J. Plant Biotechnology *Vol. 1. pp.46~55*

Journal of Plant Biotechnology

The EST Analysis and Transgene Expression System in Rice

Jukon Kim, Baek Hie Nahm*

Dept. of Biological Science, Myongji University, Namdong san 38-2, Yongin, Korea 449-728

(author for correspondence)*

Keywores; Oryza sativa, MAR, Transformation, Transit peptide, chloroplast targeting

Abstract

The expressed sequence tags(ESTs) from immature seed of rice, *Oryza sativa* cv Milyang 23, were partially sequenced and analyzed by homology. As of 1998, the partial sequences of about 6,600 cDNA clones were analyzed from normal and normalized immature seed cDNA libraries. About 2,200 ESTs were putatively identified by BLASTX deduced amino acid sequence homology analysis. About 20% of them were putatively identified as storage proteins. Also the clones were highly homologous to genes involved particularly in starch biosynthesis, glycolysis, signal transduction and defenses. Compared to 35% of redundancy in the ESTs of normal cDNA library, that from the substracted library was 15%. The Korea Rice Genome Network is maintained to provide the updated information of sequences, their homologies and sequence alignments of ESTs.

For the stable expression of transgene in rice, diverse vectors were developed for overexpression, targeting and gene dosage effect with transit peptides (Tp) and matrix attachment region (MAR) sequence from chicken lysozyme locus. The rice calli were transformed via Agrobacterium tumefaciens LBA4404(pSB1) with the triparental mating technique and selected by herbicide resistance. The green fluorescent protein(GFP) gene in expression vector under the control of rbcS promoter-Tp was overexpressed upto 10 % of the total soluble protein. In addition, the Tp-sGFP fusion protein was properly processed during translocation into chloroplast. The expression of sGFP in the presence of MAR sequences was analyzed with Northern and immunoblot analysis. All the lines in which sGFP transgene with MAR sequence, showed position independent and copy number-dependent expression, while the lines without MAR showed the varied level of expression with the integration site. Thus the MAR sequence significantly reduced the variation in transgene expression between independent transformants.

Introduction

Rice is the major crop in the world, which has been used as a staple in many countries, especially in Asia. Because of the small and compact size of the genome among the crops, rice has become as the model organism among the crops for genomic studies with the intensive mapping of rice genome and chromosomal homeology between major crops such as wheat and maize. As a major crop, the genetic information related to important agronomic trait, grain quality as well as productivity as a sink organ are the major target for improvement of grain with molecular biological approach. The importance of the EST clones and their sequence information cannot be overemphasized more than ever with the genome information, transformation and DNA chip technology.

An improvement in transformation technology has allowed novel approaches for studying gene expression and opened new avenues for the genetic modification of crop plants. During the last decade, more than 3600 field trials of transgenic plants have been performed worldwide [17]. Rice is the first cereal crop from which fertile transgenic

Table 1. The Overall Characters of ESTs analyzed

LIBRARY	Standard	Normalized
Number of clones	3,513	3315
Identified	1,226 clones(35%)	990 (30%)
Unidentified	2,287 clones (65%)	2,325 (70%)
Redundant Clones	35 %	16 %

plants were obtained [32, 33, 36]. Since then, many fertile transgenic rice lines have been produced by protoplast-mediated DNA transformation [7, 31], by microprojectile bombardment [3, 4, 5, 10, 35], or by Agrobacterium-mediated DNA transfer [15, 16, 22]. Thus, generation of transgenic plants is now becoming increasingly routine for many rice varieties. Despite many efforts have been made using transgenesis to improve important agronomic traits of rice [for review, see 5, 16], there are only few cases where transgenic plants were showing acceptable agronomic performance [5].

For efficient and stable gene expression in transgenic plants with desired phenotypes, there are several factors that need to be considered before construction of expression cassettes for transformation. These include promoters, use of introns, codon usage of genes, and localization of gene products to a specific intracellular compartment. In particular, for the genes that are expressed in cytoplasm at low level or those that are deleterious to cells when expressed in cytoplasm, localization of gene products to other intracellular compartments such as chloroplast is desirable. There are important genetic factors that affect transgene expression and that have to be carefully integrated as well before transformation. This includes transgene silencing and influence of chromosomal environment on transgene expression known as position effects. One strategy to escape from transgene silencing in rice is to use genes from genetically distant organisms such as dicotyledonous plants or microorganisms and to select transgenic lines with single copy of transgene after transformation. This is because silencing is mostly caused by sequence homology between the genes being introduced and the endogenous counterparts or between the newly introduced transgenes [26]. Recently, MAR (matrix attachment region) sequences were shown to suppress the position effects in transgenic tobacco, reducing variability in the transgene expression [1, 27, 28, 29]. Also, Iglesias et al., [19] suggested that MAR sequences play a role in decreasing chance of transgene silencing. Thus, use of MAR sequences in transgene expression can provide an alternative solution for position effects and silencing.

In this research, to provide the gene information related to grain quality and productivity, the partial sequences of the expressed sequence tags (ESTs) from immature seed of rice were analyzed. Also in order to develop systems for the secure and stable transgene expression, a chloroplast targeting sequence from the rice small subunit gene of RUBISCO and the chicken lysozyme MAR sequence were tested in transgenic rice plants.

