the *AK071961*-FOX lines, died 11 weeks after transfer to soil, but all the control plants grew normally. (**C**) Semi-quantitative RT-PCR analysis of T_0 plants of the *AK071961*-FOX transgenic lines. Upper panel shows the transcript levels of *AK071961* cDNA. Lower panel represents those of *Actin1* used for loading adjustment

Characterization of three super-dwarf FOX lines

Three FOX-rice lines (AC214, AG244, and AM249) showed an extremely dwarfed phenotype. Plant height in these lines was less than 10 cm even 120 days after transfer to soil, whereas plants in the vector-control lines became more than 90 cm at their maturity. The leaf blades of the dwarf FOX plants were darker green and shorter and wider than those of the control plants, indicative of a typical phenotype for GA-deficient dwarf rice (Fig. 7A). Although control plants flowered approximately 90 to 100 days after transfer to soil, the FOX plants had not formed floral organs even 150 days after transfer to soil. Sequence analysis of the cDNA-containing fragments obtained by genomic PCR clearly demonstrated that the same rice FL-cDNA, *AK101758*, was integrated into all of these three lines. The gene was highly expressed in line AC214 (Fig. 7B). A RAP-DB and KOME database search revealed that the ORF encoded by *AK101758* was annotated as a gibberellin 2-oxidase (GA2ox).

The phenotypes of the three super-dwarf FOX lines described above are similar to that of the transgenic rice in which the *OsGA2ox1* gene is ectopically expressed (Sakamoto et al. 2001). To confirm further that overexpression of the *AK101758* confers a super-dwarf phenotype, we fused the cDNA 3′-downstream of the P*Ubi-1* constitutive promoter in the pRiceFOX vector and introduced it into wild-type rice. All primary transformants (12 independent lines) showed the same dwarf phenotype as observed in the three FOX lines. To determine whether

the dwarf phenotype was due to reduced levels of active gibberellins (GAs), 10 mM GA_3 was applied once to the reintroduced transformants 14 days after transfer to soil. The seedlings elongated gradually with GA_3 treatment (Fig. 7C), indicating that overexpression of the *OsGA2ox* FL-cDNA in rice reduces biologically active, endogenous GAs to levels that cannot sustain normal plant growth.

Quantification of GAs in the dwarf mutants and the enzyme activity of the *AK101758* gene product will be reported elsewhere.

Discussion

We produced approximately 12 000 transgenic rice lines by introduction of the rice-FOX *Agrobacterium* library comprising a maximum of 13 980 rice FL-cDNAs from RIKEN (Rice Full-Length cDNA Consortium 2003). Single cDNAs were inserted in 81.4% of the 10 219 FOX-rice lines analyzed by genomic PCR (Fig. 2B, C), and the average number of T-DNA copies per plant was estimated at 2.1 (Fig. S1). The size distribution pattern and average size of FL-cDNAs inserted into the genomes of randomly chosen 238 FOX-rice lines were very similar to those of the original RIKEN rice FL-cDNA library (Fig. 3). As we expected, a variety of transgenes were overexpressed in the corresponding FOX lines (Fig. 4). These features in our collection of FOX-rice lines indicate that our approach would be useful for the systematic and genome-wide gain-of-function analysis of rice genes.

