농학박사학위논문

Construction, Selection and Application of Zinc Finger Nucleases as a Tool for Genome Editing in Rice

2012년 8월

서울대학교 대학원

식물생산과학부 원예육종전공

김 현 중

Abstract

Construction, Selection and Application of Zinc Finger Nucleases as a Tool for Targeted Genome Editing in Rice

Hun Joong Kim

Department of Plant Sciences

The Graduate School

Seoul National University

Zinc finger nucleases (ZFNs) are artificial restriction enzymes that can be custom-made to recognize and cleave DNA in virtually any user-defined location, leading to genome editing at user-specified genomic sites. Genome editing by ZFNs has proven to be applicable to diverse organisms including

plants. But, among monocotyledonous plants only maize has been reported to be successful (Shukla et al. 2009). With a goal of widening the applicability of ZFNs, I intended to target a rice gene *OsAP2* with ZFNs. *OsAP2* encodes a member of transcription factor with AP2 domain and is known to increase grain yield particularly during drought stress when expressed under a root-specific promoter (Oh et al. 2009). With an aim of mutagenizing *OsAP2* with ZFNs, I constructed 61 pairs of 4F-ZFNs, by modular assembly, with cleavage potential at sites in *OsAP2* protein coding region. Of the 61 ZFNs, 4 pairs of ZFNs showing high cleavage activity were selected by yeast single-strand annealing (SSA) assay. Of these, a pair of ZFNs with highest activity were chosen for delivery into rice embryogenic callus via *Agrobacterium*. The ZF modules were linked to either wild-type (F67/F96) or obligate heterodimerization variant (F67-DAS/F96-RR or F67-RR/F96-DAS) and the two ZFN monomers recognizing adjacent half-sites were linked via 2A sequence to ensure simultaneous expression of both ZFN monomers by a single promoter. The whole ZFN sequences (excluding *Fok*I variants, F67-DAS/F96-RR or F67-RR/F96-DAS) were rice codon-optimized. The resulting ZFN sequences were placed downstream of two different promoters: constitutive PGD1 and root-specific RCc3. A total of 6 distinct Ti-plasmid-based vectors were then delivered into rice embryogenic callus. As a result, a total of 100 T_0 transgenic plants were obtained and T1 seeds were harvested from the 15 plants. Genomic PCR analysis showed that 80% of T_0 transgenic plants had integrated ZFNs in their genome but one of the PCR positive plants was found to have a partial internal deletion in ZFN. GFP analysis of dried rice

seeds yielded positive signals in all transgenic plants analyzed, except one with a partial internal deletion. Immunoblot analysis also exhibited high levels of ZFN protein expression in transgenic rice tissue. To assess ZFNs' ability of targeted mutagenesis in rice, T7 endonuclease analysis, a mismatch selective endonuclease assay, was performed. In 5 T_0 plants among those analyzed, strong evidence supporting occurrence of targeted mutation in somatic tissue was observed. For further analysis, one of the plants was selected to clone the genomic region spanning the target site in *OsAP2* and subjected to sequence analysis. Of 160 clones, 4 clones were identified to have substitution mutations at the expected ZFN target region. In the 4 clones, a total of 6 substitution mutations were found, a mutation frequency far exceeding that expected for spontaneous errors by *Taq* polymerase (between 1×10^{-4} to 2×10^{-5} errors per base pair). This strongly suggests the targeting events had occurred by ZFNs. Although the lack of indel (insertion/deletion) appears to reflect the widespread occurrence of selective cell death as mutation accumulates at off-target sites from prolonged expression of ZFN proteins, genotyping for T_1 and T_2 generations will be of great help not only in finding germline indels but clarifying causes of the lack of somatic indel mutations.

Keywords : zinc finger nuclease (ZFN), genome editing, double-strand break (DSB), nonhomologous end joining, *OsAP2*

Student number : 8425-3313

CONTENTS

CHAPTER II

Application of Zinc Finger Nuclease Technology To Rice (*Oryza sativa* L.)

LIST OF TABLES

CHAPTER I

CHAPTER II

LIST OF FIGURES

LITERATURE REVIEW

CHAPTER I

CHAPTER II

LIST OF ABBREVIATIONS

DSB	Double-strand break
EtBr	Ethidium bromide
GT	Gene targeting
HR	Homologous recombination
HRM	High resolution melting
MCS	Multiple cloning site
NHEJ	Non-homologous end joining
ORF	Open reading frame
p3-1F	p3 vector fused with a single zinc finger
p3-1F-*Fok*I	p3 vector containing a single zinc finger-*Fok*I fusion
PCR	Polymerization chain reaction
RI	Random integration
ZF	Zinc finger
ZFP	Zinc finger protein
ZFN	Zinc finger nuclease
ZFN-L	4F-ZFN monomer recognizing the left half-site of the target sequence
ZFN-R	4F-ZFN monomer recognizing the right half-site of the target sequence
1F	single zinc finger

2F	two-zinc finger
4F	four-zinc finger
1F-*Fok*I	*Fok*I domain fused with a single zinc finger
1F-ZFN	Zinc finger nuclease with single zinc finger
2F-ZFN	Zinc finger nuclease with two contiguous zinc fingers
4F-ZFN	Zinc finger nuclease with four contiguous zinc fingers

GENERAL INTRODUCTION

Zinc finger nuclease (ZFN) has proven to be a powerful tool for targeted genome editing (Carroll 2011; Tzfira et al. 2012). Since its first introduction as a tailor-made restriction enzyme (Kim et al. 1996), followed by the first generation of ZFN-directed mutant organism (Bibikova et al. 2002), ZFN has found its way to a broad range of applications: reverse genetics in model organisms, efficient genome engineering in human stem cells and targeted integration in crop plants (Urnov et al. 2010).

ZFNs are engineered sequence-specific endonucleases capable of executing efficient targeted genome editing. Targeted genome editing results from targeted genome cleavage by ZFNs followed by gene modification during subsequent repair (Urnov et al. 2010). Genome editing differs from gene targeting coined by Mario Capecchi and colleagues (Thomas et al. 1986). Gene targeting refers to a replacement strategy based on homologous recombination (HR), whereas double-strand-break-assisted genome editing comprises both targeted mutagenesis mediated by nonhomologous end joining (NHEJ) and targeted gene replacement mediated by HR (Urnov et al. 2010).

As described in literature (Carroll 2011; Urnov et al. 2010; Wu et al. 2007; Tzfira et al. 2012), ZFN's capacity to modify pre-determined genomic sequences *in vivo* holds great potential for revolutionizing not only plant reverse genetics but also plant breeding. For example, the wide range of challenges to plant breeding such as efficient targeted gene modification, precise analysis of gene function, targeted homologous recombination, and

making all these in non-transgenic manner could be overcome, at least in part, by the ZFN technology.

Since the first heritable endogenous gene targeting by ZFN in crop plants (Townsend et al. 2009; Shukla et al. 2009), considerable progress has been made in implementing engineered nuclease technology (e.g. ZFNs) to plant genome-editing (Carroll 2011; Tzfira et al. 2012). Yet, it is still in the infant stage of development at least in plant applications, far from being routinely usable technology for one reason or another (Weinthal et al. 2010). For example, of monocotyledonous plants, to which all the important grain crops belong, only maize has been reported successful in ZFN-mediated gene targeting (Shukla et al. 2009; Tzfira et al. 2012), partly reflecting the recalcitrant nature of monocotyledonous plants in genetic manipulation.

With this in mind, I intended to apply ZFN technology to targeted mutagenesis of rice *OsAP2* gene. I chose rice (*Oryza sativa* L.) because it is one of the most important food crops in the world, has a relatively small genome size (about one tenth of that of human), and is amenable to biotechnology as known to be a monocot model plant for molecular studies.

The *OsAP2* gene encodes a transcription factor presumed to control flower development but precise gene function in rice has not been defined yet. Interestingly, when *OsAP2* is overexpressed under a root-specific promoter, significant yield increase ensued, particularly under drought conditions (Oh et al. 2009). But neither knock-out nor knock-down mutants for *OsAP2* were available to compare with, which prompted me to

pick up the *OsAP2* gene as a gene of interest for targeted mutagenesis.

The creation of highly specific ZFNs that recognize and cleave any target sequence depends on the reliable design and/or selection of ZFNs that can recognize and bind the chosen target with high affinity and specificity (Wu et al. 2007). The usual sequence of steps for use of ZFNs to modify endogenous genes includes: 1) identifying potential ZFN targets in gene of interest using web-based software, 2) generating multiple ZFNs, 3) screening for binding activity, 4) screening for nuclease activity, and 5) using most active pairs from screening to site-specifically modify endogenous gene (Porteus 2009; Wu et al. 2007; Maeder et al. 2009).

In my study, the step '3) screening for binding activity' was skipped to reduce workload. For construction of library of 4-finger ZFNs (Step 2), I used 'modular assembly method' (Kim et al. 2009) to take advantage of established system in the lab where I worked. As a method of selecting active ZFNs (Step 4), I chose yeast homologous recombination based assay because of ease of experiment as well as possibility of testing ZFN activity in the chromatin context in live eukaryotic cells.

In short, to induce targeted mutagenesis in the rice *OsAP2* gene, I conducted experiments as follows:

First, I constructed candidate zinc finger nucleases using modular assembly method that would target potential sites at rice *OsAP2* coding region.

Subsequently, active zinc finger nuclease pairs were selected *in vivo* by budding yeast single-strand annealing assay on the basis of their ability to induce homologous recombination.

Finally, the selected ZFN pairs were delivered to rice tissue via *Agrobacterium* to induce targeted mutagenesis at rice *OsAP2* gene.

LITERATURE REVIEW

1. Targeted mutagenesis in plant breeding

Targeted mutagenesis is important in plant breeding because many commercial varieties are derived from single gene mutations (Hanin and Paszkowski 2003; Puchta 2002; Wilde et al. 2012). For example, natural mutations in single genes of tomato are completely or mostly responsible for its determinate growth habit (Pnueli et al. 1998), resistance to powdery mildew (Bai et al. 2008), and yield heterosis (Krieger et al. 2010).

Before the advent of engineered nucleases such as zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs), the successful targeted mutagenesis in crops has been very limited because of lack of efficient tools. In the past, gene targeting (GT), the modification of an endogenous gene into a designed sequence by homologous recombination (HR), has been only rarely successful because of competing random integration (RI) of transgene. The targeted integration of transgene in the endogenous homologous sequence in higher plants are known to be in the order of 0.01% to 0.1% compared to RI (Terada et al. 2007). Thus GT using oligonucleotides has never emerged as a tool of any practical significance.

Alternative approaches are not in the category of targeted mutagenesis in a strict sense but the goal is identical. These approaches rely on highly efficient screening methods that selectively identifying mutated genes of interest. These methods first induce a large pool of random mutations throughout the genome and then using high throughput screening methods identify sequence variation at a site of interest. The

most commonly used method is TILLING (Till et al. 2007). TILLING detects mutations in the re-annealed PCR products by mismatch-sensitive endonucleases. High resolution melting (HRM) detects mutations on the basis of changes in melting kinetics of PCR products in DNA with mismatches (Gady et al. 2009). High-throughput sequencing of PCR products from candidate genes is an alternative to DNA mismatch detection (Wilde et al. 2012; Gilchrist and Haughn 2010).

The alternative approaches are, in a sense, similar to T-DNA or transposon tagging in that the process involves global mutagenesis followed by screening variations, which make them distinguishable from the targeted mutagenesis using engineered nucleases.

2. Zinc finger nucleases (ZFNs)

2.1 Development of ZFNs

Zinc-finger nuclease (ZFN), first developed by Kim et al. (1996), is a custom-made fusion protein: a sequence-specific zinc finger DNA-binding domain linked to a nonspecific *Fok*I nuclease domain. The first example of permanent genetic alterations directed by ZFNs was made by Bibikova et al. (2002). They targeted *Drosophila yellow* gene by injecting plasmids encoding ZFNs into embryos. They achieved high rates of both somatic and germline mutations resulting from nonhomologous end joining (NHEJ). They used hsp70 promoter to express ZFNs successfully although heat shock had to be moderated to avoid ZFN toxicity. Ever since the first demonstration by Bibikova et al. (2002), ZFN technology has been successfully adopted to diverse cells and organisms including human cells

(Porteus and Baltimore 2003; Urnov et al. 2005), human stem cell (Lombardo et al. 2007), zebrafish (Doyon et al. 2008; Meng et al. 2008), tobacco (Townsend et al. 2009), maize (Shukla et al. 2009), mouse (Meyer et al. 2010), *Arabidopsis* (Osakabe et al. 2010; Zhang et al. 2010) and *Xenopus* (Young et al. 2011).

2.2 Structure and function relationship of ZFNs

Structurally, ZFN can be divided into three components: i) zinc finger DNA-binding domain, ii) the inter-domain linker that connects zinc finger and *Fok*I domains, and iii) the *Fok*I cleavage domain. ZFNs act as dimers and each monomer binds one of the two inverted target half-sites. Between the two half-sites lies a spacer, 5 to 7 bp in length, where cleavage occurs (Carroll 2011; Porteus 2007; Urnov 2010; Wu et al. 2007).

The zinc finger (ZF) DNA–binding domain of ZFN typically consists of three or four tandem array of Cys_2His_2 zinc finger protein (ZFP) modules. Since each ZF module recognizes a 3 bp “sub-site”, a single ZFN monomer makes contact with 9- or 12-bp sequence (“half-site” of ZFN binding), depending on the number of ZF module, and therefore, a pair of ZFNs can recognize 18- to 24-bp of “full-site” (Pavletich and Pabo 1991). The ZFP members of the Cys_2His_2 class are the most common form of ZFP and best suited for customization because of their diversity and modular structure (Bae et al. 2003). An individual Cys_2His_2 zinc finger consists of ~30 amino acids arranged in a ββα structure that is stabilized by chelating a single zinc ion. The α-helix of the zinc finger, responsible for specific contact with DNA, lies in the major groove. The binding of DNA is

mediated by amino acid residues -1 to 6 (of which -1, 2, 3 and 6 are most important) of the α-helix with respect to the beginning of the helix (Bae et al. 2003; Jamieson et al. 2003; Pabo et al. 2001; Porteus 2007; Wu et al. 2007) (Figure 1). The ZFP modules can be assembled in the specific order necessary to recognize any given sequence in a target locus and this makes ZFNs a versatile tool for targeted genome editing (Alwin et al. 2005).

2.3 Design and Assembly of ZFNs

There are several ways to construct ZF arrays that can recognize potential target sequences. Publicly available methods for engineering ZF domains include: i) Modular Assembly (Wright et al. 2006). ii) A selection based method, Oligomerized Pool Engineering (OPEN) (Maeder et al. 2008; Maeder et al. 2009), and iii) Context-dependent Assembly (CoDA) (Sander et al. 2011).

Modular assembly methods involve assembly of ZF modules with predetermined DNA-binding specificities. This method simply links ZF modules together using standard recombinant DNA technology and would provides a rapid and easy-to-practice platform for ZFN engineering (Bae et al. 2003; Wright et al. 2006; Kim et al. 2009) but often resulted in high failure rate for making active ZFNs probably due to its inability to reflect sequence context (Segal 2011).

The OPEN method uses an ‘open-source’, combinatorial selection-based method for engineering ZF arrays and thus allows for the context-dependent engineering of three zinc fingers simultaneously. OPEN is claimed to be more effective than the modular assembly method, but it is

time-consuming and labor-intensive (Maeder et al. 2008; Meader 2009; Sander et al. 2011). OPEN requires that a new library of protein variants be constructed and screened for every desired target site, limiting the broad translation of these successes (Segal 2011).

The CoDA method is the newest and takes advantage of many hundreds of OPEN proteins now in database. In this method, two different three-zinc-finger arrays, each engineered to bind different 9-bp target sites and that have in common a middle finger unit (the F2 unit), are used to create a three-finger array with a new specificity by joining together the F1 unit from the first array, the F2 unit, and the F3 unit from the second array (Sander et al. 2011). CoDA method is a sort of compromise between context-dependent engineering and simple modular assembly of three-zinc-finger proteins (Segal 2011) and has an advantage of simplicity and efficacy due to context consideration (Sander et al. 2011). However, since it is only applicable to three-zinc-finger proteins at present, CoDA method has a potential limitation in constructing 4- or longer- zinc finger proteins.

Recently, Kim et al. (2011) reported an improved version of modular assembly method, in which they used a select set of 33 fingers that are frequently found in active ZFNs, to construct combinatorial library of two-finger modules that can be used to assemble 3-finger or 4-finger ZFNs in a single subcloning step.

2.4 Selection of active ZFNs

In their pioneering report on the first engineering of ZFNs, Kim et al.

(1996) have clearly showed that the sequence specificity of the ZFNs could be altered during the assembly of zinc finger modules. In other words, ZFNs assembled using ZFs that recognize known triplets do not always have the desired sequence specificity when assembled into arrays (Porteus 2007). Careful studies of zinc finger DNA binding have shown that in addition to helical positions -1, 3 and 6 of each zinc finger that contact three contiguous nucleotide triplet, position 2 also plays an important role in binding zinc finger to DNA. The α-helical position 2 interacts with a base on the other strand of the DNA, complementary to target recognized by the amino acid in position 6 of the preceding finger (Figure 1). The cross-strand interaction extends the binding site of a zinc finger protein from 3 to 4 bases and also synergistically links the fingers into an overlapping array (Papworth et al. 2006). Therefore, zinc finger DNA binding is not completely modular and building of a given finger depends on its neighbors. Thus, there is “context” dependence. An important caveat to the modular-assembly approach to designing new ZFPs is that this context dependence is ignored (Porteus 2007). For this reason, one needs to develop a fast and efficient screening method to accommodate the need for favorable cross-strand interaction in the neighboring fingers (Kim et al. 1996; Papworth et al. 2006).

There have been various *in vitro* and *in vivo* screening methods reported in the literature. Examples of such strategies include the electrophoretic mobility shift assay (EMSA) (Bae et al. 2003), the phage display (Alwin at al. 2005; Isalan et al. 2001), the bacterial-two-hybrid (B2H) (Joung et al. 2000; Maeder et al. 2008; Wright 2006; Zhang et al.

2010), the mammalian cell single-strand annealing (SSA) (Kim et al. 2009), and yeast SSA (Doyon et al. 2008). All of these methods rely on the affinity of ZFs for their putative cognate target sites for screening. This is because, more than likely, the specificity of ZFNs is determined solely by the DNA-binding properties of the zinc finger motifs (Kim et al. 1996). Although affinity-based screening approaches have been successful in diverse experimental settings and combining two or more screening methods would give expectations of better outcome, it is also important to consider cost and labor involved because most of such techniques usually demand considerable amount of time and effort. Yet, it appears that there is no consensus as to which approach is the best for a given experimental setting.

