JARQ 43 (2), 81 – 85 (2009) http://www.jircas.affrc.go.jp

**REVIEW**

# Towards a Highly Efficient Gene Targeting System in Higher Plants

**Hiroaki SAIKA and Seiichi TOKI***

Division of Plant Sciences, National Institute of Agrobiological Sciences
(Tsukuba, Ibaraki 305–8602, Japan)

## Abstract

Gene targeting (GT) via homologous recombination (HR) is a powerful tool for the precise modification of the target gene. In spite of many efforts, GT frequency in higher plants is still low. An efficient transformation system, enhancement of the HR and efficient selection of the gene targeted cells are thought to be essential for the development of an efficient GT system in plants. Recent reports showed that the open chromatin state of host cells and/or specific induction of the DNA double strand breaks (DSBs) in the gene of interest are important to improve HR and GT. RAD54 is involved in chromatin remodeling during HR. Chromatin assembly factor 1 (CAF-1) is involved in nucleosome assembly after DNA replication. Enhancement of RAD54 function and suppression of the CAF-1 function are expected to induce the open chromatin state in plant cells which could induce HR. Zinc finger nuclease (ZFN), an artificial endonuclease composed of zinc finger domains and a nuclease domain, induce DSBs at the target sequence and enhance GT. In the next step, determining how to apply these potential effectors is a key to establish an efficient GT system.

**Discipline**: Biotechnology
**Additional key words**: homologous recombination, molecular breeding, rice

## Introduction

Gene modification by gene targeting (GT) via homologous recombination (HR) is a powerful tool not only for the analysis of the function of the gene of interest but also for the molecular breeding of crops. The first report of GT in plants was the targeted modification of a transgene in tobacco[13]. To date, successful GT events in higher plants were reported in *Arabidopsis*[2,4,7,11,16], tobacco[8,13] and rice[3,18,20]. However, the frequency of GT events in higher plants is still low. Efficient GT involves not only maximizing the frequency of HR but also improving techniques for the identification of rare recombinants[10]. In this review, several approaches for improving the GT system in plants are discussed (Fig. 1).

## HR frequency

DNA double strand breaks (DSBs) in the genome crucially damage cells and are caused by many environmental stresses such as UV, ionizing radiation and reactive oxygen species. There are two pathways to repair the DSBs, namely HR and nonhomologous end joining (NHEJ). HR is a precise repair system because homologous templates such as sister chromatid and donor DNA (in case of GT) is used. On the other hand, NHEJ is an error-prone repair system and frequently accompanied with deletion or insertion. These pathways compete with each other in DSBs repair. In general, NHEJ is dominant in multicellular eukaryotes such as animals and plants, although HR is dominant in bacteria and unicellular eukaryotes such as yeast (reviewed by Mengiste and Paszkowski[10]). Actually, the frequency of GT via HR is between $10^{-3}$–$10^{-5}$ as the ratio of random integration of donor DNA in plant genome via NHEJ in higher plants (reviewed by Iida and Terada[5]). Thus, it is suggested that an increase of HR activity and/or decrease of NHEJ activity may enhance GT efficiency.

### 1. DSBs

Induction of DSBs at the target locus is thought to be

*Corresponding author: e-mail stoki@affrc.go.jp
Received 31 March 2008; accepted 19 August 2008.