Analysis of Rice Immature Seed ESTs

Feature of EST sequences

The EST clones from standard and normalized immature seed cDNA were randomly selected for partial sequence analysis. To identify the EST clones, the homology of the partial sequences were analyzed by BLASTX software [2], against nonredundant protein database at their deduced amino acid sequence level. As shown in Table1, about 35 and 30% of the EST clones from standard and normalized libraries, respectively, could be putatively identified. Also the number of redundant ESTs in standard and normalized libraries were found to be 35 and 16% of the total EST clones analyzed. The high number of identified clones in standard library were resulted by the highly redundant clones identified as storage proteins. As shown in Table 2, the ESTs putatively identified as storage proteins genes especially prolamines, glutelins and allergens are analyzed to be the most abundant. The results showed the genes for the biosynthesis of starch are also the highly expressed in addition to genes required in translation and basal metabolism. The comparative analysis of expression profiles of the genes related to the starch biosynthesis in diverse kernel mutants will provide the phenotypic trait of grain quality and productivity.

According to the EST clones, the genes involved in

Table 2. The Categories of Putatively Identified ESTs

Functional Category	Number of Genes	Number of ESTs
Storage Proteins	46	323
Basal Metabolism	168	273
Photosynthesis	98	253
Defense system	70	132
Structural Protein	74	102
Signal Transduction	77	100
Translation System	62	76
Starch Biosynthesis	35	68
Redox Proteins	17	23
Transcription factors	16	20
Others	119	141
Total	782	1,511

*The number of EST clones were estimated according to the blastx score higher than 100 and more than 40% of identities of amino acids sequences which showed the highly significant identification.

defense mechanisms such as heat shock, superoxide dismutase, chitinase, stress-induced genes as well as genes related to signal transduction such as G proteins and diverse protein kinase are one of the major genes actively expressed in the immature seed. The functional analysis of these genes may provide the clue for the explanation of the high resistance traits of seed to diverse pathogens and stresses. Also as structural proteins, various histone proteins, actin, tublin and diverse ubiquitine genes are found as one of the major genes expressed in immature seed.

The Distribution of Sequence and Analysis Information

The sequence information submitted to GENBANK, NCBI, NIH, USA which can be accessed by internet network. In addition to sequence information, homology results at DNA and its deduced amino acid sequence level obtained by blastn and blastx can be searched and retrieved by keyword index at the website of Korea Rice Genome DB Network (http://bioserver.myongji.ac.kr) (Figure 1). The information related to EST clones and their partial sequence are continuously updated and posted at the Website. The EST sequence information is also included in RiceGenes, rice version of ACEDB organism oriented genetic information database software, which can be run with diverse platform under various operation system, such as Windows95, MacOS, Linux and Unix. The binary executable programs are available at anony- mous FTP sites of USDA (ftp://probe.nalusda.gov/pub/ricegenes) and Korea Rice Genome DB Network at Myongji University (ftp://bioftp.myongi.ac.kr/pub).

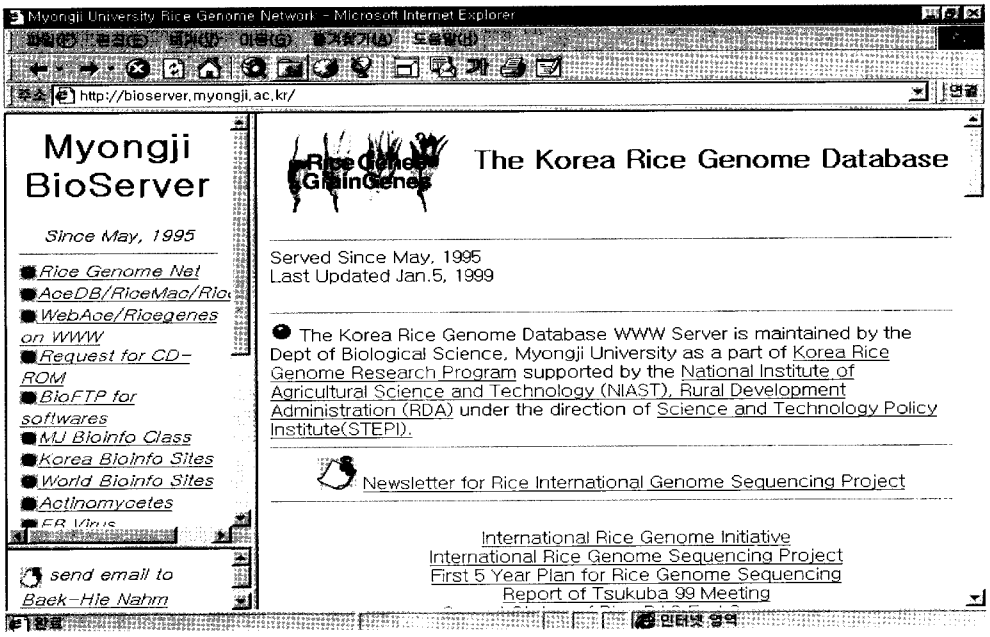
Also the CD-ROM containing RiceMac and RiceWin, Windows95 and MacOS versions of RiceGenes DB software which can be run with PC and Macintosh is available upon request with free of charge (Figure 2). Anyone who want to receive regularly can enlist on the mailing list of RiceGenes CD-ROM at the Website or by email.

Transgene Expression Systems in Rice

Subcellular targeting of protein to plastids

Ribulose bisphosphate carboxylase/oxygenase (RUBISCO) is the primary enzyme responsible for carbon dioxide fixation in the chloroplast. RUBISCO is the most abundant protein found in plant leaves, representing up to 50% of the total soluble protein [23]. The small subunit (rbcS) gene of RUBISCO is nuclear-encoded and its expression is regulated by both light and an endogenous leaf-specific pattern of activation [for review, see 9, 11]. The rbcS mRNA is translated in the cytoplasm to produce a large precursor protein which is subsequently processed to its mature size by cleavage of transit peptide during transport into the chloroplast [for review, see 6, 14]. Thus, the great abundance and chloroplast localization of protein, the high levels of regulated expression of rbcS have not only facilitated work to understand various aspect of gene expression in plants, but also provided an attractive system for expression of genes at high levels and for targeting of different proteins into the chloroplast. There are some examples of rbcS transit peptides that increased expression levels of fused genes as well as targeting of gene products [8, 12, 13]. In addition, chloroplast targeting of gene products can alleviate the toxic effects of the products on cells. Targeting of Cry1A(c) [34] and the polyhydroxybutyrate synthesizing enzymes [30] to chloroplasts of transgenic tobacco and Arabidopsis, respectively, gave rise to normal plants, which otherwise would have been stunted.