The yeast SSA assay system has an advantage of relying on eukaryotic system while retaining microbial easiness. In fact, growing number of researchers are relying on yeast selection system and its successes are remarkable (Doyon et al. 2008; Townsend et al. 2009; Zhang et al. 2010). The principle behind the SSA assay is based on the restoration of an auxotrophic marker that was previously disrupted by an insert-containing target sequence of interest (Chames et al. 2005). In budding yeast, a DSB introduced between two short direct repeats spaced by a heterologous sequence of up to 25 kb leads to homologous recombination between the direct repeats, taking away the entire intervening DNA and resulting in the restoration of chromosome integrity (Doyon et al. 2008).

2.5 ZFN toxicity

Cytotoxicity has long been a keen issue in the ZFN community (Alwin et al. 2005; Bibikova et al. 2002; Cornu et al. 2008; Kim et al. 2009; Miller et al. 2007; Osiak et al. 2011; Pattanayak et al. 2011; Ramalingam et al. 2011; Szczepek et al. 2007). Homo-dimers of ZFNs are known to cause cytotoxicity by targeting off-target sites (Carroll 2011; Durai et al. 2005; Urnov et al. 2010; Ramalingam et al. 2011). When off-target cleavage is extensive, the number of breaks outstrips the DNA repair capacity and leads to death of the treated cells or organisms (Carroll 2011; Bibikova et al. 2002; Porteus and Baltimore 2003; Alwin et al. 2005). The effect can be reduced by making *Fok*I form obligate heterodimers (Doyon et al. 2011; Guo et al. 2010; Ramalingam et al. 2011; Miller et al. 2007; Szczepek et al. 2007; Söllü et al. 2010).

3. Plant application of ZFNs

ZFNs are targetable DNA cleavage reagents that have been adopted as gene-targeting tools (Carroll 2011). Previous reports have demonstrated the potential utility of ZFN technology as a powerful tool for plant breeding by introducing heritable targeted modifications at specified genomic sites (Lloyd et al. 2005; Zhang et al. 2010; Osakabe et al. 2010; de Pater et al. 2009; Wright et al. 2005; Townsend et al. 2009; Cai et al. 2009; Marton et al. 2010; Shukla et al. 2009; Tzfira et al. 2012). Even some cultivars thus produced are in field trials (Waltz 2012). Advantages using ZFN in crop breeding include i) one-step targeted modification of an endogenous gene, eliminating the need for generations of backcrossing, ii) precise

modification which preserves surrounding genomic context, enabling precise functional studies of genes or regulatory elements, and iii) uncoupling of the site of transgene integration and the site of modification, enabling transgene (e.g. ZFN) removal by subsequent crosses, thus capable of being treated as non-transgenic as mentioned by Waltz (2012).

Only a handful of reports are available in applications of ZFNs to plant endogenous genes so far (Tzfira et al. 2012). No endogenous gene targeting was reported until Townsend et al. (2009) and Shukla et al. (2009) published their works. There is only one report for monocotyledonous plants so far (Shukla et al. 2009). Even in *Arabidopsis* targeted mutagenesis for endogenous genes was reported as recently as 2010 (Osakabe et al. 2010; Zhang et al. 2010).

They all chose a target gene with selectable or screenable phenotype to facilitate experiments and also used 2A sequence to express ZFN pair from a single plasmid. Townsend et al. (2009) and Shukla et al. (2009) expressed ZFN in plants by constitutive promoters (CaMV 35S and maize ubiquitin-1, respectively) but Osakabe et al. (2010) and Zhang et al. (2010) used inducible promoters (heat-shock promoter and estrogen-inducible promoter, respectively) to express ZFNs. Townsend et al. (2009) delivered ZFN expression plasmid to tobacco protoplasts via electroporation; Shukla et al. (2009) did by Whiskers (Petolino and Arnold 2009). Osakabe et al. (2010) and Zhang et al. (2010) introduced Ti-plasmids containing ZFN genes by the floral dip method.

More recent reports deal with similar successes but most other earlier reports deal with targeting made to the introduced target sequences

into genome. Also there are some reports concerning ZFN-directed mutagenesis in protoplasts (Wright et al. 2005).

A study by Shukla et al. (2009) is of note because it is the first and only report on ZFN-directed gene targeting in monocotyledonous crops. They delivered plasmids expressing ZFNs designed to target the IPK1 gene in maize by whisker-mediated transformation.

4. Rice *OsAP2* gene

OsAP2 refers to rice *AP2* gene. APETALA2 (AP2) is a plant-specific family of transcription factors with a highly conserved DNA-binding domain known as AP2 (Weigel 1995). Rice and *Arabidopsis* genomes are predicted to contain 139 and 122 *AP2* genes, respectively (Nakano et al. 2006). *AP2* family members are implicated in diverse functions such as flower development, spikelet meristem determinacy, plant growth and stress tolerance (Chuck et al. 1998; Liu et al. 1998; Haake et al. 2002; Dubouzet et al. 2003; Gutterson and Reuber 2004). Of these, the role in stress response has been relatively well characterized (Oh et al. 2009). Recent report (Oh et al. 2009) showed that *OsAP2* increased rice grain yield by 16% to 57% over controls under severe drought conditions when overexpressed under direction of a root-specific promoter *OsCc1*.

A B

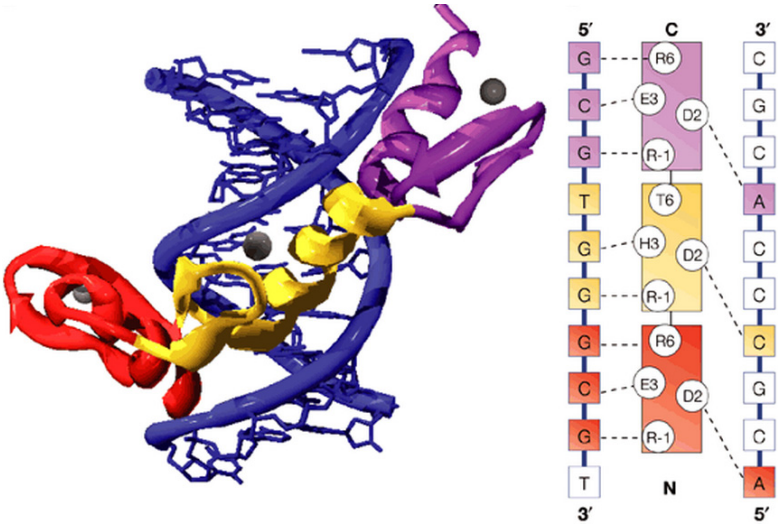


Figure 1. The Zif268–DNA complex showing the sequence-specific protein–DNA interactions. (A) The Zif268–DNA complex showing the three zinc fingers bound in the major groove of DNA5. The DNA is blue and fingers 1, 2, and 3 are red, yellow, and violet respectively. Zinc ions are shown as grey spheres. The Zif268 is a mouse immediate early protein containing zinc finger DNA-binding motif. **(B) A diagram showing the sequence-specific protein–DNA interactions between Zif268 and its DNA-binding site.** The recognition helices of the three fingers are represented in the centre of the panel and the bases on the two strands of the DNA site are shown on either side. The identity of key residues on the recognition helices (positions -1, 2, 3 and 6 with respect to the start of the helix) are also shown using the single-letter code. Contacts observed in the crystal structure are represented as dashed lines. The fingers and bases that they contact are colour-coded using the same scheme as in (A). The fingers are spaced at three-base-pair intervals and tend to contact three adjacent bases on one strand of DNA and one base on the other strand (Figure 1 is taken from Jamieson et al. (2003)).

REFERENCES

Alwin S, Gere MB, Guhl E, Effertz K, Barbas CF III, Segal DJ, Weitzman MD, Cathomen T. (2005) Custom zinc-finger nucleases for use in human cells. Mol Ther. 12: 610-617.

Bae KH, Kwon YD, Shin HC, Hwang MS, Ryu EH, Park KS, Yang HY, Lee DK, Lee Y, Park J, Kwon HS, Kim HW, Yeh BI, Lee HW, Sohn SH, Yoon J, Seol W, Kim JS. (2003) Human zinc fingers as building blocks in the construction of artificial transcription factors. Nat Biotechnol. 21: 275-280.

Bai Y, Pavan S, Zheng Z, Zappel NF, Reinstädler A, Lotti C, de Giovanni C, Ricciardi L, Lindhout P, Visser R, Theres K, Panstruga R. (2008) Naturally occurring broad-spectrum powdery mildew resistance in a Central American tomato accession is caused by loss of *Mlo* function. Mol Plant Microbe Interact. 21: 30-39.

Bibikova M, Golic M, Golic KG, Carroll D. (2002) Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases. Genetics 161: 1169-1175.

Cai CQ, Doyon Y, Ainley WM, Miller JC, Dekelver RC, Moehle EA, Rock JM, Lee YL, Garrison R, Schulenberg L, Blue R, Worden A, Baker L, Faraji F, Zhang L, Holmes MC, Rebar EJ, Collingwood TN, Rubin-Wilson B, Gregory PD, Urnov FD, Petolino JF. (2009) Targeted transgene integration in plant cells using designed zinc finger nucleases. Plant Mol Biol. 69: 699-709.

Carroll D. (2011) Genome engineering with zinc-finger nucleases. Genetics

188: 773-782.

Chames P, Epinat JC, Guillier S, Patin A, Lacroix E, Pâques F. (2005) *In vivo* selection of engineered homing endonucleases using double-strand break induced homologous recombination. Nucleic Acids Res. 33: e178.

Chuck G, Meeley RB, Hake S. (1998) The control of maize spikelet meristem fate by the *APETALA2*-like gene indeterminate spikelet1. Genes Dev. 12: 1145-1154.

Cornu TI, Thibodeau-Beganny S, Guhl E, Alwin S, Eichtinger M, Joung JK, Cathomen T. (2008) DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases. Mol Ther. 16: 352-358.

de Pater S, Neuteboom LW, Pinas JE, Hooykaas PJ, van der Zaal BJ. (2009) ZFN-induced mutagenesis and gene-targeting in *Arabidopsis* through *Agrobacterium*-mediated floral dip transformation. Plant Biotechnol J. 7: 821–835.

Doyon Y, McCammon JM, Miller JC, Faraji F, Ngo C, Katibah GE, Amora R, Hocking TD, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Amacher SL. (2008) Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. Nat Biotechnol. 26: 702-708.

Doyon Y, Vo TD, Mendel MC, Greenberg SG, Wang J, Xia DF, Miller JC, Urnov FD, Gregory PD, Holmes MC. (2011) Enhancing zinc-finger-nuclease activity with improved obligate heterodimer architectures. Nat Methods 8: 74-79.

Dubouzet JG, Sakuma Y, Ito Y, Kasuga M, Dubouzet EG, Miura S, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. (2003) *DREB* genes in rice, *Oryza sativa* L., encode transcription activators that function in drought-,

high-salt- and cold-responsive gene expression. Plant J. 33: 751-763.

Durai S, Mani M, Kandavelou K, Wu J, Porteus MH, Chandrasegaran S. (2005) Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells. Nucleic Acids Res. 33: 5978-5990.

Eisenstein M. (2005) Human gene repair at your fingertips. Nat Methods 2: 405.

Gady AL, Hermans FW, Van de Wal MH, van Loo EN, Visser RG, Bachem CW. (2009) Implementation of two high throughput techniques in a novel application: detecting point mutations in large EMS mutated plant populations. Plant Methods 5: 13.

Gilchrist E, Haughn G. (2010) Reverse genetics techniques: engineering loss and gain of gene function in plants. Brief. Functional Genomics 9: 103-110.

Guo J, Gaj T, Barbas CF III. (2010) Directed evolution of an enhanced and highly efficient *FokI* cleavage domain for zinc finger nucleases. J Mol Biol. 400: 96-107.

Gutterson N, Reuber TL. (2004) Regulation of disease resistance pathways by AP2/ERF transcription factors. Curr Opin Plant Biol. 7: 465-471.

Haake V, Cook D, Riechmann JL, Pineda O, Thomashow MF, Zhang JZ. (2002) Transcription factor CBF4 is a regulator of drought adaptation in *Arabidopsis*. Plant Physiol. 130: 639-648.

Hanin M, Paszkowski J. (2003) Plant genome modification by homologous recombination. Curr Opin Plant Biol. 6: 157-162.

Isalan M, Klug A, Choo Y. (2001) A rapid, generally applicable method to

engineer zinc fingers illustrated by targeting the HIV-1 promoter. Nat Biotechnol. 19: 656-660.

Jamieson AC, Miller JC, Pabo CO. (2003) Drug discovery with engineered zinc-finger proteins. Nat Rev Drug Discov. 2: 361-368.

Joung JK, Ramm EI, Pabo CO. (2000) A bacterial two-hybrid selection system for studying protein-DNA and protein-protein interactions. Proc Natl Acad Sci USA. 97: 7382-7387.

Kim YG, Cha J, Chandrasegaran S. (1996) Hybrid restriction enzymes: zinc finger fusions to *Fok*I cleavage domain. Proc. Natl. Acad. Sci USA 93: 1156-1160.

Kim HJ, Lee HJ, Kim H, Cho SW, Kim JS. (2009) Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. Genome Res. 19: 1279-1288.

Kim S, Lee MJ, Kim H, Kang M, Kim JS. (2011) Preassembled zinc-finger arrays for rapid construction of ZFNs. Nat Methods 8: 7.

Krieger U, Lippman ZB, Zamir D. (2010) The flowering gene *SINGLE FLOWER TRUSS* drives heterosis for yield in tomato. Nat Genetics 42: 459-465.

Liu Q, Kasuga M, Sakuma Y, Abe H, Miura S, Yamaguchi-Shinozaki K, Shinozaki K. (1998) Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in *Arabidopsis*. Plant Cell 10: 1391-1406.

Lloyd A, Plaisier CL, Carroll D, Drews GN. (2005) Targeted mutagenesis using zinc-finger nucleases in *Arabidopsis*. Proc Natl Acad Sci USA.

102: 2232–2237.

Lombardo A, Genovese P, Beausejour CM, Colleoni S, Lee YL, Kim KA, Ando D, Urnov FD, Galli C, Gregory PD, Holmes MC, Naldini L. (2007) Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. Nat Biotechnol. 25: 1298-1306.

Maeder ML, Thibodeau-Beganny S, Osiak A, Wright DA, Anthony RM, Eichtinger M, Jiang T, Foley JE, Winfrey RJ, Townsend JA, Unger-Wallace E, Sander JD, Muller-Lerch F, Fu F, Pearlberg J, Gobel C, Dassie JP, Pruett-Miller SM, Porteus MH, Sgroi DC, Iafrate AJ, Dobbs D, McCray PB Jr, Cathomen T, Voytas DF, Joung JK. (2008) Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. Mol Cell 31: 294-301.

Maeder ML, Thibodeau-Beganny S, Sander JD, Voytas DF, Joung JK. (2009) Oligomerized pool engineering (OPEN): an ‘open-source’ protocol for making customized zinc-finger arrays. Nat Protoc. 4: 1471-1501.

Marton I, Zuker A, Shklarman E, Zeevi V, Tovkach A, Roffe S, Ovadis M, Tzfira T, Vainstein A. (2010) Nontransgenic genome modification in plant cells. Plant Physiol. 154: 1079-1087.

Meng X, Noyes MB, Zhu LJ, Lawson ND, Wolfe SA. (2008) Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. Nat Biotechnol. 26: 695-701.

Meyer M, de Angelis MH, Wurst W, Kühn R. (2010) Gene targeting by homologous recombination in mouse zygotes mediated by zinc-finger nucleases. Proc Natl Acad Sci USA. 107: 15022-15026.

Miller JC, Holmes MC, Wang J, Guschin DY, Lee YL, Rupniewski I, Beausejour CM, Waite AJ, Wang NS, Kim KA, Gregory PD, Pabo CO, Rebar EJ. (2007) An improved zinc-finger nuclease architecture for highly specific geneome cleavage. Nat Biotechnol. 25: 778-785.

Nakano T, Suzuki K, Fujimura T, Shinshi H. (2006) Genome-wide analysis of the *ERF* gene family in *Arabidopsis* and rice. Plant Physiol. 140: 411-432.

Oh SJ, Kim YS, Kwon CW, Park HK, Jeong JS, Kim JK. (2009) Overexpression of the transcription factor AP37 in rice improves grain yield under drought conditions. Plant Physiol. 150: 1368-1379.

Osakabe K, Osakabe Y, Toki S. (2010) Site-directed mutagenesis in *Arabidopsis* using custom-designed zinc finger nucleases. Proc Natl Acad Sci USA. 107: 12034-12039.

Osiak A, Radecke F, Guhl E, Radecke S, Dannemann N, Lütge F, Glage S, Rudolph C, Cantz T, Schwarz K, Heilbronn R, Cathomen T. (2011) Selection-independent generation of gene knockout mouse embryonic stem cells using zinc-finger nucleases. PLoS ONE 6: e28911.

Pabo CO, Peisach E, Grant RA. (2001) Design and selection of novel Cys2His2 zinc finger proteins. Annu Rev Biochem. 70: 313-340.

Papworth M, Kolasinska P, Minczuk M. (2006) Designer zinc-finger proteins and their applications. Gene 366: 27-38.

Pattanayak V, Ramirez CL, Joung JK, Liu DR. (2011) Revealing off-target cleavage specificities of zinc-finger nucleases by *in vitro* selection. Nat Methods. 8: 765-770.

Pavletich NP, Pabo CO. (1991) Zinc finger-DNA recognition: crystal

structure of a Zif268-DNA complex at 2.1 Å. Science 252: 809-817.

Petolino JF, Arnold NL. (2009) Whiskers-mediated maize transformation. Methods Mol Biol. 526: 59-67.

Porteus M. (2007) Creating zinc finger nucleases to manipulate the genome in a site-specific manner using a modular-assembly approach. In “Gene Transfer: Delivery and Expression of DNA and RNA” (ed. Friedmann and Rossi). CSHL Press, Cold Spring Harbor, NY, USA.

Porteus MH. (2009) Zinc fingers on target. Nature 459: 337-338.

Porteus MH, Baltimore D. (2003) Chimeric nucleases stimulate gene targeting in human cells. Science 300: 763.

Pnueli L, Carmel-Goren L, Hareven D, Gutfinger T, Alvarez J, Ganal M, Zamir D, Lifschitz E. (1998) The *SELF-PRUNING* gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of *CEN* and *TFL1*. Development 125: 1979-1989.

Puchta H. (2002) Gene replacement by homologous recombination in plants. Plant Mol Biol. 48: 173-182.

Ramalingam S, Kandavelou K, Rajenderan R, Chandrasegaran S. (2011) Creating designed zinc-finger nucleases with minimal cytotoxicity. J Mol Biol. 405: 630-641.

Sander JD, Dahlborg EJ, Goodwin MJ, Cade L, Zhang F, Cifuentes D, Curtin SJ, Blackburn JS, Thibodeau-Beganny S, Qi Y, Pierick CJ, Hoffman E, Maeder ML,Khayter C, Reyon D, Dobbs D, Langenau DM, Stupar RM, Giraldez AJ, Voytas DF, Peterson RT, Yeh JR, Joung JK. (2011) Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). Nat Methods 8: 67-69.