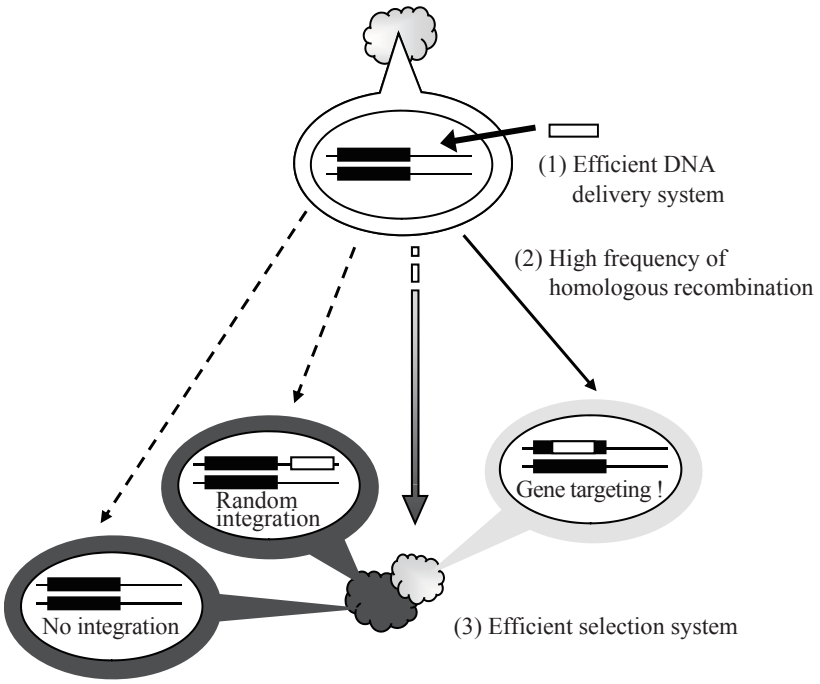


**Fig. 1. Representation of the points to improve conventional GT**

There are mainly three points to improve GT efficiency, (1) transformation system, (2) HR efficiency and (3) selection system. (1) The donor DNA should be efficiently delivered into more plant cells. To achieve this, more competent plant cells for transformation and a more efficient DNA delivery system will be necessary. (2) HR efficiency is thought to be most important to enhance GT frequency. The induction of moderate DSBs or high accessibility of donor DNA to genomic DNA is one of the ways to increase HR frequency. (3) To select only cells in which GT occurred, efficient positive and/or negative markers should be used. Especially, it is essential to prevent from growing cells in which random integration of donor DNA into plant genome occurred. Black and white boxes indicate the targeted gene in rice endogenous targeted locus and donor gene for GT, respectively.

one of the promising ways to increase GT frequency, since HR is one of the DSBs repair systems and its activity is increased by DSBs (reviewed by Puchta[14]). Nuclease such as restriction enzymes has the ability to introduce DSBs in plant genome. The creation of an artificial nuclease which could cleave the target sequence efficiently and specifically is desirable for the establishment of a highly efficient GT system since the excess induction of random DSBs causes the accumulation of undesirable mutations in plant genome. One of the artificial nucleases to cleave the target sequences is zinc finger nuclease (ZFN) which consists of a DNA recognition domain composed of several (usually three) $Cys_2His_2$ zinc fingers and a nuclease domain from endonuclease *Fok*I (Fig. 2, reviewed by Wu et al.[23]). ZFNs have been shown to cleave the target sequences and induce local HR at the target site to repair the DSB in animals such as *Xenopus* oocyte, human cells, *Caenorhabditis elegans*, and *Drosophila*[23]. In higher plants, ZFNs were used to cleave the target sequences and function as a mutagen[9] or enhancer of GT[22].

Lloyd et al. reported the efficient induction of the

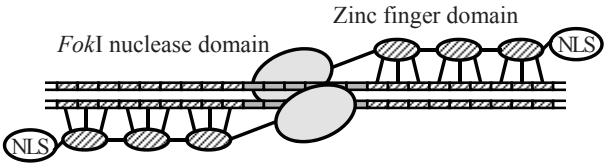


**Fig. 2. A schematic representation of a pair of 3-finger ZFNs**

Two ZFNs designed as a tail-to-tail conformation are necessary to recognize and cleave the target sequence. The 6 bp cleavage sequences are located in 2 zinc-finger binding sites. The gray and shaded circles in ZFNs indicate the *Fok*I nuclease domain and zinc finger domain to recognize the target triplet sequences, respectively.

mutations at a specific site by ZFN in *Arabidopsis*[9]. In this study, transgenic plants transformed with a HS::QQR-QEQ construct were used. HS::QQR encodes a ZFN, which recognizes 5'-GGGGAAGAA-3', driven by a heat-shock inducible promoter. Overexpression or constitu-

tive expression of ZFN causes mutation at non-canonical sites because abundant DSBs cannot be repaired completely. QEQ sequences contain QQR-ZFN binding sites and an intervening *Eco*RI site as a Cleaved Amplified Polymorphic Sequence (CAPS) marker to estimate the frequency of cleavage by QQR-ZFN. DSBs caused by QQR-ZFN are thought to be repaired by NHEJ rather than HR. NHEJ is extremely error-prone in DSBs repair in higher plants, *Eco*RI in the QEQ site is prone to be mutated and the frequency of cleavage by QQR-ZFN is estimated by the rate of mutation in *Eco*RI. The frequency of mutation induction by QQR-ZFN was estimated to be as high as 0.2 per gene. In these mutations, 78% were simple deletions of 1–52 bp, 13% were simple insertions of 1–4 bp and 8% were both deletions and insertions[9]. These data showed that ZFN could mutagenize the target gene efficiently. Wright et al. reported that ZFN enhanced GT frequency in tobacco protoplast[22]. In this study, protoplasts transformed with the construct of non-functional gus::nptII fusion gene because of a 600 bp deletion and an intervening ZFN recognition site were used. The partial GUS::NPTII fusion gene was used as donor DNA to repair the defective gus::nptII by GT. The donor DNA is not thought to be functional by itself since it lacks the 5' region sequences. GT occurred in more than 10% of protoplasts introduced donor DNA by electroporation. GT frequency with ZFN was more than three-fold higher than without ZFN[22], suggesting that ZFN is efficient for the increase in GT frequency.