(A)



(B)

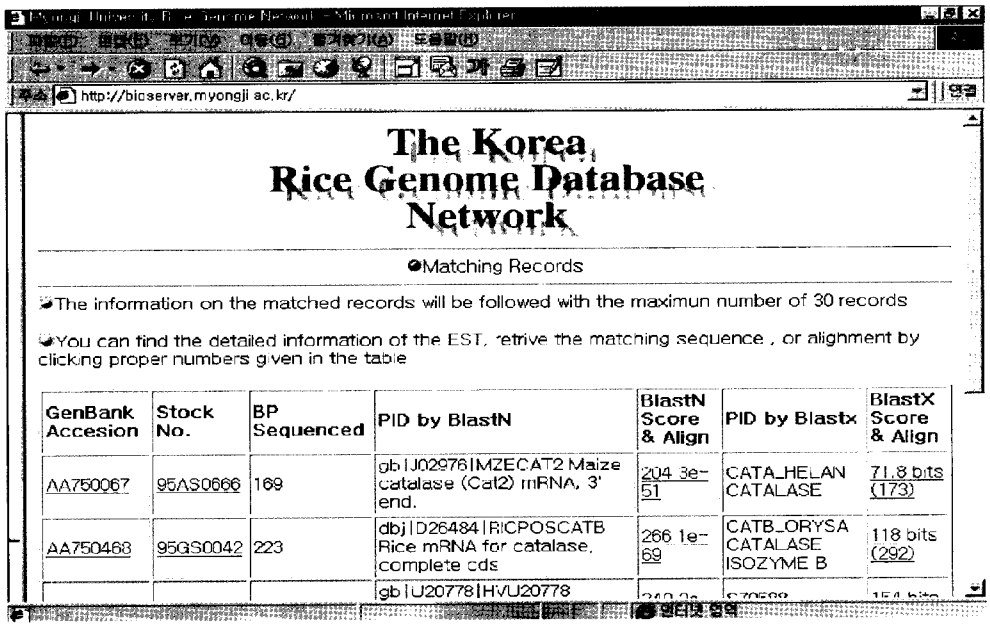


Figure 1. The Homepage of Korea Rice Genome DB Network at Myongji University (A) and its EST information search and retrieval of updated homology analysis results.(B)

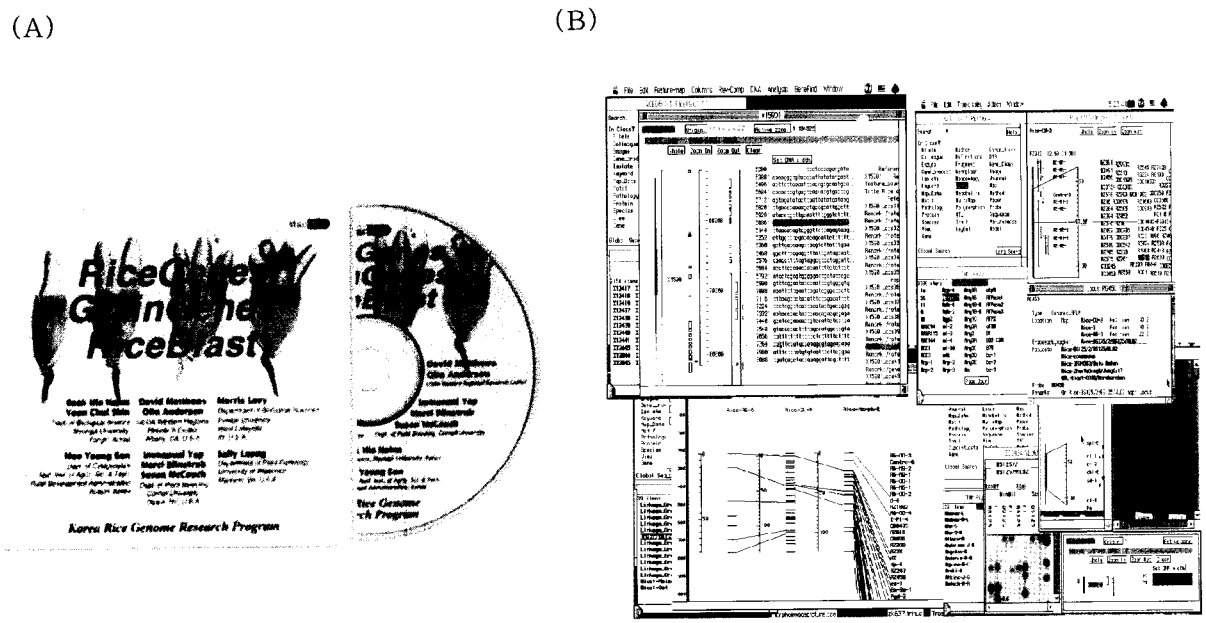


Figure 2. The CD-ROM containing RiceMac and RiceWin (A) and Multiple windows displaying contents of RiceGenes (B)