Fig. 7 Super-dwarf phenotypes appeared in the FOX line AC214. (**A**) Growth of 14-week-old T_0 plants of the control line (left) and the AC214 line (right). Upper right photograph is a magnified view of the AC214 plant. (**B**) Semi-quantitative RT-PCR analysis of the AC214 plant (T_0). Upper panel shows the transcript levels of *AK101758* cDNA. Lower panel indicates those of *UBQ5* used as an internal control. (**C**) A retransformant with *AK101758* was sprayed with water (–GA_3) or 10 mM GA_3 (+GA_3), and was photographed 0 (0 d) and 8 days (8 d) after spraying. The super-dwarf phenotype of the *AK101758*-FOX retransformant was rescued by the GA_3 treatment. Bars = 5 cm
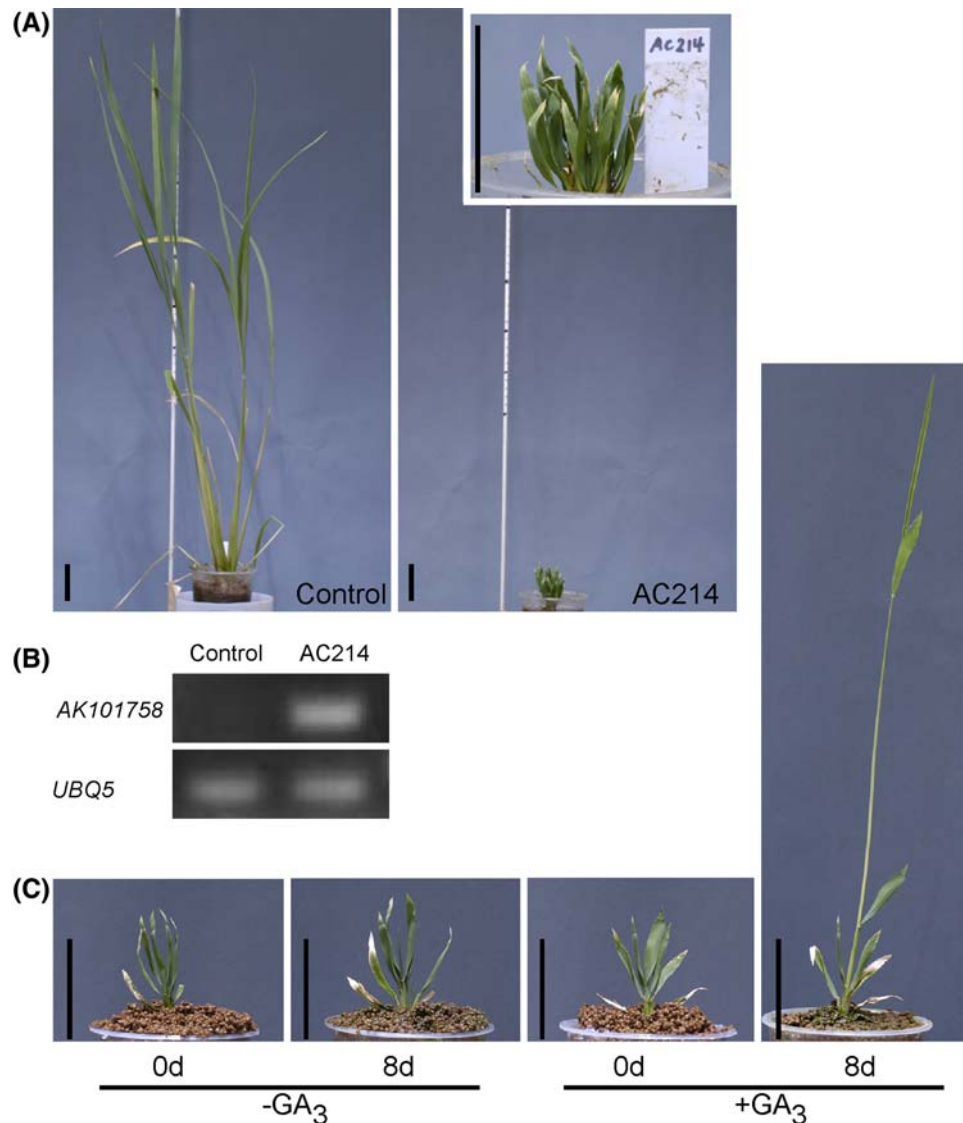
We analyzed the sequences of the cDNA-containing fragments generated by genomic PCR of 10 219 FOX-rice plants, and we identified 5462 independent rice FL-cDNAs, corresponding to 39.1% of the 13 980 rice cDNAs in the FOX *Agrobacterium* library. The number of rice genes is estimated to be ∼32 000 (Rice Annotation Project 2007). This means that we have introduced only 17.1% of them into the genomes of 10 219 FOX plants.

To increase the number of introduced cDNAs, the simplest strategy for us will be to further produce FOX-rice lines by using the overexpression library bearing at most 13 980 RIKEN FL-cDNAs. As described in the Results section, we identified 5462 independent cDNAs integrated in the 8225 FOX plants of which we could read the transgene sequences. If we apply these values to the equation of Clarke and Carbon (1976), which is an approximation of the binomial distribution, about 6930