Segal DJ. (2011) Zinc-finger nucleases transition to the CoDA. Nat Methods 8: 53-55.

Shukla VK, Doyon Y, Miller JC, DeKelver RC, Moehle EA, Worden SE, Mitchell JC, Arnold NL, Gopalan S, Meng X, Choi VM, Rock JM, Wu YY, Katibah GE,Zhifang G, McCaskill D, Simpson MA, Blakeslee B, Greenwalt SA, Butler HJ, Hinkley SJ, Zhang L, Rebar EJ, Gregory PD, Urnov FD. (2009) Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. Nature 459: 437-441.

Söllü C, Pars K, Cornu TI, Thibodeau-Beganny S, Maeder ML, Joung JK, Heilbronn R, Cathomen T. (2010) Autonomous zinc-finger nuclease pairs for targeted chromosomal deletion. Nucleic Acids Res. 38: 8269-8276.

Szczepek M, Brondani V, Büchel J, Serrano L, Segal DJ, Cathomen T. (2007) Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. Nat Biotechnol. 25: 786-793.

Terada R, Johzuka-Hisatomi Y, Saitoh M, Asao H, Iida S. (2007) Gene targeting by homologous recombination as a biotechnological tool for rice functional genomics. Plant Physiol. 144: 846-856.

Thomas KR, Folger KR, Capecchi MR. (1986) High frequency targeting of genes to specific sites in the mammalian genome. Cell 44: 419-428.

Till BJ, Comai L, Henikoff S. (2007) Tilling and ecotilling for crop improvement. In: Tuberosa R. (ed). Genomics-Assisted Crop Improvement. Springer. Netherlands. pp. 333-349.

Townsend JA, Wright DA, Winfrey RJ, Fu F, Maeder ML, Joung JK, Voytas DF. (2009) High-frequency modification of plant genes using

engineered zinc-finger nucleases. Nature 459: 442-445.

Tzfira T, Weinthal D, Marton I, Zeevi V, Zuker A, Vainstein A. (2012) Genome modifications in plant cells by custom-made restriction enzymes. Plant Biotechnol J. 10: 373-389.

Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD. (2010) Genome editing with engineered zinc finger nucleases. Nat Rev Genet. 11: 636-646.

Urnov FD, Miller JC, Lee YL, Beausejour CM, Rock JM, Augustus S, Jamieson AC, Porteus MH, Gregory PD, Holmes MC. (2005) Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature 435: 646-651.

Waltz E. (2012) Tiptoeing around transgenics. Nature Biotech. 30: 215-217.

Weigel D. (1995) The APETALA2 domain is related to a novel type of DNA binding domain. Plant Cell 7: 388-389.

Weinthal D, Tovkach A, Zeevi V, Tzfira T. (2010) Genome editing in plant cells by zinc finger nucleases. Trends Plant Sci. 15: 308-321.

Wilde HD, Chen Y, Jiang P, Bhattacharya A. (2012) Targeted mutation breeding of horticultural plants. Emir. J. Food Agric. 24: 31-41.

Wright DA, Thibodeau-Beganny S, Sander JD, Winfrey RJ, Hirsh AS, Eichtinger M, Fu F, Porteus MH, Dobbs D, Voytas DF, Joung JK. (2006) Standardized reagents and protocols for engineering zinc finger nucleases by modular assembly. Nat Protoc 1: 1637-1652.

Wright DA, Townsend JA, Winfrey RJ Jr, Irwin PA, Rajagopal J, Lonosky PM, Hall BD, Jondle MD, Voytas DF. (2005) High-frequency homologous recombination in plants mediated by zinc-finger nucleases.

Plant J. 44: 693-705.

Wu J, Kandavelou K, Chandrasegaran S. (2007) Custom-designed zinc finger nucleases: what is next? Cell Mol Life Sci. 64: 2933-2944.

Young JJ, Cherone JM, Doyon Y, Ankoudinova I, Faraji FM, Lee AH, Ngo C, Guschin DY, Paschon DE, Miller JC, Zhang L, Rebar EJ, Gregory PD, Urnov FD,Harland RM, Zeitler B. (2011) Efficient targeted gene disruption in the soma and germ line of the frog *Xenopus tropicalis* using engineered zinc-finger nucleases. Proc Natl Acad Sci USA. 108: 7052-7057.

Zhang F, Maeder ML, Unger-Wallace E, Hoshaw JP, Reyon D, Christian M, Li X, Pierick CJ, Dobbs D, Peterson T, Joung JK, Voytas DF. (2010) High frequency targeted mutagenesis in *Arabidopsis thaliana* using zinc finger nucleases. Proc Natl Acad Sci USA. 107: 12028-12033.

CHAPTER I

CONSTRUCTION AND SELECTION OF ACTIVE ZINC FINGER NUCLEASES

1. Introduction

ZFNs have been powerful tools for targeted mutagenesis as well as gene targeting (Carroll 2008 and 2011; Tzfira et al. 2012). The ability of ZFNs to stimulate targeted genome editing is a direct consequence of their ability to introduce site-specific DNA DSBs. Since DNA DSBs are detrimental, the cells have a machinery to repair such chromosomal lesion upon finding. But in the process inadvertent mutations often accompany repair. Targeted mutagenesis using ZFNs takes advantage of such consequences of cellular DSB repair processes. There are two major pathways of DNA DSB repair: nonhomologous end-joining (NHEJ) and homologous recombination (HR). In general, NHEJ is mutagenic, meaning error-prone, while HR is not, because HR makes use of normal sequence information nearby for repair. Thus when one supplies modified template sequences to the vicinity of DSBs, HR may transform itself to a gene targeting tool, replacing original sequence with external modified version. ZFN-driven targeted genome editing exploits these two cellular repair mechanisms, NHEJ and HR. Therefore the only thing ZFNs do is making DNA DSBs at a specified site. As such, designing specific and efficient ZFNs is a key to the targeted genome editing (Carroll 2011; Jamieson et al. 2003; Urnov 2010; Wu et al. 2007).

In this study, I constructed 4-finger ZFNs by modular assembly method using pre-characterized ZF modules. The choice of 4-finger ZFNs over 3-finger ZFNs was based on previous observation in 4-finger ZFNs showing increased specificity and selectivity along with reduce toxicity in human cell lines (Carroll 2008; Kandavelou et al. 2005; Kim et al. 2009;

Urnov et al. 2005). The process of my constructing ZFNs was divided into three parts: i) identifying potential full ZFN target sites using web-based software, ii) construction a library of 4-finger ZFNs, and iii) selecting the most active ZFNs.

For the selecting most active ZFNs, I used the yeast single-strand annealing (SSA) approach as a sole screening method. Yeast SSA assay is based on observation that cleavage at a site between duplicated direct repeats stimulates homologous recombination. This selection approach has proven effective for diverse higher eukaryotic gene targeting or mutagenesis (Bae et al. 2003; Doyon et al. 2008; Townsend et al. 2009). This method has an advantage not only of being *in vivo* assay but also of assaying in a eukaryotic chromosomal context not conceivable in bacterial system such as bacterial two-hybrid (B2H) selection (Joung et al. 2000). From yeast SSA assay, I was able to select four highly active pairs of ZFNs for the target sequence out of 61 pairs. I finally chose a pair of ZFNs that exhibited highest cleavage activity to apply to a rice endogenous target, which I will describe in the next Chapter.

2. Materials and Methods

2.1 Plasmids

The p3 plasmid, a derivative of pcDNA3.0 plasmid (Invitrogen, N.Y. USA), served as a backbone vector that carries one-, two-, or four-zinc finger (ZF) arrays linked to a *Fok*I nuclease domain. The p3 vector contains sequences that encode a hemagglutinin (HA) epitope, a nuclear localization signal (NLS) and a multiple cloning site (MSC) (Appendix 1 and 2). The HA tag facilitates detection and the NLS guides nuclear localization of the expressed ZFNs. The MSC provided cloning sites for both ZF tandem arrays and the *Fok*I nuclease domain. The *Fok*I nuclease domain was derived from *Flavobacterium okeanokoites* genomic DNA (Kim et al. 2010). The p3 backbone vector and single finger-*Fok*I library in p3 vector system are available at ToolGen Inc., Seoul, KOREA. Table 1 shows the list of single-zinc finger ZFN modules and their amino acid sequences and target sequences.

The p415 and p416T, galactose inducible yeast expression vectors, served as backbone vectors of the ZFN expression vectors (Mumberg et al. 1994). The p415, having Amp^R and LEU2 as selection markers, was used to express the 4F-ZFN monomers recognizing left half-sites of the target sequence (termed ZFN-L or left ZFN); the p416T, having Tet^R and URA3 as selection markers, to express the 4F-ZFN monomers recognizing right half-sites of the target sequence (termed ZFN-R or right ZFN) (Appendix 3 and 4). The p3 plasmid carrying 4F-ZFN monomers (named 'RA followed by number') served to provide the inserts for construction of ZFN expression vectors. The inserts encoding 4F-*Fok*I are flanked by restriction

sites *BamH*I and *Xho*I at each ends, which were conveniently exploited for cloning into p415 or p416T. Table 2 shows a list of p3-4F-ZFN pairs which provide 4F-*Fok*I inserts for cloning into yeast expression vectors, p415 and p416T.

The ZFN target *OsAP2* ORF sequence was carried on pGEM-T Easy-AP2-sm (a gift from Dr. Ju-Kon Kim at Myongji University, Yongin) (Figure 2). The PCR-amplified *OsAP2* ORF sequence flanked by *Xab*I and *Xma*I adapters at both ends was previously inserted in pGEM-T Easy vector and this construct was named pGEM-T Easy-AP2-sm.

Plasmids, pUC18-Trp1-HisL-HisR (abbreviated pUC18-H) and pUC18-KanMX-LysL-LysR (abbreviated pUC18-K), were used to isolate linear fragments harboring *OsAP2* ORF. These fragments were then used to transform yeast *Mat* **a** and *Mat* **α** strains, respectively, to yield the reporter strains for single-strand annealing assay (SSA) (Appendix 5 and 6). In pUC18-H and pUC18-K, overlapping truncated copies of the HIS and LYS2 genes, respectively, are interrupted by a ZFN target site (refer to Arnould et al. 2006). Truncated copies of HIS have 494 bp overlap between direct repeats and those of LYS2 have 2,017 bp overlap. Upon cleavage of ZFN target sites, tandem repeat recombination restores functional HIS and LYS2 genes by SSA. The resulting His^+ or Lys^+ cells can be selected for appropriate synthetic medium lacking histidine or lysine (Prieto et al. 2008).

Table 1. Thirty-one single-zinc finger modules

1F code[a]	ZF name[b]	Amino acid sequence	Target subsite[c]
1	DSNRa	YRCKYCDRSFSDSSNLQRHVRNIH	GAC
2	dgnv	FQCRICMRNFSDSGNLRVHIRTH	AAC
3	DSAR2	YSCGICGKSFSDSSAKRRHCILH	GTC
4	mDSCR	YTCSDCGKAFRDKSCLNRHRRTH	GCC
5	ISNR	YRCKYCDRSFSISSNLQRHVRNIH	GAT
7	QSHV	YECDHCGKSFSQSSHLNVHKRTH	YGA
10	mQSNR1	FECKDCGKAFIQKSNLIRHQRTH	GAA
12	QSSR1	YKCPDCGKSFSQSSSLIRHQRTH	GYA
13	mQTHQ	YECHDCGKSFRQSTHLTQHRRIH	AGA
15	RDER2	YHCDWDGCGWKFARSDELTRHYRKH	GYG
16	RDHT	FQCKTCQRKFSRSDHLKTHTRTH	HGG
18	rdnq	FACPECPKRFMRSDNLTQHIKTH	AAG
19	RSHR	YKCMECGKAFNRRSHLTRHQRIH	GGG
21	mVSNV	YECDHCGKAFSVSSNLNVHRRIH	MAT
23	VSTR	YECNYCGKTFSVSSTLIRHQRIH	GCT
24	mWSNR	YRCEECGKAFRWPSNLTRHKRIH	GGT
25	QSHR2	YKCGQCGKFYSQVSHLTRHQKIH	GGA
29	KSNR	YGCHLCGKAFSKSSNLRRHEMIH	GAG
30	QNTQ	YTCSYCGKSFTQSNTLKQHTRIH	ATA
35	vdyk	FHCGYCEKSFSVKDYLTKHIRTH	TAT
36	rdnt	YKCPECGKSFSREDNLHTHQRTH	TAG
37	rdne	YKCPECGKSFSRADNLTEHQRTH	CAG
38	thse	YKCPECGKSFSTSHSLTEHQRTH	CCA
41	tnse	YKCPECGKSFSTKNSLTEHQRTH	CCT
42	hghe	YKCPECGKSFSHTGHLLEHQRTH	CGC
46	QSNT	YECVQCGKGFTQSSNLITHQRVH	AAA
47	QSNV3	YKCDECGKNFTQSSNLIVHKRIH	CAA
48	dghr	YKCPECGKSFSDPGHLVRHQRTH	GGC
49	sadr	YKCPECGKSFSSPADLTRHQRTH	ACA
50	tldr	YKCPECGKSFSTHLDLIRHQRTH	ACT
51	skae	YKCPECGKSFSSKKALTEHQRTH	CAC

[a, b]The single-zinc finger modules used for construction of ZFNs and their target subsites are named as both numeric and letter codes. Each ZF name (letter code) is composed of single-abbreviation of the four amino acid residues at positions -1, 2, 3, and 6 (underlined) in the α-helix of the zinc finger. These amino acid residues are known to make contact with bases in the target DNA subsite (Bae et al. 2003). Each finger module is composed of 23 to 25 amino acids. [c]H= A or C or T; Y= C or T; M=A or C.

Table 2. Sixty-one pairs of p3-4F-ZFNs constructed to target *OsAP2*

4F-left	4F-right	4F-left	4F-right
RA 17	RA 31	RA 21	RA 16
RA 32	RA 54	RA 22	RA 16
RA 41	RA 86	RA 25	RA 82
RA 43	RA 33	RA 26	RA 82
RA 48	RA 58	RA 29	RA 10
RA 50	RA 37	RA 42	RA 87
RA 51	RA 23	RA 45	RA 4
RA 53	RA 18	RA 56	RA 39
RA 57	RA 68	RA 62	RA 47
RA 62	RA 46	RA 64	RA 3
RA 65	RA 95	RA 66	RA 63
RA 69	RA 104	RA 67	RA 96
RA 77	RA 16	RA 73	RA 59
RA 84	RA 10	RA 73	RA 60
RA 91	RA 12	RA 74	RA 98
RA 94	RA 80	RA 78	RA 35
RA 97	RA 79	RA 81	RA 52
RA 105	RA 88	RA 83	RA 30
RA 106	RA 13	RA 84	RA 19
RA 1	RA 13	RA 85	RA 28
RA 2	RA 13	RA 89	RA 38
RA 5`	RA 108	RA 90	RA 61
RA 6	RA 27	RA 91	RA 72
RA 7	RA 24	RA 92	RA 75
RA 8	RA 12	RA 93	RA 49
RA 9	RA 34	RA 94	RA 71
RA 11	RA 44	RA 97	RA 70
RA 14	RA 100	RA 101	RA 40
RA 15	RA 100	RA 102	RA 55
RA 20	RA 36	RA 103	RA 99
		RA 107	RA 76

2.2 Library of single-finger ZFNs

The thirty-one single-finger-*Fok*I modules listed in Table 1 were originally prepared by colleagues as described previously (Kim et al. 2011; Kim et al. 2010). That is, DNA segments, each encoding one of 33 zinc fingers, identified to be frequently found in active ZFNs, had been individually cloned into the p3 vector between the restriction sites, *Xma*I and *Age*I. The restriction enzymes, *Xma*I and *Age*I, recognize distinct 6-base pair DNA sequences with compatible cohesive ends and this facilitates greatly the construction of zinc finger arrays. It makes stepwise construction of zinc finger arrays possible by repetition of a simple subcloning procedure. For easier finger assembly, the *Fok*I nuclease-encoding DNA segment was cloned between *Age*I and *Xho*I sites downstream of zinc finger array in the p3 vector. The resulting 33 zinc finger modules can recognize 39 out of all possible 64 three-base-pair sequences (15 GNN and 24 non-GNN, where N is any base) (Table 1; Kim et al. 2011).

2.3 *E. coli* and yeast strains and cloning

E coli and *S. cerevisiae* strains used include *E coli* DH5α and FYBL2-7B (*MAT***a**, *ura3*Δ*851, trp1*Δ*63, leu2*Δ*1, lys2*Δ*202*) and FYC2-6A (*MAT***α**, *trp1*Δ *63, leu2*Δ*1, his3*Δ*200*). Yeast culture media and conditions were as described in Yeast Protocols Handbook (Clontech, CA, U.S.A) and Gietz and Schiestl (2007).

All subcloning procedures were performed using standard molecular biology techniques (Sambrook and Russel 2001) unless otherwise noted. Plasmid mini-prep and the gel extraction of restriction digestion products

were conducted using kits, GeneAll EXPREP™ Plasmid SV mini and QIAquick Gel Extraction Kit (Quiagen, Valencia, CA, USA) or NucleoSpin Extract II Kit (Macherey-Nagel, Germany), according to the manufacturer's instructions. For PCR cloning, Phusion DNA polymerase (F-530L, Finnzymes, Espoo, Finland) and NucleoSpin Extract II kit were used.

2.4 Construction of a library of 4F-ZFNs

Potential 4-finger ZFN target sites were identified by scanning the open reading frame (ORF) of the rice *OsAP2* gene (GeneBank accession no. AK_069833) using a computer algorithm developed by ToolGen Inc., Seoul, KOREA (http//:www.toolgen.com/ZFNfinder). The computer scan was made for the positions 82-1,167 out of the 1,677 bp *OsAP2* cDNA. The scanned region included majority of the first two exons (1-1,089) and a part of the third (last) exon (1,090-1,167) and excluded intron regions. Two rounds of scan were executed: one for 5 bp long spacer; the other for 6-bp spacer. The spacer refers to the nucleotides lying between the two half-sites recognized by cognate ZFN monomers.

Modular assembly method (Kim et al. 2010) was used to construct the 4-finger zinc finger nucleases (ZFNs). The stepwise assembly includes the two-finger ZFN assembly using single-finger modules, followed by the 4-finger ZFN assembly using the resulting 2-finger modules.

To assemble a library of 2-finger ZFNs, two separate digestions with *Age*I and *Pst*I were subjected to the appropriate pairs of p3-1F-*Fok*I plasmids to obtain the vector (p3-1F) and insert (1F-*Fok*I) fragments. Isolated vector and insert fragments were then ligated to produce a library

of the p3-2F-*Fok*I plasmid constructs, which in turn used to construct a library of 4-finger ZFN plasmid constructs (Figure 6).

Ligation products were used to transform *E. coli* and positive transformants containing correct ligation products were identified by colony PCR using primers: NLS (forward): 5´-CCTCCAAAAAAGAAGAGAAAGGTA-3´ and SP6 (reverse): 5´-CATACGATTTAGGTGACACTATAG-3´. After plasmid mini prep, the authenticity of the ligation products was confirmed by digestion with *Xma*I and *Xho*I and/or sequencing using a primer, CMV-Pro F.