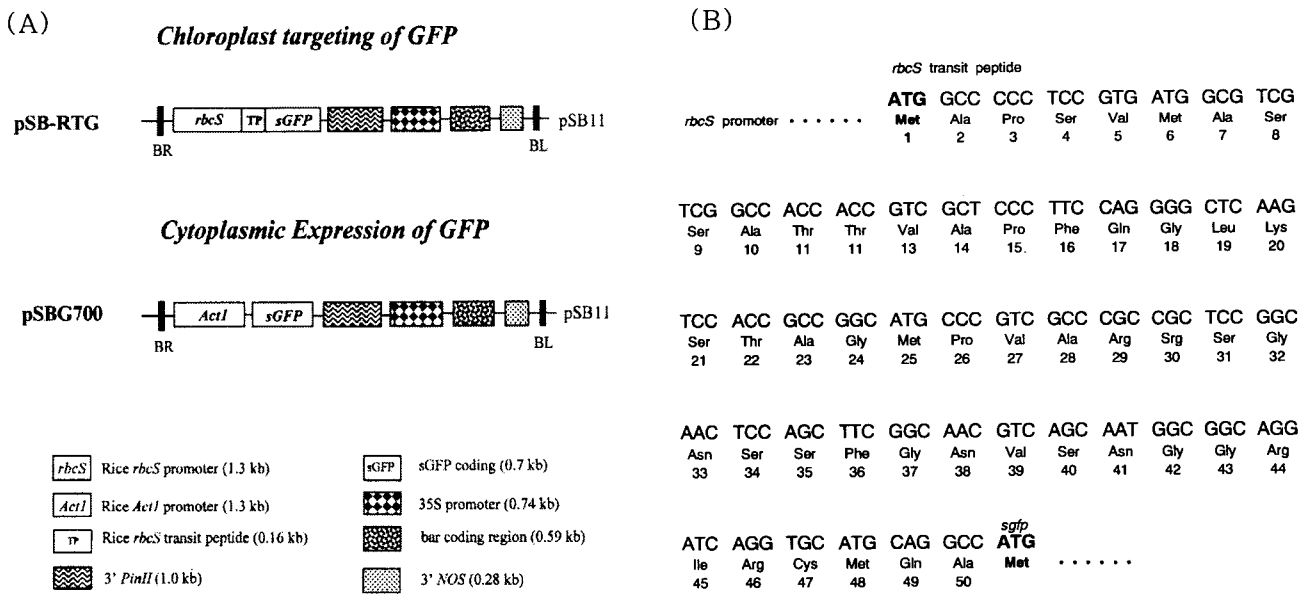


Figure 3. Expression vectors used for rice transformation. A. pSB-RTG (rbcS-Tp-sGFP) consists of the rice rbcS promoter and its transit peptide (Tp) sequence translationally fused to sGFP coding region, and 3' region of the potato proteinase inhibitor II gene (pinII) plus the bar gene expression cassette that contains 35S promoter/bar coding region/3' region of the nopaline synthase gene (Nos). pSBG700 (Act1-sGFP) consists of the rice Act1 5' region including the promoter, first non-coding exon, and first intron linked to the sGFP coding region of the pSB-RTG. BR, right border; BL, left border. B. Nucleotides and amino acid sequences of the nuclear-encoded rice rbcS transit peptide, which was translationally fused with sGFP in the plasmid pSB-RTG in A. The three amino acids added during cloning procedure by the linker sequence are underlined. The vertical arrow indicates a predicted cleavage site during the import of the protein into chloroplast.

As shown in Fig. 3A, we have designed the expression vector pSB-RTG (rbcS-Tp-sGFP) for chloroplast targeting, which consists of rice rbcS promoter and transit peptide(Tp) sequence translationally fused to the green fluorescent protein (GFP) gene. In transgenic rice plants, addition of the transit peptide to the rbcS promoter not only targeted GFP to chloroplast, also elevated expression levels of the protein 20 fold, yielding up to 10% of total soluble protein (Fig. 4) [18]. Expression of sGFP was determined by immuno-blot analysis. The protein extracts were prepared from leaf tissues or calli of six independent fertile transformants: 4 lines of rbcS-Tp-sGFP-transformed plants and 2 lines of pSB700(Act1-sGFP)-transformed plants (Fig. 4). GFP was clearly detectable in all the transgenic plants analyzed with different levels of expression (Fig. 4a). In Act1-sGFP-transformed lines, sGFP was expressed both in calli and in mature leaves at the similar levels (Fig. 4b). Surprisingly, the expression levels in the rbcS-Tp-sGFP-transformed lines were about 20 fold higher than those in the Act1-sGFP-transformed lines in mature leaves; this was measured by comparing band intensities of a series of dilutions of the protein extract (Fig. 4c, lanes 5-8). By comparing band intensities with the known amount of the purified protein (Fig. 4c, lane 1), we estimated that the GFP in Act1-sGFP-transformed lines amounted to be 0.5% of the total soluble proteins. In rbcS-Tp-sGFP-transformed lines; therefore, the levels increased up to a maximum of 10% of total soluble proteins in mature tissues. The results paralleled with those obtained from a comparison of band intensities in a protein gel stained with Coomassie blue (Fig. 4d). The difference in expression levels between the two different transgenic lines was large and unexpected. This is because the activity of the same rbcS promoter without the Tp sequence was about 5 fold higher than that of the Act1 promoter (data not shown) The two promoters were fused to a maize

As described earlier, Kyozuka et al. [24] predicted the cleavage site of the rice RbcS protein to be between 47Cys and 48Met by comparison with the N-terminal amino acid sequence of the mature RbcS protein of tobacco [25]. In this study, as shown in Figure 1B, correct removal of a putative N-terminal transit peptide from the Tp-GFP fusion protein during transport to chloroplast would yield GFP with three additional amino acids, Met-Gln-Ala, at its N-terminal. In immuno-blots with the anti-GFP antibody, the Tp-GFP fusion protein from chloro-