FOX lines are needed to obtain the 5462 cDNAs out of 13 980 cDNAs. Similarly, to obtain a population of FOX lines with 10 000 independent FL-cDNA insertions using our *Agrobacterium* library, we would need to generate approximately 17 560 FOX plants in total. We also need ∼64 380 plants to cover 99% (13 840) of FL-cDNAs. The above estimated number of lines (6930) is about 84% of the number in our results (8225 lines). One plausible reason for this discrepancy may be redundancy in the rice-FOX *Agrobacterium* library. In fact, we observed the reiterated appearance (twice or more) of 2090 of the 5462 cDNAs. If we assume that the appearance of cDNAs can be approximated to a binomial distribution, 1653 FL-cDNAs, theoretically, appear twice or more from the 8225 lines. The higher frequency for the cDNA reiteration in our FOX lines might be partly due to potential bias during construction of the rice-FOX *Agrobacterium*

library. However, the size distribution and average size of FL-cDNAs inserted in 238 lines obtained from the early experiments was similar to those of the RIKEN rice FL-cDNA library (Fig. 3). Accordingly, the bias, if any, may not be a serious issue in regard to the reiteration.

Second, the growth rates of *Agrobacterium* cells harboring a variety of cDNAs may not be synchronized. *Agrobacterium* cells carrying a P*Ubi-1:gusA* chimeric reporter gene show blue-staining when treated with X-glucuronide (Nakamura et al. unpublished observation), indicating that the maize promoter is active and can drive the expression of rice FL-cDNAs in *Agrobacterium*. This may affect the growth of some populations in the agrobacteria. The third possibility may be the effect of ectopic overexpression of various cDNAs in the rice cells. In general, the expression of genes for transcription factors, signal transducers, etc. is kept at low levels, with tissue and temporal specificity, and is often inducible by biotic or abiotic factors (Cheong et al. 2002; Rabbani et al. 2003; Eulgem 2005). We used PUbi-1 for the ectopic and constitutive expression of RIKEN FL-cDNAs in rice. Because the transcription activity of PUbi-1 is high (Cornejo et al. 1993), if the abovementioned genes were to become constitutively active, we would hopefully see altered or exaggerated phenotypes, which could give us useful insights in deducing the gene functions responsible (i.e., this is an advantage of FOX-hunting). In contrast, such overexpression may sometimes become toxic, conferring critical and lethal effects on the growth of rice (Yamamoto et al. 2007). The latter effect could be considered a disadvantage of FOX-hunting, as it would reduce the sum of FL-cDNAs integrated into the rice genome.

The alternative strategy to increase the number of introduced cDNAs in FOX-rice plants would be to use another population of FL-cDNAs from FAIS (Rice Full-Length cDNA Consortium 2003) for the construction of novel rice-FOX *Agrobacterium* libraries. We recently prepared six *Agrobacterium* expression libraries each containing 273 to 4240 normalized FAIS FL-cDNAs (13 823 in total). The novel FOX *Agrobacterium* libraries have been used to transform rice.

In 10.0% of the transformed rice lines analyzed by genomic PCR ("No amplification/Tnos–" in Fig. 2B, C), no cDNA was integrated. In the pRiceFOX—a pBI-based binary vector—FL-cDNA cassettes can be cloned next to the left border (LB) of T-DNA, whereas the chimeric *hpt* gene is positioned in the vicinity of the RB. VirD2 protein covalently binds to, and protects the 5′ end at the RB of the single-stranded T-DNA from exonucleolytic degradation (Dürrenberger et al. 1989; Jasper et al. 1994). Accordingly, transgenes adjacent to the RB are more accurately integrated into the host genomes than those at the LB (Gelvin 2000). Incomplete insertion of LB regions in the T-DNAs into the host genome could eliminate integration of the FL-cDNAs. To circumvent such a disadvantage in our current system, we made a novel binary plasmid with a T-DNA region bearing the promoter from the rice *Actin1* gene (McElroy et al. 1990) for overexpression of 3′-downstream cDNAs to be located next to the RB and the *hpt* gene positioned on the LB side (Nakamura et al. unpublished results). This vector was used for the construction of the *Agrobacterium* expression libraries carrying FAIS FL-cDNAs.

We monitored the visible phenotypes in individual FOX lines, and we took photographs 2, 7, and 14 weeks after the lines were transplanted to soil. These phenotypic data, in addition to the sequence information on FL-cDNAs integrated into the FOX-rice lines, have been compiled to build up a database (DFR, for "A Database of FOX-Rice Lines"), which will be available for plant science researchers through our web site in future. For convenience, this database will be linked to such databases for rice functional genomics as the Rice *Tos17* Insertion Mutant Database (http://www.tos.nias.affrc.go.jp/~miyao/pub/tos17/index.html.en; Miyao et al. 2003, 2007), KOME (Rice Full-Length cDNA Consortium 2003), and RAP-DB (Ohyanagi et al. 2006; Rice Annotation Project 2007).