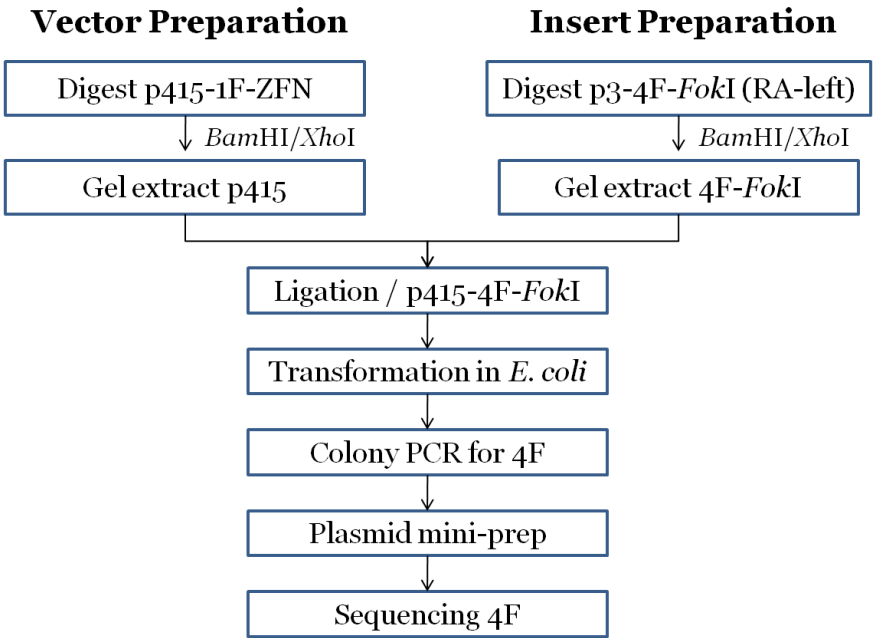
Out of 157 two-finger ZFN modules that were necessary to prepare a full set of 4F-ZFNs, 115 two-finger ZFN modules could be found in the lab stock and I assembled the rest 45 two-finger ZFN arrays (Table 4). The strategy to assemble 4-finger ZFNs was identical to that of 2-finger ZFN assembly (Figure 7).

2.5 Selection of active ZFNs

2.5.1 Construction of ZFN expression vectors

To deliver ZFN pairs inside yeast reporter strain for expression, yeast ZFN expression vectors were constructed based on galactose inducible p415 and p416T vectors (Mumberg et al. 1994). The coding sequences of 4F-ZFN monomers were cloned into galactose inducible yeast expression vectors, p415 and p416T following steps depicted in Figure 1. The p415 yeast expression constructs were made to deliver left ZFN monomers; p416T to deliver right ZFN monomers (Tables 2 and 5). A total of 61 left and 61 right ZFN monomers were cloned into *Bam*HI/*Xho*I sites of p415 and p416T backbone vectors, respectively, according to the strategy depicted in Figure 1.

A


Vector Preparation
Insert Preparation
Digest p415-1F-ZFN
Digest p3-4F-FokI (RA-left)
BamHI/XhoI
BamHI/XhoI
Gel extract p415
Gel extract 4F-FokI
Ligation / p415-4F-FokI
Transformation in E. coli
Colony PCR for 4F
Plasmid mini-prep
Sequencing 4F

B

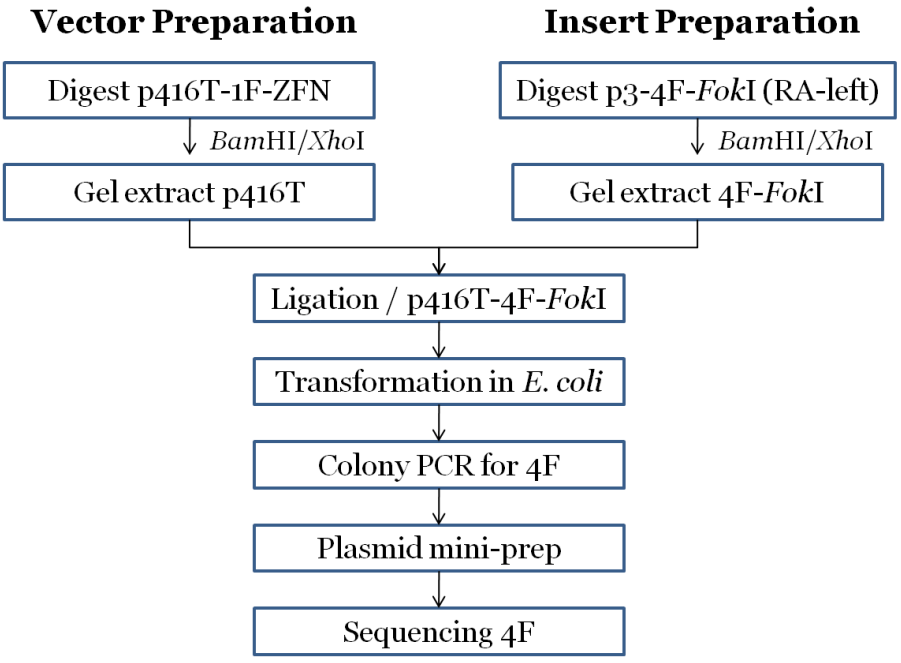

Vector Preparation
Insert Preparation
Digest p416T-1F-ZFN
Digest p3-4F-FokI (RA-left)
BamHI/XhoI
BamHI/XhoI
Gel extract p416T
Gel extract 4F-FokI
Ligation / p416T-4F-FokI
Transformation in E. coli
Colony PCR for 4F
Plasmid mini-prep
Sequencing 4F

Figure 1. Cloning of the 4F-ZFN-L and 4F-ZFN-R into yeast expression vectors, p415 and p416T. The vector and insert fragments were prepared by digesting respective plasmid constructs with *Bam*HI and *Xho*I. Gel extracted vector and insert fragments were subsequently ligated to yield 4-finger ZFN constructs, which were in turn used to transform competent *E. coli*. To identify positive transformants containing desired ligation products, colony PCR was performed using primers: forward (F40): 5′-GCCGAATTGCCTCCAAAAAAGAAGAGAAAGGTAGGGATCC-3′ and reverse (R40): 5′-GTTCAGATTTCTTCTCCTCCAGTTCACTTTTGACTAGTTG-3′. Final confirmation of the authenticity of plasmid constructs was determined by sequence analysis using primers: forward (CMV-ProF): 5′-AACATTTTCGGTTTGTATTACTTC-3′ and reverse (CYC): 5′-GCGTGAATGTAAGCGTGAC-3′. The final products are sets of p415-4F-*Fok*I (left ZFNs) and p415-4F-*Fok*I (right ZFNs). The 'L' in the first sentence indicates targeting left half-site of the target sequence; The 'R' right half-site.

2.5.2 Construction of the yeast reporter strain

To construct yeast reporter strain, a target sequence *OsAP2* coding region was inserted in yeast chromosome by homologous recombination. The target sequence was first cloned into the multiple cloning site of a pUC18 derived vector termed pUC18-H, from which the *Bsi*WI linear fragment was isolated for use to transform yeast *Mat* **a** strain. The homologous recombination between the truncated His repeats of *Bsi*WI fragment and yeast chromosomal HIS gene allowed the target sequence to integrate into yeast chromosome (Figure 4). The target sequence was also cloned into another pUC18 derived vector, pUC18-K. From this construct the *NheI* fragment was used to transform yeast *Mat* **α** strain.

2.5.2.1 Cloning of the target Os*AP2* ORF into pUC vectors

To clone the target sequence into pUC18-H, the target *OsAP2* coding region was PCR amplified using a 1,165 bp *Aat*II/*Nde*I fragment from pGEM-T Easy-AP2-sm as a template DNA. Primers were designed to include *Not*I and *Xma*I adaptors to facilitate cloning into the pUC18-derived plasmid. Primers used for cloning into pUC-H were:

OsAP2-F: 5′-ATATGCGGCCGCATGTGCGGCGGCGCCATCCTCTCCG-3′;

OsAP2-2R: 5′-GTCAACCCGGGTCAGTGAATCGGCCGACACGGGCATG-3′.

The resulting PCR products along with the pUC18 derivative vector were digested with *Not*I and *Xma*I to obtain the vector and insert fragments for ligation, yielding the final cloning products pUC-H-OsAP2 (Figure 3). The successful ligation was confirmed by both colony PCR and *Not*I/*Xma*I digestion followed by gel electrophoresis. Colony PCR was

performed by 2-step protocol using primers:
OsAP2-F : 5′-ATATGCGGCCGCATGTGCGGCGGCGCCATCCTCTCCG-3′ and
OsAP2-2R: 5′-GTCAACCCGGGTCAGTGAATCGGCCGACACGGGCATG-3′.

To clone the target sequence into pUC18-K, I used a different set of primers for amplification of *OsAP2* target sequence. The reason for this is that, while *NheI* digestion is necessary in the later step involving pUC-K-OsAP2 (the ligation product) to isolate a linear fragment for yeast transformation, there is an internal *NheI* site in *OsAP2* ORF. To avoid this problem, the reverse primer was redesigned to exclude *NheI* site in the amplified region of *OsAP2* ORF. Primers used for cloning into pUC-K were:
OsAP2-F: 5′-ATATGCGGCCGCATGTGCGGCGGCGCCATCCTCTCCG-3′;
OsAP2-MR: 5′-TCAACCCGGGAGCATTCCCCATCACAGGAGGTACC-3′.
These primers also include *Not*I and *Xma*I adaptors to facilitate cloning into the pUC18-derived plasmid.

2.5.2.2 Transformation and selection to produce yeast reporter strain

The pUC18-H construct thus obtained was digested with *Bsi*WI and the linear *Bsi*WI fragment containing *OsAP2* coding region was used to transform haploid yeast *Mat* **a** strain, yielding a yeast reporter *Mat* **a** strain (Figure 4). Transformed *Mat* **a** cells tended to grow slower than transformed *Mat* **α** cells. Nonetheless, transformed *Mat* **a** cells were used as a reporter strain for further study, that is, budding yeast single-strand annealing assay described below. Prior to yeast SSA assay, background reversion rate was assessed for the stocks of *Mat* **a** strain (Figure 8).

Spontaneous single-strand annealing (SSA) events (background reversion) were assessed for the stock of haploid yeast *Mat* **a** reporter strains using a combination of serial dilution and selection media (Figure 4). Prior to serial dilution, *Mat* **a** reporter strains were cultured in 2ml liquid SD-W media (lacking tryptophan) overnight. This condition will support growth of only transformed yeast cells ($Trp^{+}His^{-}$ phenotype). Then serially diluted cells from the overnight liquid culture were plated 10μl each per spot on both SD+H (SD media containing histidine) and SD-H (SD media lacking histidine) solid media. The SD+H plate will show viability of transformed yeast cells and the SD-H plate the rate of reversion. From this assay, the stock #14 was finally chosen as a reporter strain for further studies because it exhibited the lowest background reversion rate along with excellent viability.

The construction of reporter *Mat* **α** strain followed the same strategy except the following: *NheI* fragment isolated from pUC-K-OsAP2 contained direct repeats of 2,017 bp-overlapping truncated LYS gene, flanking KanMX selective marker (conferring kanamycin resistance) and an *OsAP2* ORF insert. Inside *Mat* **α** cells, the truncated LYS sequences trigger homologous recombination with endogenous LYS locus, causing integration of *OsAP2* ORF sequence into endogenous LYS locus with concomitant disruption of endogenous LYS gene. After recovering *NheI* fragment-transformed *Mat* **α** cells on G418 (eukaryotic kanamycin) plates, the surviving cells were re-plated on SD-K plates (lacking lysine). The surviving cells here in turn re-plated on G418 plates to identify revertants (refer to Figure 4).

2.5.3 Yeast single-strand annealing assay

This assay was used to evaluate cleavage activity of ZFNs by a principle described in Figure 5. Cleavage activity of putative ZFN pairs were evaluated by co-transforming yeast *Mat* **a** reporter strain (stock #14) with an appropriate combination of ZFN expression vector constructs of p415 and p416T. The yeast expression plasmid constructs guiding ZFN expression (p415-4F-*Fok*I and p416T-4F-*Fok*I) were prepared as described in Section 2.5.1. Co-transformation of left ZFN (p415-4F-*Fok*I) and right ZFN (p416T-4F-*Fok*I) into the *Mat* **a** reporter strain was performed following Gietz and Schiestl (2007). The four ZFN pairs shown in Figure 9 are those exhibited highest cleavage activity when plated on the –L-U-H+gal (lacking leucine, uracil and histidine in galactose media) plates. Prior to serial dilution, the *Mat* **a** reporter strain was cultured overnight in 2ml liquid SD-W media (lacking tryptophan). This condition will support growth of only transformed yeast cells ($Trp^{+}His^{-}$ phenotype). Then serially diluted cells from the overnight liquid culture were plated 10μl each per spot onto agar media containing four different compositions as indicated in the Figure 9. ZFN expression in yeast cells was induced by culturing transformed cells in the presence of galactose as a sole carbon source, instead of glucose. The cleavage efficiency of ZFNs was scored by comparing the number of colonies formed on selective and non-selective plates as well as in the presence or absence of galactose in the media.

ATGtgcggcggcgccatcctctccgacctcatcccgccgccgcggcgggtcaccgccggcgacctctggctggaga
agaccaagaagcagcagcagcagaagaagaagaacaagggcgcgaggaggctgccactgcgccaagaggagga
ggatgatttcgaggccgacttcgaggagttcgaggtggattccggcgagtgggaggtggag**tccgacgccgacgaggc**
caagccgctcgccgcgccccggagcggcttcgctaaaggtggattgaaaaacactactgttgctggtgctgatgggc
ctgcagcaaggtctgctaaaaggaagagaaagaaccaattcaggggtatccgccagcggccatggggcaaatgggc
tgcggaaatcagagatcctcgcaaaggtgtccgcgtctggcttggcaccttcaactctcctgaggaagctgccagagc
ttatgatgctgaagcacgaaggattcgaggcaagaaggccaaggtcaatttcccagatggggctccagtggcttctca
gaggagtcatgctgagccctcctccatgaacatgcctgctttcagcatcgaagagaagccggccgtcatgtcagcagg
caacaaaaccatgtacaacacaaatgcttatgcctaccctgctgttgagtacaccttacaggagccatttgtgcagatt
cagaatgtctcatttgttcctgcaatgaacgcgattgaggatactttcgtgaacctgtcctctgatcaagggagcaactc
ctttggttgctcggactttagccaggagaatgatatcaagacccctgacataacttccatgcttgcaccgaccatgaca
ggtgttgatgactccgcattcctccagaacaatgccagtgatgcaatggtacctcctgtgatggggaatgctagcattg
atcttgctgacctggagccgtacatgaaatttctgatcgatggtggttcggatgagtcgattgacacccttctgagctct
gatggatctcaggatgtggccagtagcatggacctttggagcttcgatgacatgcccgtgtcggccgagttctac**TGA**

Figure 2. *OsAP2* ORF sequence contained in pGEMT-T Easy AP2-sm. The *OsAP2* ORF sequence consists of 1,089 bp. Translation start and stop codons at both ends of the sequence are shown in upper case. The underlined sequence 'tccgacgccgacgaggccaagccgctcgc' (211-239 bp from the start) represents the ZFN 42 binding site with a 5-bp spacer ('gaggc').

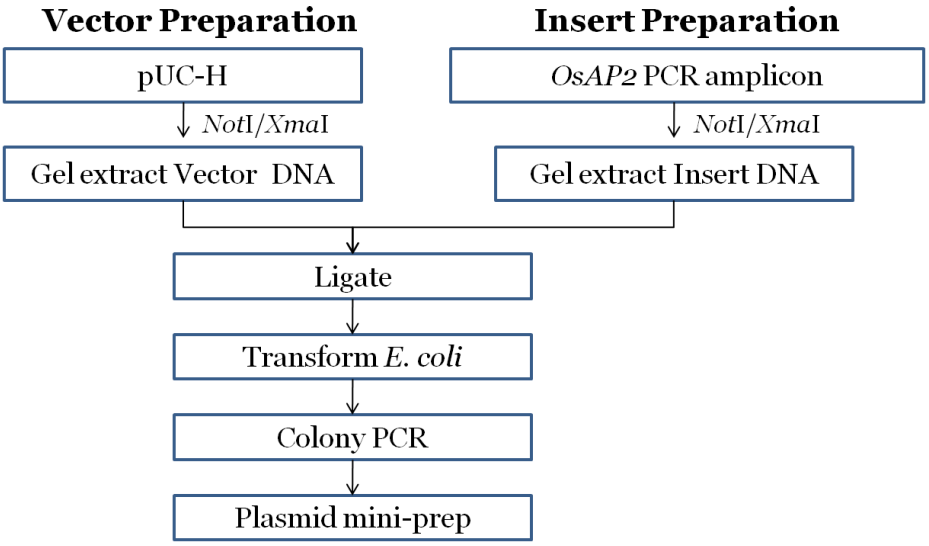


Figure 3. Cloning of the *OsAP2* target sequence into the pUC-H vector. pUC-H and the PCR amplification product spanning *OsAP2* ORF were digested with *Not*I and *Xma*I to obtain vector and insert fragments, respectively. The vector and insert fragments were isolated and ligated together. The ligation products were used to transform *E. coli*. The positive colonies containing pUC-H-OsAP2 were identified by colony PCR using primers:

OsAP2-F: 5′-ATATGCGGCCGCATGTGCGGCGGCGCCATCCTCTCCG-3′

OsAP2-R: 5′-GTCAACCCGGGTCAGTGAATCGGCCGACACGGGCATG-3′

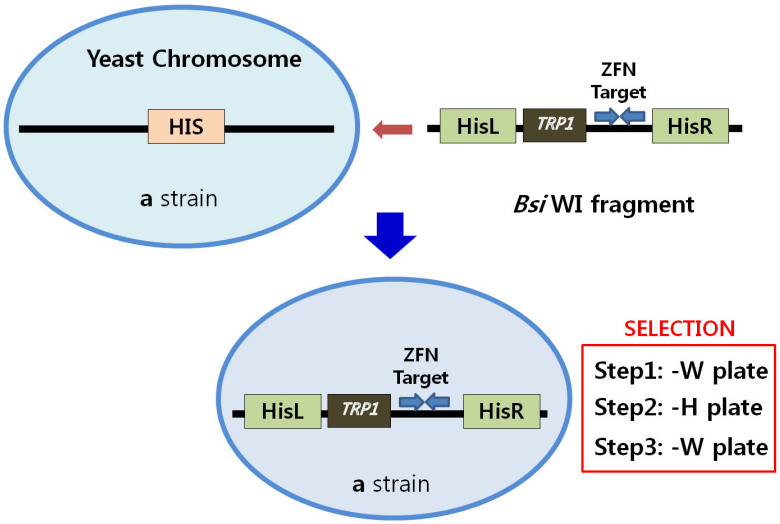


Figure 4. An Overview showing construction of yeast reporter strain and triple-selection used to preclude revertants. A haploid budding yeast strain, FYBL2-7B (*Mat* **a**) was transformed with the 2.5 kb *Bsi*WI linear DNA fragments by the LiAc/SS carrier/DNA/PEG method (Gietz and Schiestl 2007). *Bsi*WI fragment was isolated by digesting pUC-H-OsAP2 with *Bsi*WI and then used to transform *Mat* **a** strain. The *Bsi*WI fragment from pUC-H-OsAP2 contained direct repeats of 494 bp-overlapping coding region of truncated HIS genes, flanking TRP1 and an *OsAP2* ORF insert. Within the *Mat* **a** cell, the truncated HIS repeats take part in homologous recombination with endogenous host HIS gene, yielding HIS^- phenotype. To identify and preclude revertants, stepwise selection strategy was adopted. Transformed *Mat* **a** cells were first plated on SD-W plates (lacking tryptophan) to allow for growth of transformed cells (Trp^-His^+). The surviving cells were then re-plated on SD-H plates (lacking lysine) to identify spontaneous revertants (Trp^-His^+), which occur at a frequency of 1 in 10^{-4} to 10^{-5}. Again, the surviving revertants were re-plated on SD-W plates. The surviving cells here are revertants. The original colonies that went through stepwise selection but couldn't survive on the second SD-W plates were finally selected for use as a reporter strain.