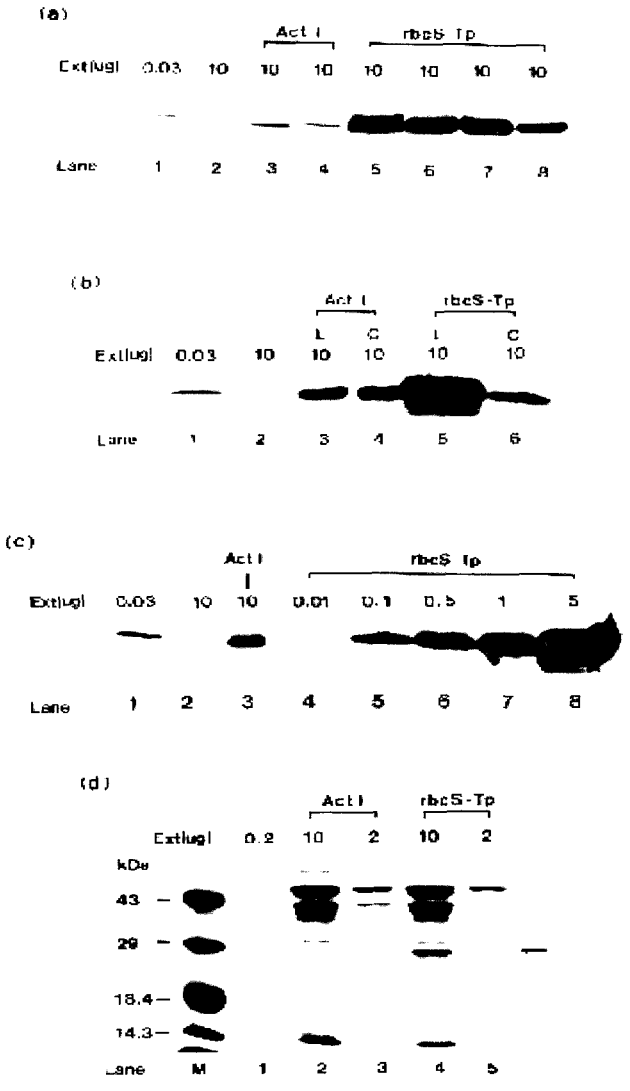


Figure 4. Expression analysis of sGFP in transgenic rice plants. Protein extracts were separated on SDS-PAGE, transferred to PVDF membrane, and GFP was detected with antibodies against GFP in A, B and C. The amounts of protein extracts loaded per lane are indicated by Ext [mg]. Proteins extracted from the Act1-sGFP-transformed (Act1) and rbcS-Tp-sGFP-transformed (rbcS-Tp) rice plants are marked in A, B, C and D. Lane 1 and 2 in A, B and C are the purified GFP and protein extracts from an untransformed plant, respectively. A. sGFP expression in different lines of the Act1-sGFP-transformed (lanes 3 and 4) and rbcS-Tp-sGFP-transformed (lanes 5-8) rice. B. sGFP expression in different tissues (L, leaves; Ca, calli) of the two transgenic lines (Act1 and rbcS-Tp) shown in lane 3 and 5 of A, respectively. C. Comparison of expression levels between the two transgenic lines (Act1 and rbcS-Tp) shown in lane 3 and 5 of A, respectively. A series of dilution of proteins from the transgenic line shown in lane 5 of A were loaded in lanes 4-8. D. SDS-PAGE gel stained with Coomassie blue of total proteins extracted from the two transgenic lines (Act1 and rbcS-Tp) shown in lane 3 and 5 of A, respectively. Lane 1, purified GFP (0.2 mg); lane M is molecular weight marker. An arrow with arrowhead at one end indicates the GFP bands.

plast-targeted leaves migrated similarly to the GFP without the transit peptide (Fig. 4a and 4b), suggesting correct processing and removal of the transit peptide from the fusion protein. Interestingly, the Tp-GFP fusion protein from calli also appeared to be properly processed (Fig. 4b, lane 6), suggesting that the Tp sequence was able to localize the fusion protein to proplastids in addition to chloroplasts. We determined the N-terminal sequences of the GFP extracted from leaf tissues and calli of the rbcS-Tp-sGFP-transformed line, obtaining the sequences, Met-Gln-Ala-Met-Val- from both samples. These sequences are exactly equivalent to those which were expected (Fig. 3b), clearly validating the correct processing of the transit peptide from the Tp-GFP fusion protein during its transport into chloroplasts and proplastids.

Taken together, our results demonstrated that the transit peptide sequence of the rice rbcS was sufficient for the targeting of the fusion protein specifically to chloroplast and proplastid, and also significantly increased expression levels of the latter. It is possible that our rice rbcS transit peptide sequence may act as a transcriptional enhancer, increasing the expression levels of the fused sGFP gene. Alternatively, continuous removal of GFP from cytoplasm may prevent the protein from reaching a toxic threshold level in the compartment causing an efficient translation. The latter is supported by our observation that, in the leaf mesophyll cells, intracellular space shared by cytoplasm and nuclei, where the untargeted GFP is present, is much smaller than that shared by chloroplasts, where the targeted GFP is present. Consequently, the amount of GFP per cell producing the targeted GFP could be larger than that per cell producing the untargeted GFP. For comparison of expression levels of pSB-RTG and pSB700 in Fig. 4, we chose those transgenic lines that contained a single copy of the corresponding transgene. Copy numbers were determined by genomic Southern blot experiments (data not shown).

Our expression system will be useful in increasing the expression of other genes which are agronomically important; in particular, genes that are expressing in cytoplasm at a low level or genes that are deleterious to cells when expressed in cytoplasm. For example, expression of Bacillus thuringiensis toxin genes, that had reached the levels of 0.02-0.1% of the total soluble protein so far, can be significantly increased. This system can also be applied for production of foreign proteins or peptides that are medically important, for example, vaccines or monoclonal antibodies. Also, GFP fluorescent chloroplasts generated in this work are readily observable as distinct from other organelles in living cells or tissues, providing a plant material for morphological studies of the important organelle at different stages of development.