Phenotype databases of rice have been developed by several groups, e.g., the International Rice Information System (IRIS; Wu et al. 2005), Oryzabase (Kurata and Yamazaki 2006), Rice Mutant Database (RMD; Zhang et al. 2006), and Rice *Tos17* Insertion Mutant Database (Miyao et al. 2007). In addition to these useful databases, our database, DFR, would be expected to provide valuable information for researchers to uncover rice gene functions.

We found that 16.6% of the observed FOX-rice plants [9021 lines (T_0 generation)] showed altered phenotypes (Table S2). This rate was much higher than that of the activation tagging population (Jeong et al. 2002). Among the 9021 FOX lines, 104 plants (1.15%) showed weak growth. Growth of transgenic regenerants often suffers from the effects of environmental factors such as humidity, temperature, macro- and micro-nutrient deficiencies or excesses, and microbes (pathogens), especially in the initial growth stage after transfer to soil. Alternatively, in the course of tissue culture for selection and regeneration of transgenic plants, a variety of both epigenetic (e.g., DNA methylation, gene silencing, activation of retrotransposons and transposable elements) and genetic (e.g., nucleotide substitutions, deletions, insertions, rearrangements) changes could happen in the plant genome (Kaeppler et al. 2000; Cheng et al. 2006; Noro et al. 2007). These factors, i.e., abiotic and biotic factors and tissue-culture–induced variations, may more or less influence the growth conditions and consequently the phenotypes of FOX plants, and may somewhat increase the rate of phenotype alteration in our FOX lines.

However, one example of a FOX line (BD108) over-expressing *AK071961* cDNA, provisionally encoding phospholipid/glycerol acyltransferase, showed weak growth and early death (Fig. 6). This phenotype was reproducible in the retransformed plants. Consequently, the overexpression of *AK071961* was responsible for the mutant phenotype. A BLAST search using the deduced protein sequence encoded by *AK071961* as a query showed that the AK071961 protein was similar to Arabidopsis membrane-bound glycerol-3-phosphate acyltransferases (AtGPATs). Zheng et al. (2003) identified seven AtGPATs and observed knockout phenotypes of the *AtGPAT1* gene. The AK071961 protein contained all of the four previously defined acyltransferase domains common to the seven AtGPATs. The *AtGPAT1*-knockout mutants showed no growth or developmental defects at the vegetative stages but severely reduced male fertility and altered development in the tapetal cells. Deficiency in AtGPAT1 caused changes in the composition of several fatty acids in flower tissues and seeds. Although the putative AK071961 protein was most similar to AtGPAT2 and neither the knockout nor overexpression phenotypes of *AtGPAT2* have been reported yet, it is likely that the ectopic overexpression of *AK071961* caused alterations in fatty acid compositions, and these changes made the FOX-rice plants weak and lethal on the soil. Kachroo et al. (2003, 2004) demonstrated that a loss-of-function mutation in the *ACT1* gene in Arabidopsis encoding the soluble chloroplastic enzyme glycerol-3-phosphate acyltransferase reversed the salicylic-acid- and the jasmonic-acid (JA)-mediated defense phenotypes of the *ssi2* mutant (encoding stearoyl–acyl carrier protein desaturase) by increasing oleic acid levels. Although the AK071961 protein contains a conserved transmembrane domain, unlike the ACT1 protein, it may be speculated that alterations in fatty acid compositions by the overexpression of *AK071961* affected the actions of fatty-acid–derived signaling molecules such as JA and caused the severe growth defect in the FOX-rice lines. To confirm these working hypotheses, we need at least to answer the question of whether or not fatty acid composition is altered in the *AK071961*-FOX lines.