However, there are some points to improve ZFN technology for the application to GT. All available zinc finger motifs that recognize every DNA triplet bases have not been found yet. Some ZF motifs which recognize CNN and TNN triplets are not available, while all ZF motifs which recognize GNN and ANN triplets are available[23]. Moreover, ZFN applied for endogenous genes in plants have not been reported yet.

**2. Chromatin remodeling**

Genomic DNA is tightly compacted with histone and non-histone proteins and forms the chromatin structure. The state of the chromatin structure changes reversibly. Open chromatin state, which is a relaxed chromatin structure, results in the high accessibility of donor DNA to genomic DNA and the increase in HR frequency. This suggests that the regulation of chromatin state by chromatin remodeling factors affects GT efficiency.

RAD54 is a member of the SWI2/SNF2 chromatin remodeling factors. In the step of HR, RAD54 protein interacting with RAD51 nucleoprotein and single strand DNA is implicated in the search for homologous sequence and chromatin remodeling to form a D-loop structure. RAD54 is one of the promising factors to enhance HR frequency. In *Arabidopsis*, T-DNA insertion mutant or RNAi line of *RAD54* (*AtRAD54*) gene was sensitive to genotoxic stresses such as gamma ray, UV-C and DNA cross-linking reagent, cisplatin[12,17]. Moreover, in a T-DNA insertion mutant of *AtRAD54*, the frequency of intra- and inter-chromosomal HR was lower than that in the wild-type[12]. These data show that RAD54 protein plays important roles in DNA repair and HR in *Arabidopsis* and is possible to enhance GT frequency. Actually, overexpression of yeast RAD54 (ScRAD54) enhanced the GT by one or two orders more than the wild-type but did not affect the random integration of donor DNA[16]. However, the frequency of intrachromosomal HR in a RNAi line of *AtRAD54* was comparable with the wild-type[17]. One possible reason is that the gene expression of *AtRAD54* in the RNAi line was not completely repressed so that a small amount of RAD54 affects HR frequency. Moreover, the overexpression of *AtRAD54* gene did not enhance the tolerance to genotoxic stresses or the frequency of HR[12]. It is discussed that overexpression of AtRAD54 alone was not sufficient for the increase in HR efficiency unlike ScRAD54 or excess AtRAD54 protein has little effect on intrachromosomal HR. Thus, RAD54 protein functions to enhance GT efficiency, but fine regulation of RAD54 expression might be necessary to apply to an efficient GT system.

Chromatin assembly factor 1 (CAF1) functions as a nucleosome chaperone in DNA replication (reviewed by Ramirez-Parra and Gutierrez[15]), suggesting that replicated DNA is thought to be left naked for a longer period in the *caf1* mutant. In *Arabidopsis caf-1* (*atcaf-1*) mutant, gene expression involved in HR such as *RAD51* or *RAD54* was up-regulated, although the gene expression involved in NHEJ such as *Ku70, Ku80* or *LIGIV* was comparable to wild-type[1]. Moreover, the level of DSBs was increased and G2 retardation in the cell cycle was observed in the *atcaf-1* mutant, resulting in an increase of the HR frequency; 40-fold in *atcaf-1* mutant[1]. These results suggest that the suppression of CAF-1 activity may potentially increase GT frequency.

## Efficient selection of the gene modified plants

To select only cells or plants in which GT occurs successfully, it is very important to distinguish clearly between cells in which GT occurred and others, non-transformed cells and cells in which donor DNA is inserted at random. There are mainly two strategies to select GT cells; the selection using exogenous markers and the selection using endogenous target gene specific markers (Fig. 3). As exogenous selection markers, hygromycin

phosphotransferase (HPT)[18,20] and neomycin phosphotransferase (NPTII)[7,11] were used as positive markers to confer tolerance against antibiotics, and GUS[7,11] and GFP[16] were used as positive visible markers. Diphtheria toxin A fragment (DT-A)[18,20] was used as a negative marker to stop growth of cells in which T-DNA integrates at random. DT-A protein inactivates the elongation factor 2 and was demonstrated to work as a strong negative marker of GT in rice[18–20]. On the other hand, *protoporphyrinogen oxidase* (*PPO*) gene and *acetolactate synthase* (*ALS*) gene were used as target gene specific positive selection markers[2–4,8].