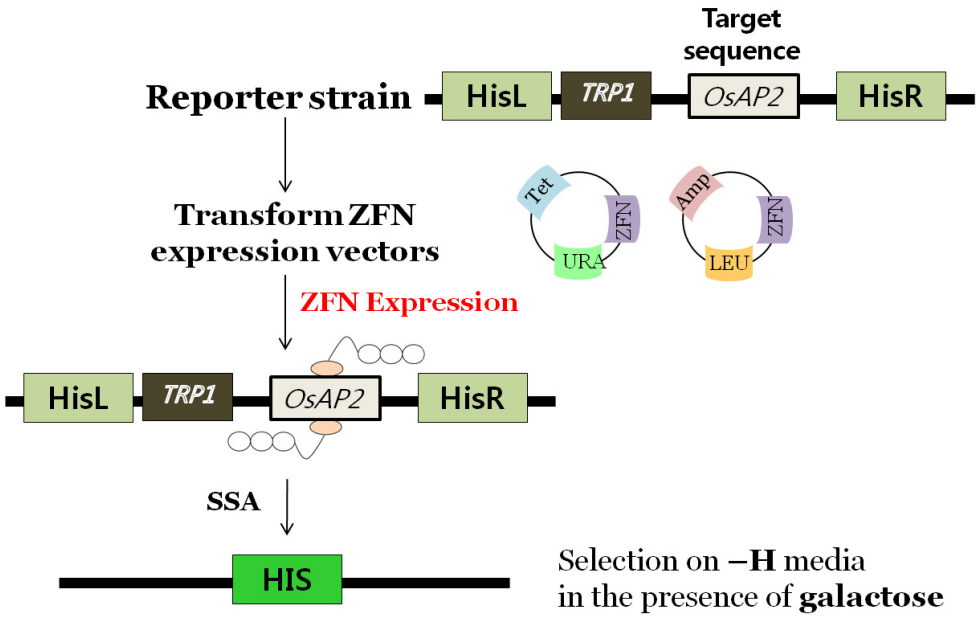


Figure 5. Schematic depicting process of yeast single-strand annealing assay. Reporter strain containing a target sequence (*OsAP2* ORF) flanking truncated His direct repeats was transformed with two distinct yeast expression constructs simultaneously, each directing expression of left-ZFN and right-ZFN monomers. In the presence of galactose, ZFNs are expressed and form dimers at target sites. The cleavage by ZFNs at the target stimulates homologous recombination between the direct repeats of truncated His via a process called single-strand annealing, resulting in recovered HIS phenotype. Thus on SD-H plates (lacking histidine) only cells with active ZFNs can survive.

3. Results

I constructed 4-finger ZFNs that would drive genome editing by generating a site-specific DNA double-strand break at rice *OsAP2* gene. The 4-finger ZFN construction was done by modular assembly method using 31 pre-characterized ZF modules (Table 1). The process of constructing my ZFNs was divided into three parts: i) identifying potential full ZFN target sites using web-based software, ii) construction a library of 4-finger array, iii) selecting most active ZFNs.

3.1 Identifying potential full target cleavage sites

For the design of 4-finger ZFNs, I chose 31 ZF modules that had been characterized in Dr. Jin-Soo Kim's lab at Soul National University and used a computer algorithm (available at http://www.toolgen.com/ZFNfinder) to identify potential ZFN target sites in the DNA sequence of the *OsAP2* coding region. *OsAP2* consists of three exons but the scan was made essentially on the first two exons to maximize mutagenic effect. The actual scan made includes the coding region between about 100 bp position from the translation start to 1 kb position, thus covering a region of approximately 900 bp in length.

Table 1 lists the single-finger zinc finger modules and their cognate target 3-bp sub-sites on which computer scan was based. From this computer scan, a total of 61 potential target cleavage sites were identified (Table 2). Among these sites, 42 sites constitute potential ZFN targets with 5-bp spacers (4F-5bp-4F); 19 sites 6-bp spacers (4F-6bp-4F). On the average, 5-bp spacers were found every 21 bp; 6-bp spacers every 47 bp.

Overall potential ZFN targets were found every 15 bp over the scanned region of the *OsAP2* coding region.

3.2 Construction of a library of 4-finger array

A total of 61 pairs of ZFNs were constructed. The modular compositions of zinc fingers were as indicated in Table 3. To prepare 61 pairs of ZFNs, I had to assemble a total of 108 different 4-finger ZFN monomers as shown in Table 5. To secure these 108 four-finger ZFN monomers, I needed a total of 155 different 2-finger ZFN arrays, of which 110 2-finger ZFN arrays were available from the lab stock. Table 4 shows a list of 2-finger ZFN arrays I made besides ones available from the lab stock.

3.3 Selection of active ZFNs

The computer search results only give a list of potential ZFN pairs and the ZFN pairs with actual cleavage activity have to be identified by one way or another. There are various *in vitro* and *in vivo* approaches available (Rebar and Pabo 1994; Greisman and Pabo 1997; Joung et al. 2000; Bae et al. 2003; Hurt et al. 2003; Bae and Kim 2006; Maeder et al. 2008). But I have chosen an *in vivo* yeast homologous recombination-based approach termed yeast single-strand annealing (SSA) assay (Arnould et al. 2006; Prieto et al. 2008). This approach requires the construction of both yeast ZFN expression vectors and yeast reporter strain.

To evaluate cleavage activity of putative ZFN pairs on the target site, the yeast reporter *Mat* **a** strain was co-transformed with appropriate pairs of p415 and p416T yeast expression constructs carrying ZFN

monomers, according to Gietz and Schiestl (2007).

ZFNs were expressed in the reporter yeast by plating cells on galactose media. As expected, in the presence of galactose (induced condition), cells were able to grow on SD-H plates (devoid of histidine), suggesting cleavage event at the target sequence stimulated homologous recombination between truncated His direct repeats. In contrast, in the absence of galactose (un-induced condition), only spontaneous revertants were survived on SD-H plates, indicating no ZFN expression and thus no cleavage induced.

Cleavage activity of a total of 61 putative ZFN pairs was scored by comparing colony formation of transformed reporter cells as shown in Figure 9. The –L-U plate (lacking both leucine and uracil) selects for only cells containing ZFN pairs. The colony formation on the –L-H plate in the presence of galactose (Top left panel) identifies toxic effect of ZFN expression on the cell. The –L-U plate in the absence of galactose (Top right panel) signifies cell viability in the absence of ZFN expression. The –L-U-H plate in the presence of galactose (Bottom left panel) shows recovered His^+ cells resulting from target cleavage by ZFN pairs, giving a measure of ZFN activity. The –L-U-H plate in the absence of galactose measures background nuclease activity because cells survived in these plates constitute spontaneously recovered His^+ cells in the absence of ZFN expression.

Table 6 shows four highly active ZFNs selected from yeast SSA assay. Their finger arrays on left and right ZFN monomers and target sequences are also shown. The ZFN pair 42 turned out to be the best.

Selected ZFNs are arranged in the order of cleavage activity, the highest at the top and the lowest at the bottom. All 4 selected ZFN pairs have 5-bp spacers.

By combining all the information obtained from the four types of plates as in Figure 9, I have chosen the ZFN pair 42 for further plant application. The ZFN 42 has shown to be the highest in ZFN activity and very low in toxicity. The activity of the ZFN 42 was comparable to the zif 268 which was used as a positive control; but the background cleavage activity of the ZFN 42 was much lower than that of zif 268. Zif 268 is a very well-characterized ZFN showing consistently excellent cleavage activity in diverse organisms and commonly used as a positive control.

Figure 10 depicts relative chromosomal sites on the *OsAP2* gene of highly active ZFN pairs 60, 39, 42, and 20. The cleavage sites of ZFN pairs 60, 39 and 42 reside in the first exon; the ZFN pair 20 in the second exon. The ZFN pair 42 chosen for further studies is marked with a red star. The target DNA sequence and recognizing zinc-finger arrays of the ZFN 42 is depicted in expanded image below ZFN42 target site. The small Arabic numerals in the upper diagram indicate the number of base pairs from the transcription start site. The full length of *OsAP2* genomic DNA is 2,542 bp.

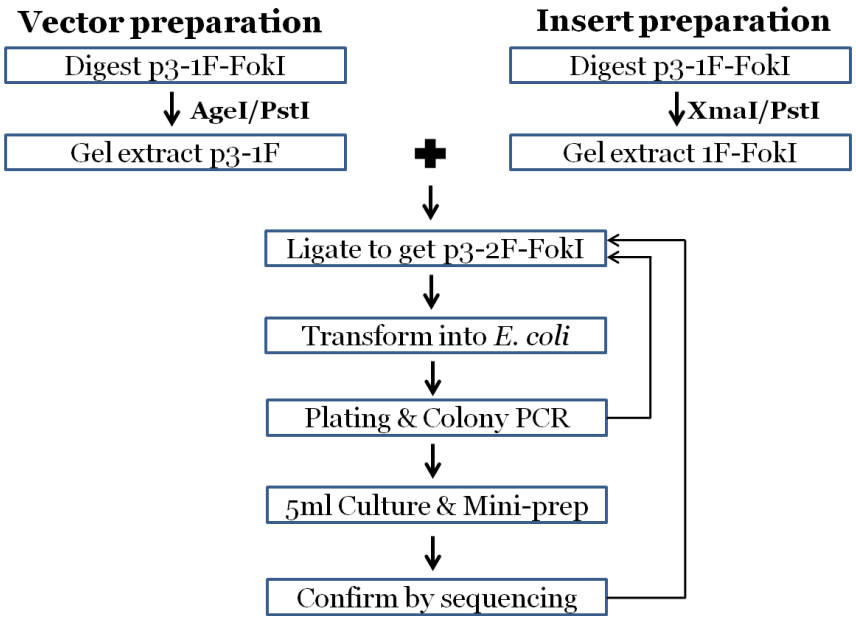
Table 3. Potential 4F-ZFN target sites in rice *OsAP2* cDNA[a]

Seq. No.		4F-6bp-4F								DB: TBS1 / cDNA: 1677bp / ORF: 82-1167
1	98	QSHR2 (25)	mQTHQ (13m	RDHT (16)	mWSNR (24m)	-	dghr (48)	hghe (42)	hghe (42)	hghe (42)
2	101	mQTHQ (13m)	RDHT (16)	mWSNR (24m)	QSHV (7)	-	RSHR (19)	dghr (48)	hghe (42)	hghe (42)
3	114	RSHR (19)	dghr (48)	dghr (48)	hghe (42)	-	QSHV (7)	RDHT (16)	hghe (42)	skae (51)
4	117	dghr (48)	dghr (48)	hghe (42)	hghe (42)	-	tnse (41)	QSHV (7)	RDHT (16)	hghe (42)
5	131	QSHV (7)	RDHT (16)	RDHT (16)	hghe (42)	-	mQTHQ (13m)	mQTHQ (13m)	RDHT (16)	dghr (48)
6	134	RDHT (16)	RDHT (16)	hghe (42)	mWSNR (24m)	-	thse (38)	mQTHQ (13m)	mQTHQ (13m)	RDHT (16)
7	137	RDHT (16)	hghe (42)	mWSNR (24m)	mQTHQ (13m)	-	mQTHQ (13m)	thse (38)	mQTHQ (13m)	mQTHQ (13m)
8	167	VSTR (23)	VSTR (23)	VSTR (23)	VSTR (23)	-	RDHT (16)	sadr (49)	mQTHQ (13m)	mQTHQ (13m)
9	199	hghe (42)	tnse (41)	tnse (41)	rdne (37)	-	KSNR (29)	KSNR (29)	QSNV3 (47)	hghe (42)
10	202	tnse (41)	tnse (41)	rdne (37)	RDHT (16)	-	KSNR (29)	KSNR (29)	KSNR (29)	QSNV3 (47)
11	205	tnse (41)	rdne (37)	RDHT (16)	rdne (37)	-	ISNR (5)	KSNR (29)	KSNR (29)	KSNR (29)
12	208	rdne (37)	RDHT (16)	rdne (37)	RDER2 (15)	-	ISNR (5)	ISNR (5)	KSNR (29)	KSNR (29)
13	502	rdnq (18)	mDSCR (4)	mWSNR (24m)	mQSNR1 (10m)	-	VSTR (23)	mQSNR1 (10m)	KSNR (29)	tnse (41)
14	780	sadr (49)	KSNR (29)	mVSNV (21m)	sadr (49)	-	ISNR (5)	hghe (42)	mQSNR1 (10m)	mVSNV (21m)
15	783	KSNR (29)	mVSNV (21m	sadr (49)	QSHR2 (25)	-	QSHV (7)	ISNR (5)	hghe (42)	mQSNR1 (10m)
16	786	mVSNV (21m)	sadr (49)	QSHR2 (25)	QSSR1 (12)	-	QSHR2 (25)	QSHV (7)	ISNR (5)	hghe (42)
17	827	mWSNR (24m)	sadr (49)	RDHT (16)	rdne (37)	-	tnse (41)	tldr (50)	QSSR1 (12)	QSHR2 (25)
18	855	rdnq (18)	thse (38)	QSNV3 (47)	KSNR (29)	-	mQSNR1 (10m)	QSHR2 (25)	thse (38)	rdnt (36)
19	908	rdnq (18)	RDHT (16)	QSSR1 (12)	QSNV3 (47)	-	RDER2 (15)	rdne (37)	QSHV (7)	thse (38)

Seq. No		4F-5bp-4F								DB: TBS1 / cDNA: 1677bp / ORF: 82-1167
1	100	KSNR (29)	QSHR2 (25)	DSAR2 (3)	KSNR (29)	-	RDER2 (15)	RDER2 (15)	mDSCR (4)	mDSCR (4)
2	103	QSHR2 (25)	DSAR2 (3)	KSNR (29)	ISNR (5)	-	mWSNR (24m)	RDER2 (15)	RDER2 (15)	mDSCR (4)
3	106	DSAR2 (3)	KSNR (29)	ISNR (5)	RDHT (16)	-	skae (51)	mWSNR (24m)	RDER2 (15)	RDER2 (15)
4	109	KSNR (29)	ISNR (5)	RDHT (16)	RDHT (16)	-	hghe (42)	skae (51)	mWSNR (24m)	RDER2 (15)
5	112	ISNR (5)	RDHT (16)	RDHT (16)	RDHT (16)	-	RDHT (16)	hghe (42)	skae (51)	mWSNR (24m)
6	133	mWSNR (24m)	dghr (48)	mDSCR (4)	DSAR2 (3)	-	DSNRa (1)	mQSNR1 (10m)	QSHR2 (25)	VSTR (23)
7	136	dghr (48)	mDSCR (4)	DSAR2 (3)	KSNR (29)	-	QSNV3 (47)	DSNRa (1)	mQSNR1 (10m)	QSHR2 (25)
8	139	mDSCR (4)	DSAR2 (3)	KSNR (29)	thse (38)	-	mQSNR1 (10m)	QSNV3 (47)	DSNRa (1)	mQSNR1 (10m)
9	142	DSAR2 (3)	KSNR (29)	thse (38)	rdne (37)	-	QSSR1 (12)	mQSNR1 (10m)	QSNV3 (47)	DSNRa (1)
10	167	VSTR (23)	VSTR (23)	VSTR (23)	VSTR (23)	-	rdnq (18)	dgnv (2)	rdnq (18)	rdnq (18)
11	196	mDSCR (4)	hghe (42)	tnse (41)	tnse (41)	-	mQTHQ (13m)	thse (38)	RDER2 (15)	tldr (50)
12	199	hghe (42)	tnse (41)	tnse (41)	rdne (37)	-	QSHR2 (25)	mQTHQ (13m)	thse (38)	RDER2 (15)
13	202	tnse (41)	tnse (41)	rdne (37)	RDHT (16)	-	QSHR2 (25)	QSHR2 (25)	mQTHQ (13m)	thse (38)
14	205	tnse (41)	rdne (37)	RDHT (16)	rdne (37)	-	QSHR2 (25)	QSHR2 (25)	QSHR2 (25)	mQTHQ (13m)
15	206	mDSCR (4)	QSSR1 (12)	RDER2 (15)	QSSR1 (12)	-	ISNR (5)	KSNR (29)	KSNR (29)	KSNR (29)
16	208	rdne (37)	RDHT (16)	rdne (37)	RDER2 (15)	-	QSHV (7)	QSHR2 (25)	QSHR2 (25)	QSHR2 (25)
17	209	QSSR1 (12)	RDER2 (15)	QSSR1 (12)	dghr (48)	-	ISNR (5)	ISNR (5)	KSNR (29)	KSNR (29)
18	257	tldr (50)	QSHV (7)	tnse (41)	thse (38)	-	KSNR (29)	RDHT (16)	KSNR (29)	dghr (48)
19	260	QSHV (7)	tnse (41)	thse (38)	mVSNV (21m)	-	RDER2 (15)	KSNR (29)	RDHT (16)	KSNR (29)
20	263	tnse (41)	thse (38)	mVSNV (21m)	RDHT (16)	-	KSNR (29)	RDER2 (15)	KSNR (29)	RDHT (16)
21	284	tnse (41)	thse (38)	tldr (50)	RDHT (16)	-	rdnq (18)	mDSCR (4)	KSNR (29)	DSNRa (1)
22	292	QSHR2 (25)	DSAR2 (3)	dghr (48)	DSAR2 (3)	-	hghe (42)	VSTR (23)	mDSCR (4)	QSNV3 (47)
23	362	rdnt (36)	rdne (37)	QSNV3 (47)	rdne (37)	-	QSSR1 (12)	tnse (41)	RSHR (19)	ISNR (5)
24	365	rdne (37)	QSNV3 (47)	rdne (37)	skae (51)	-	QSSR1 (12)	QSSR1 (12)	tnse (41)	RSHR (19)
25	368	QSNV3 (47)	rdne (37)	skae (51)	rdne (37)	-	RDHT (16)	QSSR1 (12)	QSSR1 (12)	tnse (41)
26	432	qntq (30)	RDHT (16)	RDHT (16)	hghe (42)	-	RSHR (19)	mVSNV (21m)	QSSR1 (12)	RSHR (19)
27	434	QSHR2 (25)	dghr (48)	VSTR (23)	mDSCR (4)	-	VSTR (23)	RDHT (16)	QSNT (46)	dghr (48)
28	434	QSHR2 (25)	dghr (48)	VSTR (23)	mDSCR (4)	-	VSTR (23)	RDHT (16)	QSNV3 (47)	dghr (48)
29	575	tnse (41)	RDHT (16)	tnse (41)	QSHV (7)	-	VSTR (23)	RSHR (19)	ISNR (5)	thse (38)
30	578	RDHT (16)	tnse (41)	QSHV (7)	mVSNV (21m)	-	thse (38)	VSTR (23)	RSHR (19)	ISNR (5)
31	581	tnse (41)	QSHV (7)	mVSNV (21m)	QSHR2 (25)	-	RDER2 (15)	thse (38)	VSTR (23)	RSHR (19)
32	676	RDHT (16)	dghr (48)	DSNRa (1)	mVSNV (21m)	-	dgnv (2)	QSNV3 (47)	QSNV3 (47)	RDHT (16)
33	679	dghr (48)	DSNRa (1)	mVSNV (21m)	QSHV (7)	-	mVSNV (21m)	dgnv (2)	QSNV3 (47)	QSNV3 (47)
34	781	DSNRa (1)	QSHV (7)	QSNT (46)	dgnv (2)	-	ISNR (5)	hghe (42)	mQSNR1 (10m)	mVSNV (21m)
35	781	DSNRa (1)	QSHV (7)	QSNV3 (47)	dgnv (2)	-	ISNR (5)	hghe (42)	mQSNR1 (10m)	mVSNV (21m)
36	784	QSHV (7)	QSNT (46)	dgnv (2)	RDHT (16)	-	QSHV (7)	ISNR (5)	hghe (42)	mQSNR1 (10m)
37	784	QSHV (7)	QSNV3 (47)	dgnv (2)	RDHT (16)	-	QSHV (7)	ISNR (5)	hghe (42)	mQSNR1 (10m)
38	862	QSSR1 (12)	QSHV (7)	DSAR2 (3)	QSNT (46)	-	vdyk (35)	QSHV (7)	mQSNR1 (10m)	QSHR2 (25)
39	862	QSSR1 (12)	QSHV (7)	DSAR2 (3)	QSNV3 (47)	-	vdyk (35)	QSHV (7)	mQSNR1 (10m)	QSHR2 (25)
40	865	QSHV (7)	DSAR2 (3)	QSNT (46)	VSTR (23)	-	QSNV3 (47)	vdyk (35)	QSHV (7)	mQSNR1 (10m)
41	865	QSHV (7)	DSAR2 (3)	QSNV3 (47)	VSTR (23)	-	QSNV3 (47)	vdyk (35)	QSHV (7)	mQSNR1 (10m)
42	965	mVSNV (21m)	RDHT (16)	skae (51)	mVSNV (21m)	-	RDER2 (15)	tnse (41)	tnse (41)	QSSR1 (12)