In the activation tagging of rice, 35S enhancers have been shown to enhance endogenous gene expression without altering the expression patterns of most T-DNA-tagged genes in transgenic rice (Jeong et al. 2002, 2006). In the FOX-hunting system of rice, the maize *Ubi-1* promoter directs ectopic overexpression of attached transgenes of any origin in transgenic plants. A typical example was found in the *GA2ox*-FOX lines, which showed an extremely dwarf phenotype with dark-green color. This overemphasized phenotype is a good indicator of the fundamental gene function(s). However, phenotypes caused by ectopic overexpression are not necessarily related to the native functions of the transgenes. Accordingly, much attention should be paid to deducing whether or not an observed phenotype is related to the native function of each transgene. To make such deductions with accuracy, the search for, and use of, loss-of-function mutants such as *Tos17* insertion lines (Miyao 2003, 2007) and T-DNA insertion lines (An et al. 2005b) would be of great use. If such gene-knockout lines are not available, the production of transgenic lines with targeted interruption of gene expression using RNAi (Miki and Shimamoto 2004) and of those carrying a dominant chimeric repressor to suppress the expression of each transcription factor (Hiratsu et al. 2003) would be effective. Appropriate combinations of gain-of-function resources, such as the FOX-rice plants described here, and loss-of-function resources could be powerful tools to promote functional genomics in rice. From another perspective, we may be able to utilize the potentially useful traits that appear in FOX-rice plants in molecular breeding programs for crop improvement, even though the traits are brought about not by simple exaggeration of native gene function(s) but by ectopic overexpression of the genes responsible.

Here, we describe only the visible phenotypes in our FOX-rice lines. It will be difficult to screen FOX lines exhibiting various agriculturally important traits (e.g., biotic and abiotic stress tolerance) under ordinary growth conditions. Using T_1 progeny seeds of FOX lines, screening for genes that confer resistance to blast fungus, salt tolerance, cadmium tolerance and/or accumulation, and tolerance to cold temperature and submergence is in progress, in collaboration with other researchers. These time-consuming screens are of great importance for gene discovery and in the manufacture of plant materials for crop improvement. Of similar importance are our efforts to generate the resources (rice cDNA expression libraries, FOX-rice plants and the progeny seeds) and the information on the resources (database of the FOX-rice lines) for the acceleration of rice gene-hunting.

rice-FOX *Agrobacterium* library; the Rice Genome Resource Center of the National Institute of Agrobiological Sciences for providing the rice FL-cDNA plasmid DNAs. We are grateful to Drs. Yoshihito Suzuki and Masatoshi Nakajima for their helpful suggestions on the GA2-oxidase study, and to Dr. Masaki Mori for his useful information on functional analysis of the AK071961-FOX line.

References

An G, Jeong DH, Jung KH, Lee S (2005a) Reverse genetic approaches for functional genomics of rice. Plant Mol Biol 59:111–123

An G, Lee S, Kim SH, Kim SR (2005b) Molecular genetics using T-DNA in rice. Plant Cell Physiol 46:14–22

An S, Park S, Jeong DH, Lee DY, Kang HG, Yu JH, Hur J, Kim SR, Kim YH, Lee M, Han S, Kim SJ, Yang J, Kim E, Wi SJ, Chung HS, Hong JP, Choe V, Lee HK, Choi JH, Nam J, Kim SR, Park PB, Park KY, Kim WT, Choe S, Lee CB, An G (2003) Generation and analysis of end sequence database for T-DNA tagging lines in rice. Plant Physiol 133:2040–2047

Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796–815

Chen S, Jin W, Wang M., Zhang F, Zhou J, Jia Q, Wu Y, Liu F, Wu P (2003) Distribution and characterization of over 1000 T-DNA tags in rice genome. Plant J 36:105–113

Cheng C, Daigen M, Hirochika H (2006) Epigenetic regulation of the rice retrotransposon *Tos17*. Mol Genet Genomics 276:378–390

Cheong YH, Chang HS, Gupta R, Wang X, Zhu T, Luan S (2002) Transcriptional profiling reveals novel interactions between wounding, pathogen, abiotic stress, and hormonal responses in Arabidopsis. Plant Physiol 129:661–677

Christensen AH, Sharrock RA, Quail PH (1992) Maize *polyubiquitin* genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. Plant Mol Biol 18:675–689

Clarke L, Carbon J (1976) A colony bank containing synthetic Col E1 hybrid plasmids representative of the entire *E. coli* genome. Cell 9:91–99

Cornejo MJ, Luth D, Blankenship KM, Anderson OD, Blechl AE (1993) Activity of a maize ubiquitin promoter in transgenic rice. Plant Mol Biol 23:567–581

Dekeyser R, Claes B, Marichal M, Van Montagu M, Caplan A (1989) Evaluation of selectable markers for rice transformation. Plant Physiol 90:217-223