Terada et al. succeeded in generating the knockout plants of *Waxy* gene and *ADH2* gene by positive-negative selection system using exogenous markers in rice[18,20]. In their GT vector, HPT expression cassette is inserted into each gene and DT-A expression cassettes are added at both ends (Fig. 3A). In this system, cells in which double crossovers occur at both of the HPT-flanking homologous regions on T-DNA and DT-A cassettes are not inserted into genome DNA, can grow. GT frequency (targeted calli per surviving callus) was estimated at 0.94% and 1.9% in *Waxy* gene or *ADH2* gene, respectively and was comparable with the GT frequency in mouse embryogenic stem cells. Moreover, in this system, no ectopic GT, in which the modified sequence copied from the target locus integrates elsewhere in the genome, was observed[18,20]. On the other hand, Endo et al. succeeded in regenerating plants which were tolerant to the herbicide, bispyribac (BS) by introducing two base substitutions which confer tolerance to BS in rice *ALS* gene using GT[3]. The partial *ALS* coding sequence with two mutations was used as GT vector for rice *ALS* gene (Fig. 3B). The partial *ALS* gene in T-DNA is not thought to be functional when it integrates into rice genome because it has no chloroplast-targeting signal sequence. Moreover, ALS with two point mutations by GT works as a positive selection marker against BS. In this strategy, Endo et al. could regenerate 66 independent plants from 1,500 calli and true GT, in which the wild-type gene was modified as expected, in *ALS* locus occurred without any rearrangements in approximately two-thirds of regenerated BS-tolerant plants[3]. Interestingly, the analysis of enzymatic activity showed that GT plants which had homozygous modified ALS locus conferred extreme tolerance to BS[3], suggesting that GT technology can completely eliminate the wild-type BS-sensitive ALS locus and be a powerful tool for complete substitution of a target gene. This selection method is applied to modify the gene, which is a target of toxic reagents such as herbicides and antibiotics, to confer tolerance to such reagents. Indeed, we recently succeeded in the modification of *OASA2* gene encoding α subunit of anthranilate synthase in rice using the same method (Saika et al., unpublished data).

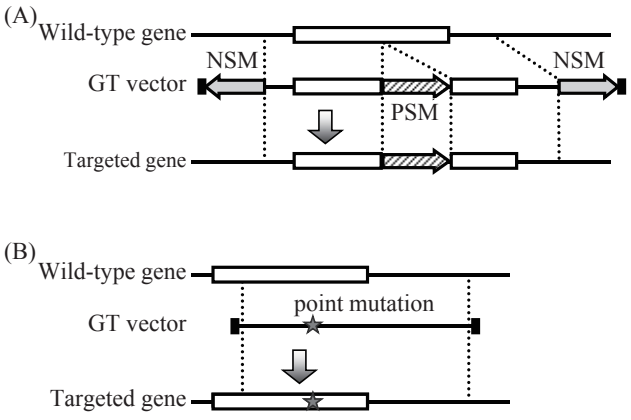


**Fig. 3. Vector and selection system in rice GT**
(A) GT in rice *Waxy* and *ADH2* genes[18,20]. To knockout the targeted gene, positive selection marker (PSM) expression cassette is inserted into each gene. At both ends of each gene, negative selection marker (NSM) expression cassettes are added to prevent inserting T-DNA into rice genome at random. (B) GT in rice *ALS* gene[3]. To introduce mutations (shown as a star) into the targeted gene, the sequence of *ALS* gene induced the mutation and deleted chloroplast targeting signal sequence was used as a GT vector.

GT frequency is thought to be dependent on not only the selection system but also the vector construction. Especially, the length of homologous sequence between the GT vector and the targeted locus is suggested to affect GT efficiency. The analyses in moss showed that there is a significant dependency of the frequency of HR on length of homology[6]. However, the length of homologous sequence in GT vector for HR to occur efficiently at the targeted locus has not been elucidated in angiosperms. There still remain points to improve in vector construction.

## Efficient transformation system

The establishment of a highly efficient transformation system is thought to be one of the approaches to increase GT efficiency, since more transformed calli result in the increase of the number of cells in which GT events occurred and initial amounts of rice callus for transformation can be reduced. In the rice variety Nipponbare, a highly efficient transformation system has been established. In our system, several hundreds of independent transformation events are observed in every small callus[21]. The GT system will be widely utilized if a very highly efficient transformation system in other plants

such as maize is established.