Position-Independent Transgene Expression by MAR

Genetic transformation of plants generally results in a large and random variation in the expression of the newly introduced transgene between individual transformants. This variability is attributed to different integration sites of the transgene, reflecting the influence of the surrounding chromatin known as position effects. MAR (Matrix Attachment Region) sequences are repeatedly shown to insulate transgenes from surrounding chromatin, reducing variability in the transgene expression in transgenic tobacco [1, 27, 28, 29]. Addition of MAR sequences to expression vectors resulted in an increase in expression levels of the linked reporters in transgenic plants [1]. Interestingly, in a recent study of genomic context around transgenic loci [19], the authors found that the stably expressed loci were flanked by a MAR-like sequence whereas the unstably expressed loci were not, suggesting that MAR sequences play a role in decreasing risks of transgene silencing. Thus, use of MAR sequences in construction of expression vectors can minimize position effects, increase expression levels and possibly stabilize transgene expression.

To investigate whether the MAR sequence is working in transgenic rice plants, we have inserted the chicken lysozyme MAR sequences into borders of the T-DNA of the plasmid pSBG700 containing Act1 promoter linked to GFP coding sequence, generating the plasmid pSBG-MAR as shown in Fig. 5. A number of fertile transgenic rice plants for each construct were produced by the Agrobacterium-mediated method. Southern blot analysis of 40 primary transformants resulted in distinct band patterns, indicating that all the transformants had been generated by independent integration events (Fig. 6). Northern and immunoblot analysis demonstrated that, in the absence of MAR, expression of sGFP transgenes was observed in all of the lines, but expression levels markedly varied with the integration sites on chromosomes (Fig. 7A). In contrast, all the lines in which sGFP transgenes were flanked by the MAR sequence at both ends showed position-independent expression, resulting in similar levels in ex-

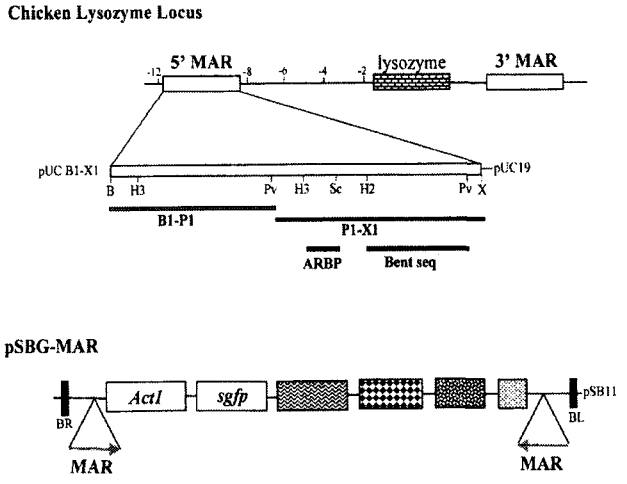


Figure 5. Schematic representation of the chicken lysozyme MAR sequence and pSBG-MAR. The 5'MAR sequence present around the chicken lysozyme gene with important restriction sites is shown in upper panel. B1-P1 and P1-X1 DNA fragments were isolated by PCR and inserted into the plasmid pSBG700 (Fig. 3) as shown in lower panel, generating the plasmid pSBG-MAR. B1-P1, the 5' half of 5' MAR sequence between BamHI and PvuII sites; P1-X1, the 3' half of 5' MAR between PvuII and XbaI sites; APBP, binding region for the nuclear matrix protein ARBP; Bent seq, location of intrinsically bent sequence; lysozyme, coding sequence of chicken lysozyme gene. Other abbreviations are the same as those in Fig. 3.

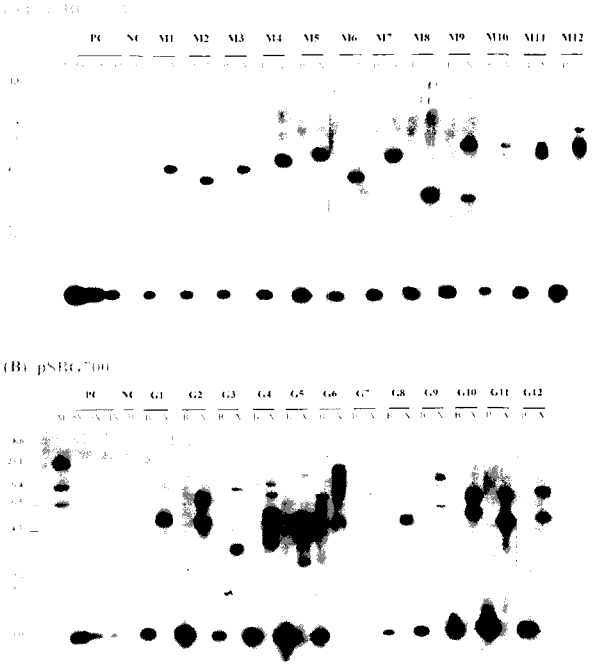


Figure 6. Genomic DNA blot analysis of pSBG-MAR (A) or pSBG 700 (B)-transformed rice plants and estimation of the transgene copy numbers. 1X, 3X and 5X in PC represent 1, 3 and 5 genome equivalents of pSBG-MAR in A and pSBG700 DNA in B, relative to 5 ug of rice genomic DNA, respectively. The DNA molecular size markers (M) are indicated on the left-hand side.

A. Genomic DNAs isolated from the leaf tissues of pSBG-MAR-transformed rice plants (M1-M12) and from an untransformed control plant (NC) were digested with XbaI (X), BamHI (B) and hybridized with a 32P-labeled 0.7-kb DNA fragment containing the sGFP coding region. PC contains BamHI-digested pSBG-MAR.