Dürrenberger F, Crameri A, Hohn B, Koukolíková-Nicola Z (1989) Covalently bound VirD2 protein of *Agrobacterium tumefaciens* protects the T-DNA from exonucleolytic degradation. Proc Natl Acad Sci USA 86:9154–9158

Ebitani T, Takeuchi Y, Nonoue Y, Yamamoto T, Takeuchi K, Yano M (2005) Construction and evaluation of chromosome segment substitution lines carrying overlapping chromosome segments of *indica* rice cultivar 'Kasalath' in a genetic background of *japonica* elite cultivar 'Koshihikari'. Breed Sci 55:65–73

Eulgem T (2005) Regulation of the *Arabidopsis* defense transcriptome. Trends Plant Sci 10:71–78

Fujita M, Mizukado S, Fujita Y, Ichikawa T, Nakazawa M, Seki M, Matsui M, Yamaguchi-Shinozaki K, Shinozaki K (2007) Identification of stress-tolerance-related transcription-factor genes via mini-scale full-length cDNA overexpressor (FOX) gene hunting system. Biochem Biophys Res Commun, in press

Gelvin SB (2000) *Agrobacterium* and plant genes involved in T-DNA transfer and integration. Annu Rev Plant Physiol Plant Mol Biol 51:223–256

Hiratsu K, Matsui K, Koyama T, Ohme-Takagi M (2003) Dominant repression of target genes by chimeric repressors that include the EAR motif, a repression domain, in *Arabidopsis*. Plant J 34:733–739

Hirochika H, Guiderdoni E, An G, Hsing YI, Eun MY, Han C, Upadhyaya N, Ramachandran S, Zhang Q, Pereira A, Sundaresan V, Leung H (2004) Rice mutant resources for gene discovery. Plant Mol Biol 54:325–334

Hood EE, Gelvin SB, Melchers LS, Hoekema A (1993) New *Agrobacterium* helper plasmids for gene transfer to plants. Transgenic Res 2:208–218

Hsing YI, Chern CG, Fan MJ, Lu PC, Chen KT, Lo SF, Sun PK, Ho SL, Lee KW, Wang YC, Huang WL, Ko SS, Chen S, Chen JL, Chung DI, Lin YC, Hour AL, Wang YW, Chang YC, Tsai MW, Lin YS, Chen YC, Yen HM, Li CP, Wey CK, Tseng CS, Lai MH., Huang SC, Chen LJ, Yu SM (2007) A rice gene activation/ knockout mutant resource for high throughput functional genomics. Plant Mol Biol 63:351–364

Ichikawa T, Nakazawa M, Kawashima M, Iizumi H, Kuroda H, Kondou Y, Tsuhara Y, Suzuki K, Ishikawa A, Seki M, Fujita M, Motohashi R, Nagata N, Takagi T, Shinozaki K, Matsui M (2006) The FOX hunting system: an alternative gain-of-function gene hunting technique. Plant J 48:974–985

International Rice Genome Sequencing Project (IRGSP) (2005) The map-based sequence of the rice genome. Nature 436:793–800

Jain M, Nijhawan A, Tyagi AK, Khurana JP (2006) Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR. Biochem Biophys Res Commun 345:646–651

Jasper F, Koncz C, Schell J, Steinbiss HH (1994) *Agrobacterium* T-strand production *in vitro*: sequence-specific cleavage and 5′ protection of single-stranded DNA templates by purified VirD2 protein. Proc Natl Acad Sci USA 91:694–698

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

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, An G (2006) Generation of a flanking sequence-tag database for activation-tagging lines in japonica rice. Plant J 45:123–132

Kachroo A, Lapchyk L, Fukushige H, Hildebrand D, Klessig D, Kachroo P (2003) Plastidial fatty acid signaling modulates salicylic acid- and jasmonic acid-mediated defense pathways in the Arabidopsis *ssi2* mutant. Plant Cell 15:2952–2965

Kachroo A, Venugopal SC, Lapchyk L, Falcone D, Hildebrand D, Kachroo P (2004) Oleic acid levels regulated by glycerolipid metabolism modulate defense gene expression in *Arabidopsis*. Proc Natl Acad Sci USA 101:5152–5157

Kaeppler SM, Kaeppler HF, Rhee Y (2000) Epigenetic aspects of somaclonal variation in plants. Plant Mol Biol 43:179–188