## Conclusion

GT is one of the powerful tools for the modification of the targeted gene. However, much more effort is needed for the establishment of a GT system which could be applied to any gene of interest as a routine technique. There might be more efficient factors other than ZFN and RAD54, which are possible factors to enhance HR activity. Anyway, the establishment of a model GT system, in which GT frequency is estimated rapidly and accurately, is thought to be essential to improve the conventional GT technology. Now, we are going to establish a model GT system for evaluating GT efficiency and find out the factors that function to enhance GT efficiency. Then, the next step to establish an efficient GT system is determining how to apply these selected factors.

## References


1. Endo, M. et al. (2006) Increased frequency of homologous recombination and T-DNA integration in *Arabidopsis* CAF-1 mutants. *EMBO J.*, **25**, 5579–5590.
2. Endo, M. et al. (2006) Molecular characterization of true and ectopic gene targeting events at the acetolactate synthase gene in *Arabidopsis*. *Plant Cell Physiol.*, **47**, 372–379.
3. Endo, M. et al. (2007) Molecular breeding of a novel herbicide-tolerant rice by gene targeting. *Plant J.*, **52**, 157–166.
4. Hanin, M. et al. (2001) Gene targeting in *Arabidopsis*. *Plant J.*, **28**, 671–677.
5. Iida, S. & Terada, R. (2005) Modification of endogenous natural genes by gene targeting in rice and other higher plants. *Plant Mol. Biol.*, **59**, 205–219.
6. Kamisugi, Y., Cuming, A. C. & Cove, D. J. (2005) Parameters determining the efficiency of gene targeting in the moss *Physcomitrella patens*. *Nucleic Acids Res.*, **33**, e173.
7. Kempin, S. A. et al. (1997) Targeted disruption in *Arabidopsis*. *Nature*, **389**, 802–803.
8. Lee, K. Y. et al. (1990) Homologous recombination in plant cells after *Agrobacterium*-mediated transformation. *Plant Cell*, **2**, 415–425.
9. Lloyd, A. et al. (2005) Targeted mutagenesis using zinc-finger nucleases in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA*, **102**, 2232–2237.
10. Mengiste, T. & Paszkowski, J. (1999) Prospects for the precise engineering of plant genomes by homologous recombination. *Biol. Chem.*, **380**, 749–758.
11. Miao, Z. H. & Lam, E. (1995) Targeted disruption of the TGA3 locus in *Arabidopsis thaliana*. *Plant J.*, **7**, 359–365.
12. Osakabe, K. et al. (2006) Isolation and characterization of the *RAD54* gene from *Arabidopsis thaliana*. *Plant J.*, **48**, 827–842.
13. Paszkowski, J. et al. (1988) Gene targeting in plants. *EMBO J.*, **7**, 4021–4026.
14. Puchta, H. (2005) The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. *J. Exp. Bot.*, **56**, 1–14.
15. Ramirez-Parra, E. & Gutierrez, C. (2007) The many faces of chromatin assembly factor 1. *Trends Plant Sci.*, **12**, 570–576.
16. Shaked, H., Melamed-Bessudo, C. & Levy, A. A. (2005) High-frequency gene targeting in *Arabidopsis* plants expressing the yeast *RAD54* gene. *Proc. Natl. Acad. Sci. USA*, **102**, 12265–12269.
17. Shaked, H., Avivi-Ragolsky, N. & Levy A. A. (2006) Involvement of the Arabidopsis SWI2/SNF2 chromatin remodeling gene family in DNA damage response and recombination. *Genetics*, **173**, 985–994.
18. Terada, R. et al. (2002) Efficient gene targeting by homologous recombination in rice. *Nat. Biotechnol.*, **20**, 1030–1034.
19. Terada, R., Asao, H. & Iida, S. (2004) A large-scale *Agrobacterium*-mediated transformation procedure with a strong positive-negative selection for gene targeting in rice (*Oryza sativa* L.). *Plant Cell Rep.*, **22**, 653–659.
20. Terada, R. et al. (2007) Gene targeting by homologous recombination as a biotechnological tool for rice functional genomics. *Plant Physiol.*, **144**, 846–856.
21. Toki, S. et al. (2006) Early infection of scutellum tissue with *Agrobacterium* allows high-speed transformation of rice. *Plant J.*, **47**, 969–976.
22. Wright, D. A. et al. (2005) High-frequency homologous recombination in plants mediated by zinc-finger nucleases. *Plant J.*, **44**, 693–705.
23. Wu, J., Kandavelou, K. & Chandrasegaran, S. (2007) Custom-designed zinc finger nucleases: What is next? *Cell Mol. Life Sci.*, **64**, 2933–2944.