[a]Potential cleavage sites in the rice *OsAP2* gene were identified by scanning cDNA of *OsAP2* (AK 069833) with a computer algorithm developed by ToolGen Inc. The putative target sites shown in Table 3 include only those found on open reading frame, not on introns. Two separate computer scans were executed: one for ZFN pairs with 5-bp spacers and the other for 6-bp spacers. A total of 61 potential target sites were identified: 19 sites with 6-bp spacer (the upper table) and 42 with 5-bp spacers (the lower table). Sequence numbers shown in the second column indicate the positions of nucleotides that begin 12-bp left finger half-sites. The names of 3-bp sub-sites are shown in both letter codes, composed of four amino acid abbreviations, and alternative numeric codes in parentheses. Flanking the 5-bp or 6-bp spacers in the middle, the four 3-bp subsites on the left are called left half-site; the four 3-bp subsites on the right are called right half-site. The ZFN monomers, 4F-ZFN-L and 4F-ZFN-R recognize left and right half-sites, respectively. The dimers of 4F-ZFN-L and 4F-ZFN-R constitute active ZFN pairs recognizing both half-sites.

A


Vector preparation
Digest p3-1F-FokI
AgeI/PstI
Gel extract p3-1F
Insert preparation
Digest p3-1F-FokI
XmaI/PstI
Gel extract 1F-FokI
Ligate to get p3-2F-FokI
Transform into E. coli
Plating & Colony PCR
5ml Culture & Mini-prep
Confirm by sequencing

B

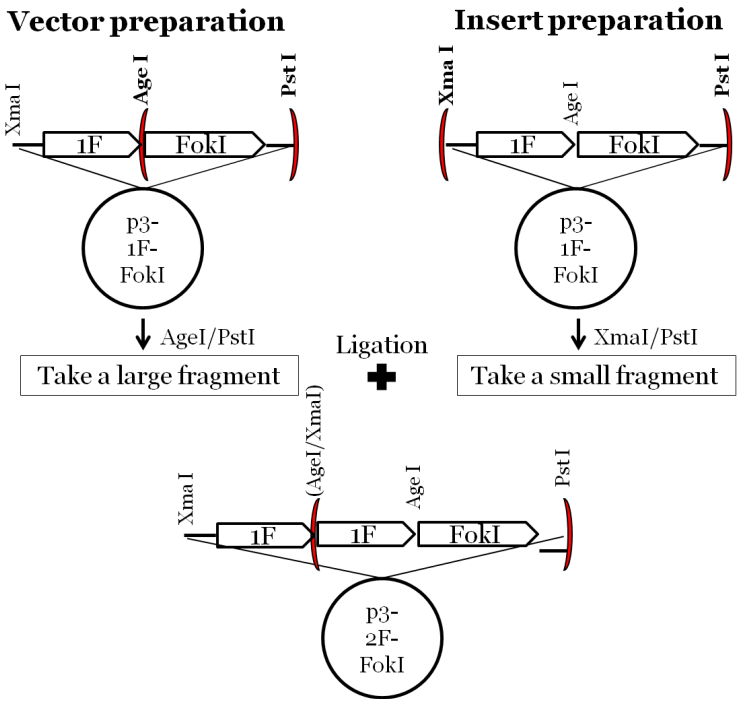

Vector preparation
Xma I
Age I
Pst I
1F
FokI
p3-1F-FokI
AgeI/PstI
Take a large fragment
Insert preparation
Xma I
Age I
Pst I
1F
FokI
p3-1F-FokI
XmaI/PstI
Take a small fragment
Ligation
Xma I
(AgeI/XmaI)
Age I
Pst I
1F
1F
FokI
p3-2F-FokI

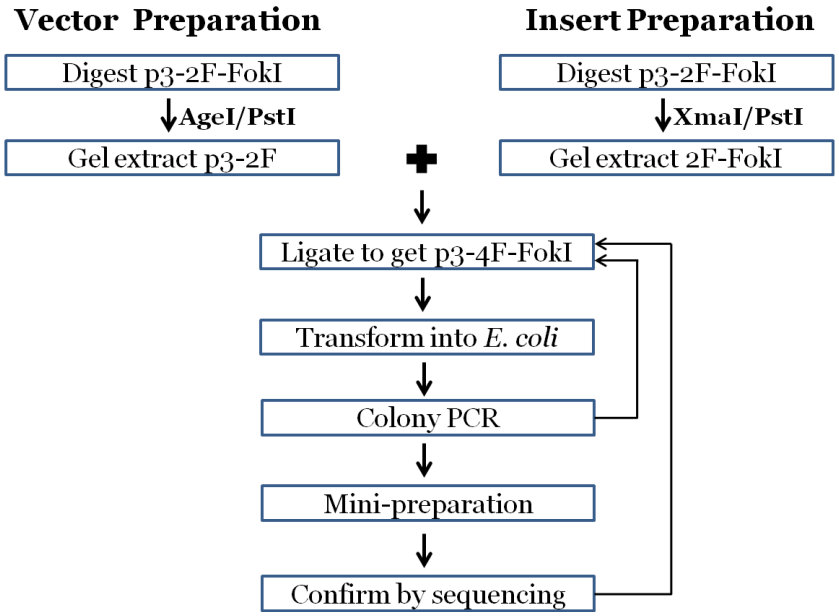
Figure 6. Schematic diagram of the construction of 2-finger ZFN library. **(A)** The vector and insert fragments were prepared from plasmid with single-finger ZFN modules by digesting respective plasmids. The purified vector and insert fragments were used to ligate each other to produce 2-finger ZFN monomers. The ligation reaction mixture was transformed into competent *E. coli* to isolate colonies harboring plasmid with 2-finger ZFNs (p3-2F-*Fok*I). The identification of colonies with desired plasmids was confirmed by colony PCR as described in section 2.4. The confirmed colonies were then inoculated into LB liquid media to prepare plasmid DNA. The final confirmation of the validity of 2-finger ZFNs was made by sequencing the zinc finger regions using primer CMV-Pro F. **(B)** *Age*I with recognition sequence A/CCGGT, and *Xma*I, C/CCGGG, produce compatible ends, enabling fragments with such ends to ligate each other. But the ligation product with a novel sequence, A/CCGGG, at the ligation juction no longer functions as a substrate for either restriction enzyme. This nature greatly facilitates stepwise addition of zinc fingers to ZFNs. In vector preparation, larger fragments (equivalent to p3 fused with a single zinc finger) were excised from the agarose gel; in insert preparation, smaller fragments (equivalent to a single-finger fused with *Fok*I domain) were excised. The purified vector and insert fragments are then ligated to produce p3 fused with 2-finger ZFNs.

Table 4. List of forty-five 2-finger ZFN modules assembled *de novo*.

F1[a]	F2[b]	F1	F2	F1	F2
3[c] DSAR2[d]	47[c] QSNV3[d]	19 RSHR	48 dghr	42 hghe	51 skae
4 DSCR	42 hghe	21 VSNV	49 sadr	46 QSNT	23 VSTR
4 DSCR	47 QSNV3	23 VSTR	4 DSCR	46 QSNT	48 dghr
5 ISNR	38 thse	24 WSNR	13 QTHQ	47 QSNV3	1 DSNRa
5 ISNR	42 hghe	24 WSNR	48 dghr	47 QSNV3	37 rdne
7 QSHV	10 QSNR1	25 QSHR2	3 DSAR2	47 QSNV3	42 hghe
7 QSHV	41 tnse	25 QSHR2	23 VSTR	47 QSNV3	47 QSNV3
10 QSNR1	21 VSNV	29 KSNR	1 DSNRa	48 dghr	1 DSNRa
10 QSNR1	47 QSNV3	29 KSNR	48 dghr	48 dghr	3 DSAR2
18 rdnq	2 dgnv	38 thse	36 rdnt	48 dghr	42 hghe
18 rdnq	4 DSCR	41 tnse	12 QSSR1	48 dghr	48 dghr
18 rdnq	16 RDHT	41 tnse	50 tldr	49 sadr	25 QSHR2
18 rdnq	18 rdnq	42 hghe	10 QSNR1	49 sadr	29 KSNR
18 rdnq	38 thse	42 hghe	23 VSTR	51 skae	21 VSNV
19 rdnq	21 VSNV	42 hghe	42 hghe	51 skae	24 WSNR

[a]F1 represents the first zinc finger starting from the N-terminus; [b]F2 the second zinc finger. Each zinc finger is named either [c]numeric code or [d]letter code. The letter codes, consisting of four-amino acid abbreviations, represent four amino acid residues at positions -1, 2, 3, and 6 in the α-helix of the zinc finger.

A


Vector Preparation
Insert Preparation
Digest p3-2F-FokI
AgeI/PstI
Gel extract p3-2F
Digest p3-2F-FokI
XmaI/PstI
Gel extract 2F-FokI
Ligate to get p3-4F-FokI
Transform into E. coli
Colony PCR
Mini-preparation
Confirm by sequencing

B

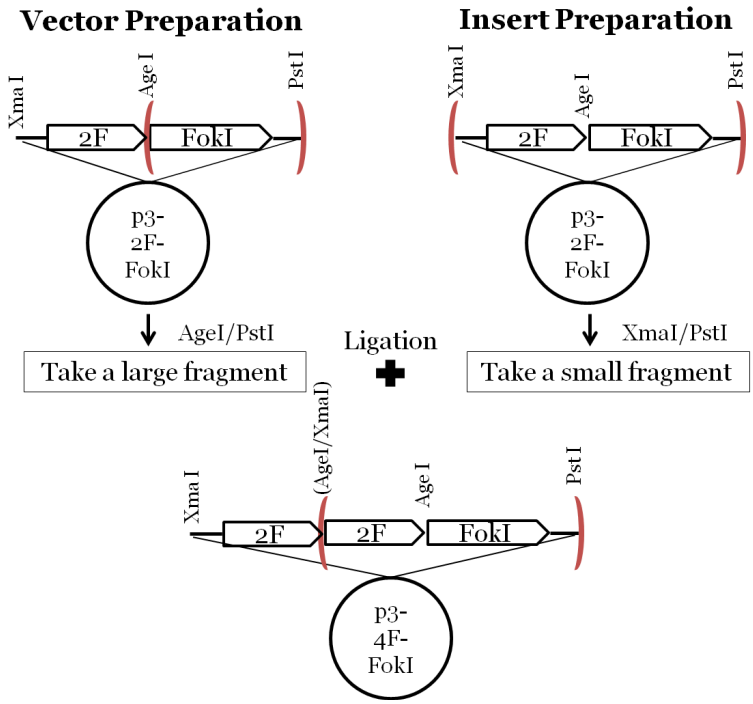

Vector Preparation
Insert Preparation
Xma I
Age I
Pst I
2F
FokI
p3-
2F-
FokI
AgeI/PstI
Take a large fragment
Ligation
XmaI/PstI
Take a small fragment
(AgeI/XmaI)
2F
p3-
4F-
FokI

Figure 7. Schematic diagram of the construction of 4-finger ZFN library. **(A)** The vector and insert fragments were prepared from plasmid with two-finger ZFNs by digesting respective plasmids. The purified vector and insert fragments were then ligated each other to form 4-finger ZFN monomers. The ligation reaction mixtures were transformed into *E. coli* to isolate colonies harboring plasmid with 4-finger ZFNs. The confirmation of this was done by colony PCR as described in section 2.4. The confirmed colonies were then inoculated into LB liquid media to harvest plasmids. The final confirmation of the identity of 4-finger ZFNs was made by sequencing the zinc finger regions.

(B) *Age*I with recognition sequence A/CCGGT, and *Xma*I, C/CCGGG, produce compatible ends, enabling fragments with such ends to ligate each other. But the ligation product with a novel sequence, A/CCGGG, at the ligation juction no longer functions as a substrate for either restriction enzyme *Age*I or *Xma*I. This nature greatly facilitates stepwise addition of zinc fingers to ZFNs. For vector preparation, larger fragments (equivalent to p3 fused with a two zinc fingers) were excised from the agarose gel; for insert preparation, smaller fragments (equivalent to a 2-finger fused with *Fok*I domain) were excised. The purified vector and insert fragments are then ligated to produce p3 fused with 4-finger ZFNs.

Table 5. Four-finger array ZFN library[a]

4F Array ID	Vector		Insert	
RA 1	1	7	46	2
RA 2	1	7	47	2
RA 3	1	10	25	23
RA 4	2	47	47	16
RA 5	3	29	5	16
RA 6	3	29	38	37
RA 7	4	3	29	38
RA 8	4	12	15	12
RA 9	4	42	41	41
RA 10	5	5	29	29
RA 11	5	16	16	16
RA 12	5	29	29	29*
RA 13	5	42	10	21
RA 14	7	3	46	23
RA 15	7	3	47	23
RA 16	7	5	42	10
RA 17	7	16	16	42
RA 18	7	16	42	51
RA 19	7	25	25	25
RA 20	7	41	38	21
RA 21	7	46	2	16
RA 22	7	47	2	16
RA 23	10	25	38	36
RA 24	10	47	1	10
RA 25	12	7	3	46
RA 26	12	7	3	47
RA 27	12	10	47	1
RA 28	12	12	41	19
RA 29	12	15	12	48
RA 30	12	41	19	5
RA 31	13	13	16	48
RA 32	13	16	24	7
RA 33	13	38	13	13
RA 34	13	38	15	50
RA 35	15	15	4	4
RA 36	15	29	16	29
RA 37	15	37	7	38
RA 38	15	38	23	19
RA 39	15	41	41	12
RA 40	16	12	12	41
RA 41	16	16	42	24
RA 42	16	41	7	21
RA 43	16	42	24	13
RA 44	16	42	51	24
RA 45	16	48	1	21
RA 46	16	49	13	13
RA 47	18	2	18	18
RA 48	18	4	24	10
RA 49	18	4	29	1
RA 50	18	16	12	47
RA 51	18	38	47	29
RA 52	19	21	12	19
RA 53	19	48	48	42
RA 54	19	48	42	42
RA 55	21	2	47	47
RA 56	21	16	51	21
RA 57	21	49	25	12
RA 58	23	10	29	41
RA 59	23	16	46	48
RA 60	23	16	47	48
RA 61	23	19	5	38
RA 62	23	23	23	23
RA 63	24	15	15	4
RA 64	24	48	4	3
RA 65	24	49	16	37
RA 66	25	3	29	5
RA 67	25	3	48	3
RA 68	25	7	5	42
RA 69	25	13	16	24
RA 70	25	13	38	15
RA 71	25	25	13	38
RA 72	25	25	25	13
RA 73	25	48	23	4
RA 74	29	5	16	16
RA 75	29	15	29	16
RA 76	29	16	29	48
RA 77	29	21	49	25
RA 78	29	25	3	29
RA 79	29	29	47	42
RA 80	29	29	29	47
RA 81	30	16	16	42
RA 82	35	7	10	25
RA 83	36	37	47	37
RA 84	37	16	37	15
RA 85	37	47	37	51
RA 86	38	13	13	16
RA 87	38	23	19	5
RA 88	41	7	16	42
RA 89	41	7	21	25
RA 90	41	16	41	7
RA 91	41	37	16	37
RA 92	41	38	21	16
RA 93	41	38	50	16
RA 94	41	41	37	16
RA 95	41	50	12	25
RA 96	42	23	4	47
RA 97	42	41	41	37
RA 98	42	51	24	15
RA 99	47	1	10	25
RA 100	47	35	7	10
RA 101	47	37	51	37
RA 102	48	1	21	7
RA 103	48	4	3	29
RA 104	48	42	42	42
RA 105	48	48	42	42
RA 106	49	29	21	49
RA 107	50	7	41	38
RA 108	51	24	15	15

[a]Four-finger array ZFN library assembled to make ZFN pairs to target putative cleavage sites in *OsAP2*. A total of 108 four-finger ZFN monomers, designated RA 1 to RA 108, were constructed via modular assembly method as describe in text. These ZFN monomers recognize ZFN half-sites and contribute to make a total of 61 ZFN pairs that are designed to cleave potential 61 target sites in *OsAP2* ORF. Of the 108 ZFN monomers, 14 were used more than once in making ZFN pairs. The zinc finger arrays that made up ZFN monomers were named by numeric codes. The first two fingers were derived from vector fragments; the last two fingers from insert fragments.