B.Genomic DNAs isolated from the leaf tissues of pSBG700-transformed rice plants (G1-G12) and from an untransformed control plant (NC) were digested with XbaI (X), BamHI (B) and hybridized with a 32P-labeled 0.7-kb DNA fragment containing the sGFP coding region. PC contains BamHI-digested pSBG700.

pression among different lines (Fig. 7B). Thus, the MAR sequence significantly reduced the variation in transgene expression between independent transformants. Moreover, the sequence appears to confer copy number-dependence in gene expression (Fig. 7B).

Thus, variations in expression levels of the two different set of transformants were measured and compared, evaluating MAR activities in transgenic rice. Since the chicken MAR sequence turned out to be effective in rice chromosomes, this system should provide two significant advantages in transformation practices. One is to increase the frequency of transformation because transgenes that are flanked by MAR sequences will be expressed regardless of chromosomal regions they are integrated into. Transferred genes without MAR sequences are expected to be expressed only when they are landed on transcriptionally active regions of chromosomes. The genes integrated into transcriptionally inactive regions are not expressed, being excluded during selection in tissue culture. The other advantage is to reduce the number of transgenic plants that one should obtain in the initial stage of transformation. Generally, for a few fertile transgenic plants with desired levels of transgene expression, a large number of primary transformants have to be generated and analyzed. Because the MAR sequence can direct evenly high levels of expression of transgenes in different transgenic lines by suppressing position effect, one do not need to make so many transgenic lines. These advantages will synergistically help us to perform efficient transformations by reducing the amount of work for one transformation and by obtaining transgenic plants stably expressing genes of interest.

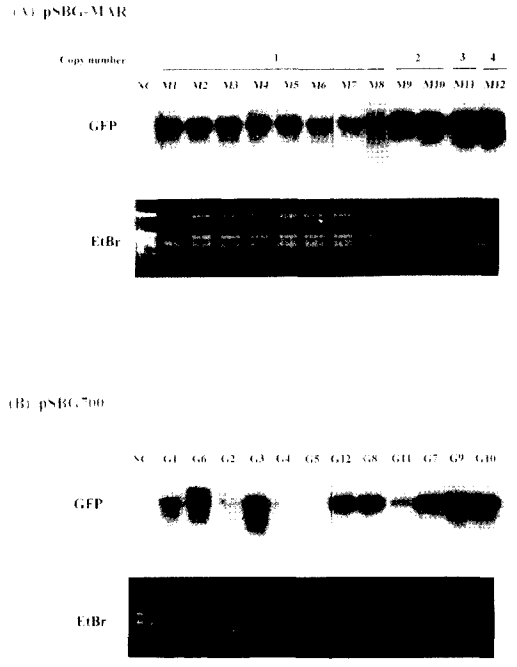


Figure 7. RNA gel blot analysis showing levels of GFP transcripts in pSBG-MAR (A) or pSBG700 (B)-transformed rice plants. EtBr-staining of total RNA was used as a control for equal RNA loading. The identification numbers of the transgenic plants are marked on the top corresponding to numbers of DNA gel blot analysis described in Fig 3.

A. Total RNAs extracted from the leaf tissues of pSBG-MAR-transformed rice plants (M1-M12) and from an untransformed control plant (NC) were hybridized with a 32P-labeled 0.7-kb DNA fragment containing the sGFP coding region.

B. Total RNAs extracted from the leaf tissues of pSBG700-transformed rice plants (G1-G12) and from an untransformed control plant (NC) were hybridized with a 32P-labeled 0.7-kb DNA fragment containing the sGFP coding region.