Karlowski WM, Schoof H, Janakiraman V, Stuempflen V, Mayer KF (2003) MOsDB: an integrated information resource for rice genomics. Nucleic Acids Res 31:190–192

Kolesnik T, Szeverenyi I, Bachmann D, Kumar CS, Jiang S, Ramamoorthy R, Cai M, Ma ZG, Sundaresan V, Ramachandran S (2004) Establishing an efficient Ac/Ds tagging system in rice: large-scale analysis of Ds flanking sequences. Plant J 37: 301–314

Komatsu S (2005) Rice proteome database: a step toward functional analysis of the rice genome. Plant Mol Biol 59:179–190

Khush GS (2005) What it will take to feed 5.0 billion rice consumers in 2030. Plant Mol Biol 59:1–6

Kurata N, Yamazaki Y (2006) Oryzabase. An integrated biological and genome information database for rice. Plant Physiol 140: 12–17

Lin SY, Sasaki T, Yano M (1998) Mapping quantitative trait loci controlling seed dormancy and heading date using backcross inbred lines in rice, *Oryza sativa L.* Theor Appl Genet 96: 997–1003

McElroy D, Zhang W, Cao J, Wu R (1990) Isolation of an efficient actin promoter for use in rice transformation. Plant Cell 2:163–171

Miki D, Shimamoto K (2004) Simple RNAi vectors for stable and transient suppression of gene function in rice. Plant Cell Physiol 45:490–495

Miyao A, Tanaka K, Murata K, Sawaki H, Takeda S, Abe K, Shinozuka, Onosato K, Hirochika H (2003) Target site specificity of the *Tos17* retrotransposon shows a preference for insertion within genes and against insertion in retrotransposon-rich regions of the genome. Plant Cell 15: 1771–1780

Miyao A, Iwasaki Y, Kitano H, Itoh J, Maekawa M, Murata K, Yatou O, Nagato Y, Hirochika H (2007) A large-scale collection of phenotypic data describing an insertional mutant population to facilitate functional analysis of rice genes. Plant Mol Biol 63:625–635

Mori M, Tomita C, Sugimoto K, Hasegawa M, Hayashi N, Dubouzet JG, Ochiai H, Sekimoto H, Hirochika H, Kikuchi S (2007) Isolation and molecular characterization of a *Spotted leaf 18* mutant by modified activation-tagging in rice. Plant Mol Biol 63:847–860

Nakazawa M, Ichikawa T, Ishikawa A, Kobayashi H, Tsuhara Y, Kawashima M, Suzuki K, Muto S, Matsui M (2003) Activation tagging, a novel tool to dissect the functions of a gene family. Plant J 34:741–750

Noro Y, Takano-Shimizu T, Syono K, Kishima Y, Sano Y (2007) Genetic variations in rice in vitro cultures at the *EPSPs-RPS20* region. Theor Appl Genet 114:705–711

Ohyanagi H, Tanaka T, Sakai H, Shigemoto Y, Yamaguchi K, Habara T, Fujii A, Antonio BA, Nagamura Y, Imanishi T, Ikeo K, Itoh T, Gojobori T, Sasaki T (2006) The rice annotation project database (RAP-DB): hub for *Oryza sativa* ssp. *japonica* genome information. Nucleic Acids Res 34:D741-D744

Ouyang S, Zhu W, Hamilton J, Lin H, Campbell M, Childs K, Thibaud-Nissen F, Malek RL, Lee Y, Zheng L, Orvis J, Haas B, Wortman J, Buell CR (2007) The TIGR Rice Genome Annotation Resource: improvements and new features. Nucleic Acids Res 35:D883-D887

Park SH, Jun NS, Kim CM, Oh TY, Huang J, Xuan YH, Park SJ, Je BI, Piao HL, Park SH, Cha YS, Ahn BO, Ji HS, Lee MC, Suh SC, Nam MH, Eun MY, Yi G, Yun DW, Han CD (2007) Analysis of gene-trap *Ds* rice populations in Korea. Plant Mol Biol, doi:10.1007/s11103-007-9192-5

Rabbani MA, Maruyama K, Abe H, Khan MA, Katsura K, Ito Y, Yoshiwara K, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2003) Monitoring expression profiles of rice genes under cold, drought, and high-salinity stresses and abscisic acid application using cDNA microarray and RNA gel-blot analyses. Plant Physiol 133:1755–1767