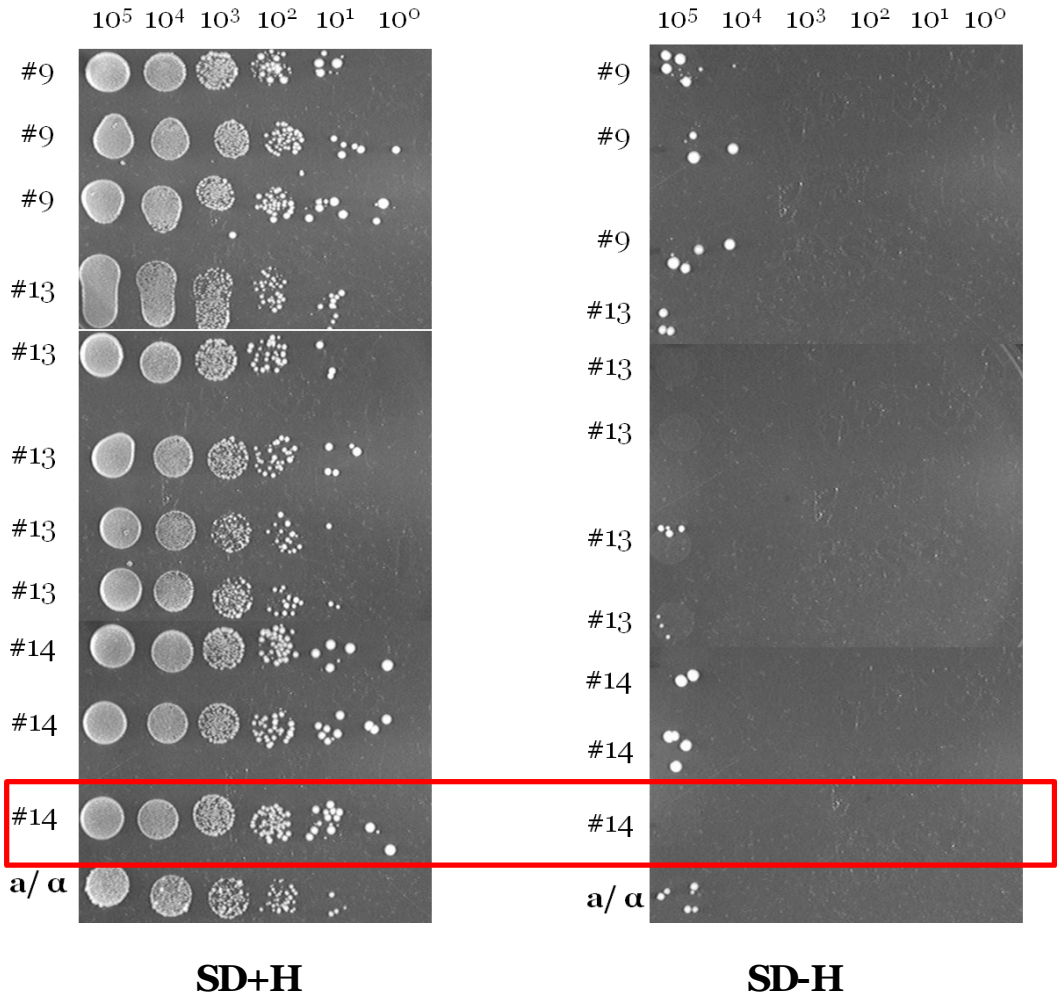


Figure 8. Estimation of the background reversion rate of *Mat* a reporter strains. Spontaneous single-strand annealing (background reversion) rate for a stock of haploid yeast *Mat* **a** reporter strains was estimated using a combination of serial dilution and different selection media. The strain numbers are designated on the left of each photo. The number of yeast cells in each spotting is shown on the top. Each colony formed represents each reversion event.

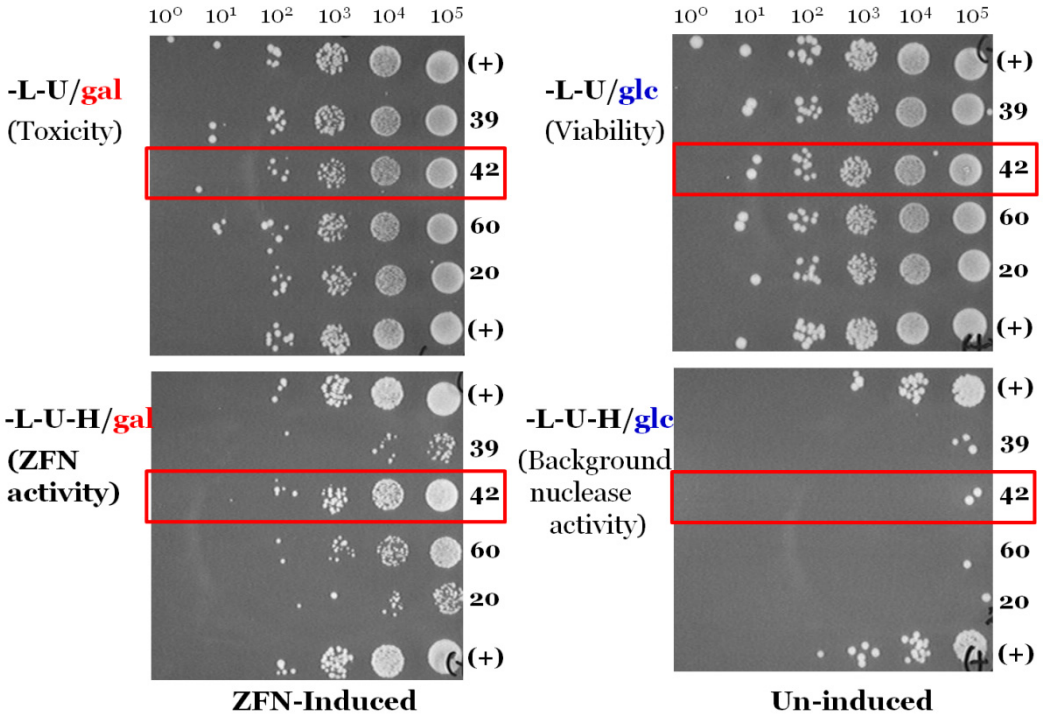


Figure 9. Budding yeast SSA spotting assay showing ZFN pairs with high cleavage activity on the target sequence *OsAP2*. The starting number of cells inoculated in each spot is given on the top. The agar medium composition is shown on the left of each photo and the identity of ZFN pairs on the right. (+): zif 268; 'ZFN-induced': media contain galactose; 'Un-induced': media contain glucose, instead of galactose.

Table 6. Highly active ZFNs for *OsAP2* target selected by budding yeast SSA assay

ZFN ID[a]	Plasmid ID[b]	4F arrangement[c]	Target DNA sequence[d]
42	p415-RA 67	25-3-48-3	5'-TCCGACGCCGAC$(N)_5$CAAGCCGCTCGC-3'
	p416T-RA 96	42-23-4-47	
60	p415-RA 103	48-4-3-29	5'-GCCGGCGACCTC$(N)_5$GGAGAAGACCAA-3'
	p416T-RA 99	47-1-10-25	
21	p415-RA 2	1-7-47-2	5'-GTCTCATTGGTT$(N)_5$AATGAACGCGAT-3'
	p416T-RA 13-3	5-42-10-21	
20	p415-RA 1	1-7-46-2	5'-GTCTCATTGGTT$(N)_5$AATGAACGCGAT-3'
	p416T-RA 13-2	5-42-10-21	

[a]All 4 selected ZFN pairs have 5bp spacers. The four ZFNs shown are arranged in the order of cleavage activity from top to bottom, the highest at the top. ZFN 21 and 20 recognize the same target site and have shown comparable cleavage activity. [b]RA codes in individual ZFN constructs correspond to those in Table 5. [c]The numerical designation of 4-finger arrangements is based on the numerical codes in Table 1. [d]The spacer sequences ($(N)_5$) of the ZFN targets are, from top to bottom, 5′-gaggc-3′, 5′-tggct-3′, 5′-cctgc-3′ and 5′-cctgc-3′, respectively.

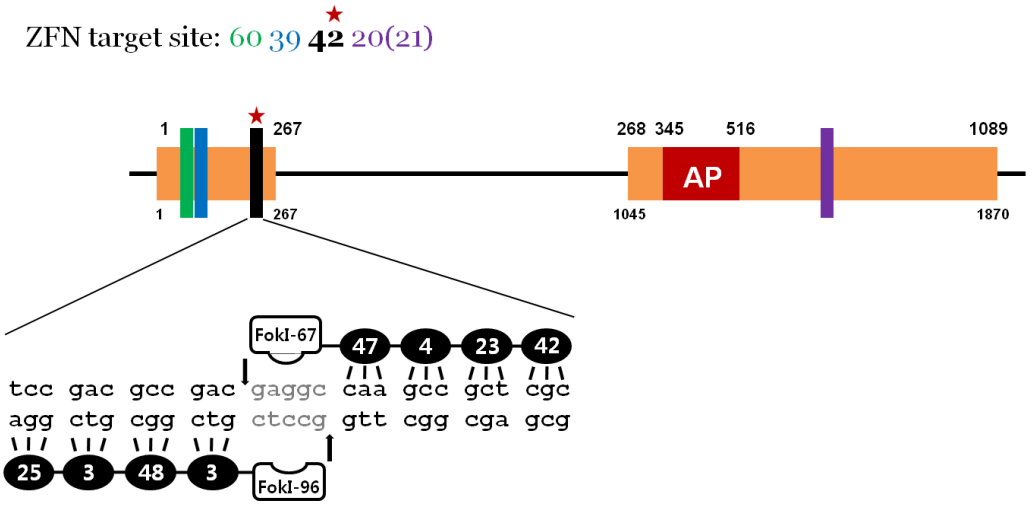


Figure 10. Highly active ZFN pairs selected, 60, 39, 42, and 20, and their respective *OsAP2* genomic sites of targeting. Four highly active ZFN pairs capable of cleaving their respective target sites in *OsAP2* gene were selected based on yeast SSA assay. The cleavage sites of the ZFN pairs, 60, 39 and 42 reside in the first exon; the ZFN pair, 20 in the second exon. Of these four active ZFN pairs, the ZFN pair, 42, marked with a red stars was finally chosen for subsequent transformation experiment. The target DNA sequence and recognizing zinc-finger arrays of the ZFN 42 is depicted in expanded image below ZFN 42 target site. The small Arabic numerals in the upper diagram indicate the number of base pairs from the transcription start site. The full length of *OsAP2* genomic DNA is 2,542 bp.

4. Discussion

In this study, the modular assembly method and the budding yeast-based selection system was used to construct highly active ZFN pairs, at least in yeast (Table 6). Among 61 pairs of ZFNs I constructed, the ZFN pair 42 has shown to be exceptional in terms of the highest cleavage activity, minimal toxicity and negligible spontaneous SSA rate. The cleavage activity of the ZFN 42 was comparable to that of zif 268 but the background reversion rate of the ZFN 42 was much lower than that of zif 268, the well-characterized highly active ZFN that was used as a positive control (Figure 9). In this respect, I would argue the ZFN 42 is superior to zif 268 and could find a use as a positive control. In my assay, zif 268 showed consistently high cleavage activity but it accompanied high background cleavage activity as well. In contrast, the background nuclease acitivity of the ZFN 42 was negligible (Bottom right photo in Figure 9). Repeated pairwise assay for zif 268 and the ZFN 42 consistently gave rise to the same outcome. With such superior characteristics, I dare to say the ZFN 42 has at least a qualification as a use of positive control as has zif 268.

Kim et al. (2009) has pointed out the importance of target site compositions in the design of functional ZFNs. Most early successful studies (Bibikova et al. 2002; Lloyd et al. 2005; Morton et al. 2006) involving ZFN technology indeed focused on a so-called canonical ZF target sequences 5′-NNCNNCNNC$(N)_6$GNNGNNGNN-3′ (N=any nucleotide). It is interesting to note that my four select ZFNs all target the sites with 5-bp spacers instead of 6-bp spacers as in the canonical target. The second

point of note is the significant deviation in base composition from the canonical GNN-repeat sequences. Except the ZFN 42 whose deviation involves only two bases, the rest three select ZFNs have only half or less bases in common with canonical GNN-repeats (the ZFN 60: 4 bp deviated; the ZFN 21 and 20: 5 bp deviated). This suggests that the yeast SSA assay can be used to expand the utility of ZFN technology because the canonical target sequence composed of GNN-repeats occur only rarely in a given gene of interest. GNN-repeat sites for 4-finger ZFNs occur, on average, only once in a 65,536-bp ($=4^8$) sequence. Therefore, it is likely that such sites do not exist in many genes of interest (Kim et al. 2009).

Another interesting finding involves the subsite specificity of one of single-finger ZFN modules used to construct 'ZFN 20', whose 4-finger arrangement for the left ZFN monomer is 1-7-46-2. The single-finger ZF module with code name 46 (1F-module 46) is supposed to recognize a subsite AAA (Table 1). But the 1F-module 46 actually recognized the subsite CAA in my assay with little influence on cleavage activity of the ZFN 20 of which it is a part of (Table 6), compared to the ZFN 21 having the same target sequence (data not shown). This suggests the 1F-module 46 has dual specificity recognizing both AAA and CAA.

At this point, there is no way of telling whether the select ZFNs actually work at the rice endogenous target. Chromosomal context may be different between yeast and rice and also other factors may affect the cleavage efficiency and thus mutagenicity. Various *in vitro* and *in vivo* selection approaches have been tested (Rebar and Pabo 1994; Greisman and Pabo 1997; Joung et al. 2000; Bae et al. 2003; Hurt et al. 2003; Bae

and Kim 2006; Maeder et al. 2008). In general, *in vivo* system works better than *in vitro* system (Kim et al. 2009). But other than that, apparently there is no consensus as to what the best selection approach is. Each system has cons and pros. This reflects, at least in part, the diversity of chromosomal environment in the vicinity of endogenous target sites. Thus the best selection system would be the systems involving the cells of organisms of interest themselves. For instance, if you are interested in targeting a rice endogenous gene, then selecting ZFNs in a rice cellular environment would possibly work best. Townsend et al. (2009) demonstrated that ZFNs that was selected in the tobacco protoplast successfully targeted tobacco endogenous target. Even better would be developing a system that would allow for enrichment of cells with active ZFNs. A mammalian version using surrogate reporter system was reported previously (Kim et al. 2011).

The yeast SSA assay system (Arnould et al. 2006; Prieto et al. 2008) has an advantage of relying on eukaryotic system while retaining microbial easiness. In fact, growing number of researchers are relying on yeast selection system (Doyon et al. 2008; Townsend et al. 2009; Zhang et al. 2010). There are different versions of strategies of conducting yeast SSA assay (Chames et al. 2005; Doyon et al. 2008; Prieto et al. 2008; Townsend et al. 2009). Initially, I tried to construct the reporter strain containing two chromosomal reporters without success. There is plasmid reporter system which I didn't try but it appears to be easier and quicker than chromosomal reporter system.

To identify potential full 4-finger ZFN target sites, cDNA sequence

of the rice *OsAP2* gene (AK069833) was scanned by a computer algorithm developed by ToolGen Inc. (available at http://www.toolgen.com/ZFNfinder). This step is necessary because zinc finger modules cannot recognize all potential sequences. Currently available 'pools' of zinc finger modules are capable of recognizing only subset of all possible three-base-pair 'sub-sites' at each finger position (Maeder et al. 2009; Kim et al. 2011). Thus as progress in improvements in ZF modules are made, so the scanned results will be.

In designing ZFNs, it is important to consider three aspects of ZFNs. They are affinity, specificity and cytotoxicity. Affinity and specificity to a target site sequence is largely determined by zinc finger moiety, whereas toxicity in general results from off-targeting led by homodimerization of *Fok*I moiety. Many strategies have been attempted to enhance affinity and specificity while reducing toxicity. One of successful approaches is attaching additional zinc fingers to ZFN (Carroll 2011; Kim et al. 2009; Urnov et al. 2005), which led me to make 4-finger ZFNs rather than 3-finger ZFNs.

Modular assembly method has an advantage of straightforwardness because it involves simple joining of single fingers with pre-characterized specificities (Maeder et al. 2008) but it is criticized that it ignores subtle effects of context, often resulting in high failure rates (Ramirez et al. 2008). For this reason, several improved assembly methods including OPEN (Maeder et al. 2008; Maeder et al. 2009) and CoDA (Sander et al. 2011; Segal 2011) were devised that take account of context effects. Unfortunately, at the time when I started my ZFN experiment, only

modular assembly method was publically available. Recently, Kim et al. (2011) devised preassembled zinc-finger arrays that could greatly accelerate the construction of ZFNs. In ZFN technology, construction of ZFNs is the rate-determining step and it is hoped that great improvements in designing and engineering highly active ZFNs would be seen in the foreseeable future. Partly due to keen competition for earlier publication, most reports concerning engineered nuclease-mediated genome editing has remained proof-of-concept in nature. To enable this valued technology to find more practical applications, comparative studies testing different approaches in design and construction of ZFNs will be desired in the future.

5. References

Arnould S, Chames P, Perez C, Lacroix E, Duclert A, Epinat JC, Stricher F, Petit AS, Patin A, Guillier S, Rolland S, Prieto J, Blanco FJ, Bravo J, Montoya G,Serrano L, Duchateau P, Pâques F. (2006) Engineering of large numbers of highly specific homing endonucleases that induce recombination on novel DNA targets. J. Mol. Biol. 355: 443-458.

Bae KH, Kwon YD, Shin HC, Hwang MS, Ryu EH, Park KS, Yang HY, Lee DK, Lee Y, Park J, Kwon HS, Kim HW, Yeh BI, Lee HW, Sohn SH, Yoon J, Seol W, Kim JS. (2003) Human zinc fingers as building blocks in the construction of artificial transcription factors. Nat Biotechnol. 21: 275-280.

Bibikova M, Golic M, Golic KG, Carroll D. (2002) Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. Genetics 161: 1169-1175.

Carroll D. (2008) Progress and prospects: Zinc-finger nucleases as gene therapy agents. Gene Ther 15: 1463-1468.

Carroll D. (2011) Genome engineering with zinc-finger nucleases. Genetics 188: 773-782.

Chames P, Epinat JC, Guillier S, Patin A, Lacroix E, Pâques F. (2005) *In vivo* selection of engineered homing endonucleases using double-strand break induced homologous recombination. Nucleic Acids Res. 33: e178.