References

1. Allen, GC, Hall G, Michalowski S, Newman W, Spiker S, Weissinger AK, Thompson WF (1996) High-level transgene expression in plant cells: effects of a strong scaffold attachment region from tobacco. Plant Cell 8: 899-913.
2. Altschul SF, Thomas LM, Alejandro AS, Jinghui Z, Zheng Z, Webb M, David JL (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389-3402.
3. Cao J, Duan X, McElroy D, Wu R (1992) Regeneration of herbicide resistant transgenic rice plants following microprojectile-mediated transformation of suspension cells. Plant Cell Reports 11 586-591.
4. Christou P, Ford TL, Kofron M (1991) Production of transgenic rice (*Oryza sativa* L.) plants from agronomically important indica and japonica varieties via electric discharge particle acceleration of exogenous DNA into immature zygotic embryos. Bio/Technology 9 957-962.
5. Christou P (1997) Rice transformation: bombardment. Plant Mol Biol. 35: 197-203.
6. Cline K Henry R (1996) Import and routing of nucleus-encoded chloroplast proteins. Ann Rev Cell Dev Biol 12: 1-26.
7. Datta SK, Peterhans A, Datta K, Potrykus I (1990) Genetically engineered fertile indica-rice recovered from protoplasts. Bio/Technology 8: 736-740.
8. DeAlmeida ERP, Gossele V, Muller CG, Dock J, Reynaerts A, Botterman J, Krebbers E, Timko MP (1989) Transgenic expression of two marker genes under the control of an Arabidopsis rbcS promoter: sequences encoding the rubisco transit peptide increase expression levels. Mol Gen Genet 218: 78-86.
9. Dean C. Pichersky E, Dunsmuir P (1989) Structure, evolution, and regulation of RbcS genes in higher plants. Ann Rev Plant Physiol. Plant Mol Biol 40: 415-439.
10. Duan X, Li X, Xue Q, Abo-El-Saad M, Xu D, Wu R (1996) Transgenic rice plants harboring an introduced potato proteinase inhibitor II gene are insect resistant. Nature Biotechnology 14: 494-498.
11. Gilmartin PM, Sarokin L, Memelink J Chua N-H (1990) Molecular light switches for plant genes. Plant Cell 2: 369-378.
12. Herminghaus S, Schreier PH, McCarthy JEG, Landsmann J, Botterman J, Berlin J (1991) Expression of a bacterial lysine decarboxylase gene and transport of the protein into chloroplasts of transgenic tobacco. Plant Mol Biol 17: 475-486.
13. Herminghaus S, Tholl D, Rugenhagen C, Fecker LF, Leuschner C, Berlin J (1996) Improved metabolic action of a bacterial lysine decarboxylase gene in tobacco hairy root cultures by its fusion to a rbcS transit peptide coding sequence. Transgenic Res 5: 193-201.
14. Von Heijne G, Steppuhn J, Herrmann RG (1989) Domain structure of mitochondrial and chloroplast targeting peptides. Eur J Biochem 180: 535-545.
15. Hiei Y, Ohta S, Komari T, Kumashiro T (1994) Efficient transformation of rice (*Oryza sativa* L.) mediated by Agrobacterium and sequence analysis of the boundaries of the T-DNA. Plant J 6: 271-282.
16. Hiei Y, Komari T, Kubo T (1997) Transformation of rice mediated by *Agrobacterium tumefaciens*. Plant Mol Biol 35: 205-218.
17. James C, Krattiger AF (1996) Global review of the field testing and commercialization of transgenic plants, 1986 to 1995: the first decade of crop biotechnology. The Internation Service for the Acquisition of Agri-biotech Applications (ISAAA) Briefs No. 1, ISAAA, Ithaca, NY.
18. Jang I-C, Nahm BH, Kim J-K (1999) Subcellular targeting of green fluorescent protein to plastids in transgenic rice plants provides a high-level expression system. Mol Breeding in press.
19. Iglesias VA, Moscone EA, Papp I, Neuhuber F, Michalowski S, Phelan T, Spiker S, Matzke M, Matzke AJM (1997) Molecular and cytogenetic analysis of stably and unstably expressed transgene loci in tobacco. Plant Cell 9: 1251-1264.
20. Kim J-K, Xu D, Duan X, Hwang Y-S Wu R (1994) Expression of a maize ribosome-inactivating protein gene in transgenic rice. Rice Genet Newslett 11: 171-174.
21. Kim J-K, Duan X, Wu R, Seok SJ, Boston RS, Jang I-C, Eun M-Y, Nahm BH (1999) Molecular and genetic analysis of

transgenic rice plants expressing the maize ribosome-inactivating protein b-32 gene and the herbicide resistance bar gene. Mol Breeding 5: 85-94.

22. Komari T, Hiei Y, Saito Y, Murai N, Kumashiro T (1996) Vectors carrying two separate T-DNAs for co-transformation of higher plants mediated by *Agrobacterium tumefaciens* and segregation of transformants free from selection markers. Plant J 10: 165-174.
23. Kung SD (1976) Tobacco fraction I protein: A unique genetic marker. Science 191: 429-434.
24. Kyozuka J, McElroy D, Hayakawa T, Xie Y, Wu R Shimamoto K (1993) Light-regulated and cell-specific expression of tomato rbcS-gusA and rice rbcS-gusA fusion gene in transgenic rice. Plant Physiol 102: 991-1000.
25. Mazur BJ Chui CF (1985) Sequence of a genomic DNA clone for the small subunit of ribulose bis-phosphate carboxylase-oxygenase of tobacco. Nucleic Acids Res 13: 2373-2386.
26. Matzke MA, Matzke AJM. (1995) How and why do plants inactivate homologous (trans)genes. Plant Physiol 107: 679-685.
27. Mlynarova L, Loonen A, Heldens A, Jansen RC, Keizer LCP, Stiekema WJ, Nap J-P (1994) Reduced position effect in mature transgenic plants conferred by the chicken lysozyme matrix-associated region. Plant Cell 6: 417-426.
28. Mlynarova L, Jansen RC, Conner AJ, Stiekema WJ, Nap J-P (1995) The MAR-mediated reduction in position effect can be uncoupled from copy-dependent expression in transgenic plants. Plant Cell 7: 599-609.
29. Mlynarova L, Keizer LCP, Stiekema WJ, Nap J-P (1996) Approaching the lower limits of transgene variability. Plant Cell 8: 1589-1599.
30. Nawrath C, Poirier Y, Somerville C (1994) Targeting of the polyhydroxybutyrate biosynthetic pathway to the plastids of *Arabidopsis thaliana* results in high levels of polymer accumulation. Proc Natl Acad Sci USA 91: 12760-12764.
31. Peng J, Kononowicz H, Hodges TK (1992) Transgenic indica rice plants. Theor Appl Genet 83: 855-863.
32. Shimamoto K, Terada R, Izawa T, Fujimoto H (1989) Fertile transgenic rice plants regenerated from transformed protoplasts. Nature 338: 274-276.
33. Toriyama K, Arimoto Y, Uchimiya H, Hinata K (1988) Transgenic rice plants after direct gene transfer intp protoplasts. Bio/Technology 6 1072-1074.
34. Wong EY, Hironaka CM, Fischoff DA (1992) Arabidopsis thaliana small subunit leader and transit peptide enhance the expression of *Bacillus thuringiensis* proteins in transgenic plants. Plant Mol Biol 20: 81-93.
35. Xu D, Duan X, Wang B, Hong B, Ho T-HD, Wu R (1996) Expression of a late embryogenesis abundant protein gene, HVA1, from barley confers tolerance to water deficit and salt stress in transgenic rice. Plant Physiol 110: 249-257.
36. Zhang W, Wu R (1988) Efficient regeneration of transgenic plants from rice protoplasts and correctly regulated expression of the foreign gene in the plants. Theor Appl Genet 76: 835-840.