Rice Annotation Project (2007) Curated genome annotation of *Oryza sativa* ssp. *japonica* and comparative genome analysis with *Arabidopsis thaliana.* Genome Res 17:175–183

Rice Full-Length cDNA Consortium (2003) Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice. Science 301:376–379

Ryu CH, You JH, Kang HG, Hur J, Kim YH, Han MJ, An K, Chung BC, Lee CH, An G (2004) Generation of T-DNA tagging lines using a bi-directional gene trap vector and the establishment of an insertion-site database. Plant Mol Biol 54:489–502

Sakamoto T, Kobayashi M, Itoh H, Tagiri A, Kayano T., Tanaka H, Iwahori S, Matsuoka M (2001) Expression of a gibberellin 2-oxidase gene around the shoot apex is related to phase transition in rice. Plant Physiol 125:1508–1516

Sallaud C, Gay C, Larmande P, Bès M, Piffanelli P, Piégu B, Droc G, Regad F, Bourgeois E, Meynard D, Périn C, Sabau X, Ghesquière A, Glaszmann JC, Delseny M, Guiderdoni E (2004) High throughput T-DNA insertion mutagenesis in rice: a first step towards *in silico* reverse genetics. Plant J 39:450–464

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, New York

Sasaki T, Matsumoto T, Antonio BA, Nagamura Y (2005) From mapping to sequencing, post-sequencing and beyond. Plant Cell Physiol 46:3–13

Seki M, Narusaka M, Kamiya A, Ishida J, Satou M, Sakurai T, Nakajima M, Enju A, Akiyama K, Oono Y, Muramatsu M, Hayashizaki Y, Kawai J, Carninci P, Itoh M, Ishii Y, Arakawa T, Shibata K, Shinagawa A, Shinozaki K (2002) Functional annotation of a full-length Arabidopsis cDNA collection. Science 296:141–145

Toki S, Hara N, Ono K, Onodera H, Tagiri A, Oka S, Tanaka H (2006) Early infection of scutellum tissue with *Agrobacterium* allows high-speed transformation of rice. Plant J 47:969–976

Upadhyaya NM, Zhou X-R, Ramm K, Zhu Q-H, Wu L, Eamens AL, Sivakumar R, Kato T, Yun D-W, Kumar S, Narayanan KK, Peacock WJ, Dennis ES (2002) An *iAc/Ds* gene and enhancer trapping system for insertional mutagenesis in rice. Funct Plant Biol 29:547–559

van Enckevort LJ, Droc G, Piffanelli P, Greco R, Gagneur C, Weber C, González VM, Cabot P, Fornara F, Berri S, Miro B, Lan P, Rafel M, Capell T, Puigdomènech P, Ouwerkerk PB, Meijer AH, Pe' E, Colombo L, Christou P, Guiderdoni E, Pereira A (2005) EU-OSTID: a collection of transposon insertional mutants for functional genomics in rice. Plant Mol Biol 59:99–110

Wu JL, Wu C, Lei C, Baraoidan M, Bordeos A, Madamba MR, Ramos-Pamplona M, Mauleon R, Portugal A, Ulat VJ, Bruskiewich R, Wang G, Leach J, Khush G, Leung H (2005) Chemical- and irradiation-induced mutants of indica rice IR64 for forward and reverse genetics. Plant Mol Biol 59:85–97

Yamamoto Y, Kamiya N, Morinaka Y, Matsuoka M, Sazuka T (2007) Auxin biosynthesis by the *YUCCA* genes in rice. Plant Physiol 143:1362–1371

Zhang J, Li C, Wu C, XiongL, Chen G, Zhang Q, Wang S (2006) RMD: a rice mutant database for functional analysis of the rice genome. Nucleic Acids Res 34:D745–D748

Zhao W, Wang J, He X, Huang X, Jiao Y, Dai M, Wei S, Fu J, Chen Y, Ren X, Zhang Y, Ni P, Zhang J, Li S, Wang J, Wong GK, Zhao H, Yu J, Yang H, Wang J (2004) BGI-RIS: an integrated information resource and comparative analysis workbench for rice genomics. Nucleic Acids Res 32:D377–D382

Zheng Z, Xia Q, Dauk M, Shen W, Selvaraj G, Zou J (2003) Arabidopsis *AtGPAT1*, a member of the membrane-bound glycerol-3-phosphate acyltransferase gene family, is essential for tapetum differentiation and male fertility. Plant Cell 15:1872–1887