Doyon Y, McCammon JM, Miller JC, Faraji F, Ngo C, Katibah GE, Amora R, Hocking TD, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Amacher SL. (2008) Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. Nat Biotechnol. 26: 702-708.

Gietz RD, Schiestl RH. (2007) High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat Protoc. 2: 31-34.

Greisman HA, Pabo CO. (1997) A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target sites. Science 275: 657-661.

Jamieson AC, Miller JC, Pabo CO. (2003) Drug discovery with engineered zinc-finger proteins. Nat Rev Drug Discov. 2: 361-368.

Joung JK, Ramm EI, Pabo CO. (2000) A bacterial two-hybrid selection system for studying protein-DNA and protein-protein interactions. Proc Natl Acad Sci USA. 97: 7382-7387.

Kandavelou K, Mani M, Durai S, Chandrasegaran S. (2005) “Magic” scissors for genome surgery. Nat Biotechnol. 23: 686-687.

Kim H., Um E., Cho S.R, Jung C., Kim H., Kim J.S. (2011) Surrogate reporters for enrichment of cells with nuclease-induced mutations. Nat Methods 8: 941-943.

Kim HJ, Lee HJ, Kim H, Cho SW, Kim JS. (2009) Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. Genome Res. 19: 1279-1288.

Kim S, Kim EJ, Kim JS. (2010) Construction of combinatorial libraries that encode zinc finger-based transcription factors. Methods Mol Biol. 649: 133-147.

Kim S, Lee MJ, Kim H, Kang M, Kim JS. (2011) Preassembled zinc-finger arrays for rapid construction of ZFNs. Nat Methods 8: 7.

Lloyd A, Plaisier CL, Carroll D, Drews GN. (2005) Targeted mutagenesis using zinc-finger nucleases in Arabidopsis. Proc Natl Acad Sci USA. 102: 2232-2237.

Maeder ML, Thibodeau-Beganny S, Osiak A, Wright DA, Anthony RM, Eichtinger M, Jiang T, Foley JE, Winfrey RJ, Townsend JA, Unger-Wallace E, Sander JD,Müller-Lerch F, Fu F, Pearlberg J, Göbel C, Dassie JP, Pruett-Miller SM, Porteus MH, Sgroi DC, Iafrate AJ, Dobbs D, McCray PB Jr, Cathomen T, Voytas DF, Joung JK. (2008) Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. Mol Cell 31: 294-301.

Maeder ML, Thibodeau-Beganny S, Sander JD, Voytas DF, Joung JK. (2009) Oligomerized pool engineering (OPEN): an 'open-source' protocol for making customized zinc-finger arrays. Nat Protoc. 4: 1471-1501.

Morton J, Davis MW, Jorgensen EM, Carroll D. (2006) Induction and repair of zinc-finger nuclease-targeted double-strand breaks in Caenorhabditis elegans somatic cells. Proc Natl Acad Sci USA. 103: 16370-16375.

Mumberg, D, Müller R, Funk M. (1994) Regulatable promoters of *Saccharomyces cerevisiae*: comparison of transcriptional activity and their use for heterologous expression. Nucleic Acids Res. 22: 5767-5768.

Prieto J, Epinat JC, Redondo P, Ramos E, Padró D, Cédrone F, Montoya G, Pâques F, Blanco FJ. (2008) Generation and analysis of mesophilic variants of the thermostable archaeal I-DmoI homing endonuclease. J Biol Chem. 283: 4364-4374.

Porteus, MH. (2006) Mammalian Gene Targeting with Designed Zinc Finger Nucleases. Mol Ther. 13: 438-446.

Ramirez CL, Foley JE, Wright DA, Müller-Lerch F, Rahman SH, Cornu TI, Winfrey RJ, Sander JD, Fu F, Townsend JA,Cathomen T, Voytas DF,

Joung JK. (2008) Unexpected failure rates for modular assembly of engineered zinc fingers. Nat Methods 5: 374-375.

Rebar EJ, Pabo CO. (1994) Zinc finger phage: affinity selection of fingers with new DNA-binding specificities. Science 263: 671-673.

Sambrook J, Russell DW. (2001) Molecular cloning: a laboratory manual. 3rd ed. Cold Spring Harbor Press, N.Y.

Sander JD, Dahlborg EJ, Goodwin MJ, Cade L, Zhang F, Cifuentes D, Curtin SJ, Blackburn JS, Thibodeau-Beganny S, Qi Y, Pierick CJ, Hoffman E, Maeder ML,Khayter C, Reyon D, Dobbs D, Langenau DM, Stupar RM, Giraldez AJ, Voytas DF, Peterson RT, Yeh JR, Joung JK. (2011) Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). Nat Methods 8: 67-69.

Segal DJ. (2011) Zinc-finger nucleases transition to the CoDA. Nat Methods 8: 53-55.

Townsend JA, Wright DA, Winfrey RJ, Fu F, Maeder ML, Joung JK, Voytas DF. (2009) High-frequency modification of plant genes using engineered zinc-finger nucleases. Nature 459: 442-445.

Tzfira T, Weinthal D, Marton I, Zeevi V, Zuker A, Vainstein A. (2012) Genome modifications in plant cells by custom-made restriction enzymes. Plant Biotechnol J. 10: 373-389.

Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD. (2010) Genome editing with engineered zinc finger nucleases. Nat Rev Genet. 11: 636-646.

Urnov FD, Miller JC, Lee YL, Beausejour CM, Rock JM, Augustus S, Jamieson AC, Porteus MH, Gregory PD, Holmes MC. (2005) Highly

efficient endogenous human gene correction using designed zinc-finger nucleases. Nature 435: 646-651.

Wu J, Kandavelou K, Chandrasegaran S. (2007) Custom-designed zinc finger nucleases: what is next? Cell Mol Life Sci. 64: 2933-2944.

Zhang F, Maeder ML, Unger-Wallace E, Hoshaw JP, Reyon D, Christian M, Li X, Pierick CJ, Dobbs D, Peterson T, Joung JK, Voytas DF. (2010) High frequency targeted mutagenesis in *Arabidopsis thaliana* using zinc finger nucleases. Proc Natl Acad Sci USA. 107: 12028-12033.

CHAPTER II

APPLICATION OF ZINC FINGER NUCLEASE TECHNOLOGY TO RICE (*Oryza sativa* L.)

1. Introduction

Efficient targeted mutagenesis has been a long sought-after goal among plant breeders because it can allow them to bypass arduous and time-consuming series of backcross and subsequent selection processes. The ability of zinc finger nuclease (ZFN) to make a double-strand break (DSB) at a user-specified genomic site has offered great promise in realizing such goal.

Zinc-finger nucleases (ZFNs) are hybrids between a nonspecific DNA-cleavage domain and a DNA-binding domain composed of Cys_2His_2 zinc fingers. Because the array of zinc finger proteins can be manipulated to recognize a broad range of sequences, these enzymes are capable of directing cleavage and thus sequence modification to arbitrarily chosen targets (Biblikova et al. 2002).

Since the first targeted mutagenesis *in vivo* at a chromosomal site in higher eukaryotes (Biblikova et al. 2002), ZFN-directed genome editing has demonstrated remarkable successes (Baker 2012a; Carroll 2008 and 2011). Nonetheless, only a handful of reports so far deal with plant applications of ZFNs. The plant species subjected to ZFN-mediated genome editing include *Arabidopsis*, tobacco, maize, soybean and petunia (Cai et al. 2009; Curtin et al. 2011; de Pater et al. 2009; Even-Faitelson 2011; Lloyd et al. 2005; Marton et al. 2010; Osakabe et al. 2010; Petolino et al. 2010; Shukla et al. 2009; Tovkach et al. 2009; Townsend et al. 2009; Tzfira et al. 2012; Wright et al. 2005; Zhang et al. 2010). Among them maize is the only species of the monocotyledonous category where all the important grain crops belong.

With this in mind, I have decided to apply ZFN technology to rice, a monocotyledonous model plant in molecular biology. The monocotyledonous plant species has long been intractable for genetic manipulation. But rice is an exception, is amenable to *Agrobacterium*-mediated transformation and has a small genome size, about three times as that of *Arabidopsis*. The *OsAP2* gene has been chosen as a target gene. *OsAP2* is a transcription factor containing AP2-domain and has shown to increase grain yield when over-expressed by a root-specific promoter particularly under drought conditions (Oh et al. 2009). There is no knock-down or knock-out mutants available yet.

With an eventual goal of obtaining *OsAP2* knock-out mutants, I assembled a library of 4-finger ZFNs and selected a pair of 4-finger ZFNs that showed the highest cleavage activity as measured by yeast SSA assay (refer to Chapter I). This chosen ZFN pair was delivered to rice embryonic calli via *Agrobacterium* to generate transgenic plants expressing ZFNs.

Here I describe the analyses of 100 transgenic plants thus obtained and provide an evidence of targeted mutagenesis occurring at the desired target. Further considerations for the future studies will be discussed.

2. Materials and Methods

2.1 Plant material and growth conditions

Both transgenic and non-transgenic rice (*Orysa sativa* L. ssp. japonica cv. Ilmi) plants were grown in 16 h light 8 h dark cycle at 28-28℃ in the greenhouse. The seeds of non-transgenic (NT) rice plants that were used as a negative control were chemical sterilized and germinated on Murashige-Skoog (MS) plates at 28℃ in a growth chamber. During the 4 days of germination period, the light was turned off for the first three days before turning on the fourth day. The young seedlings, both non-transgenic and transgenic, were transplanted individually in a pot (4x4x5 cm^3) filled with rice nursery soil (Bio-media, Kyeongju, KOREA) and placed in the greenhouse year-round.

For DNA or protein extraction, appropriate amounts of leaf or root tissues were cut from individual plants and dumped immediately into liquid nitrogen before transferred to a -80℃ freezer in the lab. Leaf or root tissue powder was prepared by grinding frozen samples with a pestle and mortar in the presence of liquid nitrogen. The sample powder was divided and stored in the 1.5-ml E-tubes at -80℃ before use. Whenever possible, T1 seeds were collected and saved for later genotyping.

2.2. DNA constructs

A ZFN pair, ZFN 42 (Table 6 and Figure 10 in Chapter I) was chosen to induce targeted mutagenesis in rice *OsAP2* because of its high cleavage activity combined with low toxicity and low background cleavage activity

(Figure 9 in Chapter I). The ZFN 42 has a finger arrangement of 25-3-48-3 for the left ZFN monomer and 42-23-4-47 for the right ZFN monomer. The *Fok* I nuclease used includes wild-type (F67/F96) and obligate heterodimeric variants (F67-RR/F96-DAS and F67-DAS/F96-RR) (Guo et al. 2010; Appendix 7). Zinc finger array-*Fok* I fusions that recognize left and right half-sites of the target sequence were expressed in plant cells from a single plasmid by placing between them an in-frame 20-amino acid ribosome skipping signal from the foot-and-mouth disease virus (Ha et al. 2010). Each zinc finger array is preceded by the SV40 nuclear localization signal (NLS) and the HA tag (Appendix 2). The zinc finger arrays, wild-type *Fok* I nuclease and 2A sequences were codon-optimized for expression in rice. The codon-optimized sequence including 'HA tag - NLS - left ZFN - 2A - HA tag - NLS - right ZFN' was synthesized at Entelechon GmBH, Germany (Appendix 8).

Expression plasmids were constructed by cloning the synthesized sequence into plant expression vector, pSB11 (Komari et al. 1996). As a Gateway-compatible destination vector, pSB11 can recombine with an entry vector that contains 'HA tag - NLS - left ZFN - 2A - HA tag - NLS - right ZFN'. To enable the overexpression of the ZFN protein in rice, the ZFN gene was linked to the *PDG1* promoter for constitutive expression, and the *RCc3* promoter for root specific expression using the Gateway system (Invitrogen, Carlsbad, CA). The rice transformation vector contained two marker genes in tandem, adjacent to the left border (BL) of T-DNA region (Figure 2): the GFP gene driven by stress-inducible *Wsi18* promoter (Yi et al. 2010) for visual screening; the *bar* gene under *OcCc1* promoter (Jang et

al. 2002) for herbicide-based selection. All PCR amplifications, sequencing reactions, DNA digestions and agarose gel electrophoresis were carried out according to standard protocols (Sambrook and Russell 2001).

2.3 Transgenic plants

Callus induction, co-cultivation with *Agrobacterium*, and the selection of transformed calli were carried out as previously described (Jang et al. 1999; Hiei et al. 1994), except for adding 7 mg/l and 4 mg/l of phosphinothricin to the selection and the regeneration medium, respectively (Jang et al. 2002). Plasmids were introduced into *Agrobacterium tumefaciens* LBA4404 by tri-parental mating and embryogenic calli from mature seeds were transformed as previously described (Jang et al. 1999). Overall about 2,000 seed scutellum-derived calli (2~5 mm in diameter) were co-cultivated with *Agrobacterium* carrying Ti plasmid during the 2-year period. Through the course of about 70 days of tissue culture from *Agrobacterium*-infected calli, a total of 100 transformed rice plants (T_0 plants) were regenerated (Table 1).

2.4 Analysis of gene targeting events

2.4.1 Genomic PCR for ZFN integration

Genomic DNA was isolated from rice tissue using a DNA purification kit (GeneAll Exgene^{TM} Plant SV mini, GeneAll Biotechnology Co., Seoul). PCR analysis for evaluation of ZFN integration in rice genome was carried out using forward primer 5′-GGGGACAAGTTTGTACAAAAAAGCAGGCT-3′

and reverse primer 5′-GGGGACCACTTTGTACAAGAAAGCTGGGT-3′. *Taq* DNA polymerase (Cat. No. 501-025) was purchased from GeneAll Biotechnology, Seoul, KOREA.

2.4.2 Detection of GFP fluorescence

GFP assay on dry T_1 seeds was performed to assess ZFN expression indirectly using a research stereomicroscope (LAS-3000, Olympus, Tokyo, Japan). Images were captured with exposure time between 60 to 210 seconds using a C5060-ZOOM digital camera (Olympus, Tokyo) under blue light using a 460–480 nm excitation filter (Park et al. 2010).

2.4.3 Immunoblot analysis

Western blot analysis of ZFN expression was carried out on the leaf and root proteins using Hoefer SDS-PAGE kit (Hoefer Pharmacia Biotech Inc., San Francisco, CA; Amersham Pharmacia Biotech Inc.) and ECL Semi-dry Blotter (Amersham Biosciences, Piscataway, NJ), according to the manufacturer's instruction. Polyclonal antibody against HA-tag was purchased from Santa Cruz Biotechnology Inc., Santa Cruz, U.S.A.

2.4.4 T7 endonuclease assay

To evaluate ZFN-induced mutagenesis for the targeted site in *OsAP2*, a mismatch-based assay was performed using T7 endonuclease (T7E1) as described (Figure 1; Kim et al. 2011). The primers used for PCR amplification of the ZFN target were forward primer 5′-CCATCCTCTCCGACCTCATC-3′ and reverse primer

5′-CTAATCTCCCCACTGACCGC-3′. The primers were designed to amplify about 300 bp region spanning the ZFN target site.

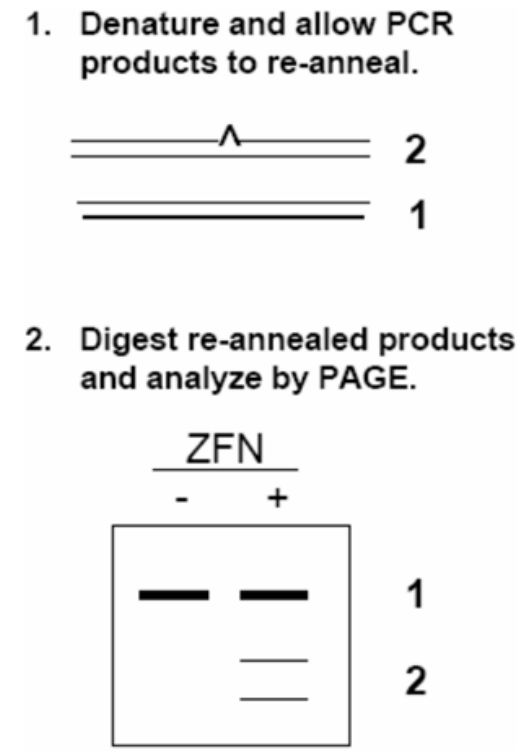


Figure 1. Principle of mismatch selective endonuclease assay for measuring ZFN-mediated gene disruption. The level of ZFN-induced mutations is quantified by PCR amplification of the ZFN target region from genomic DNA, followed by denaturing and allowing wild type and mutant alleles to re-anneal together to create hetero-duplexes. The re-annealed PCR products are then digested with the mismatch-selective enzyme such as T7EI or Surveyor nuclease (Transgenomic) that preferentially cuts DNA at sites of duplex distortions (Figure 1 is taken from Perez et al. 2008).

3. Results

3.1. Regeneration of transgenic plants

By delivering the 6 different vector constructs (Figure 2), a total of 100 transgenic plants were obtained (Table 1). The individual T_0 plants are listed in the Appendix 9, where the transplanting dates, the harvested seeds and the results of genomic PCR and GFP assays are indicated.

Marked difference was observed between ZFNs with wild-type (F67/F96) and variant *Fok*I domains (F67-DAS/F96-RR or F67-RR/F96DAS) in the rate of plant regeneration from calli, particularly among those under the control of the constitutive promoter PGD1 (Table 1). From calli containing PGD1::ZFN with wild-type *Fok*I, only a single plant was recovered from a total of 720 calli, the regeneration rate being only 0.14%. Interestingly, among calli containing RCc3::ZFN, no significant difference between wild-type and variant *Fok*I domains was detectable in the rate of regeneration from transformed calli. The regeneration rate of the calli containing RCc3::ZFN ranged from 7.2% to 8.3%.

Initially, only wild-type *Fok*I-ZFN was delivered to embryogenic calli to recover transgenic plants without knowing any possible adverse effect on the regeneration capability. After experiencing such difficulty, variant *Fok*I-ZFN was solely used for later co-cultivations.

The appearance of regenerated rice plants was rather diverse as is typical to transgenic plants due to somaclonal variation. But when cancelling out the variations among plants containing the same plasmid constructs no significant phenotypic difference could be seen among transgenic plants expressing different types and levels of ZFNs. This was

true not only among transgenic plants but also between non-transgenic and transgenic rice plants (Data not shown).

T_1 seeds were harvested from only 15 T_0 plants that were early regenerants. Since majority of transgenic plants were still in vegetative stage it would be too early to tell any possible effect of *OsAP2* targeting on the fertility of transgenic rice plants. Nonetheless, it appeared that most transgenic plants looked fertile but with varying degrees.

3.2. ZFN Integration and expression

3.2.1. Genomic PCR

To assess ZFN integration in the rice genome genomic PCR analysis was performed using primers that were designed to amplify a 300 bp genomic region spanning 'HA tag - NLS - left ZFN - 2A - HA tag - NLS - right ZFN'. The genomic PCR results are indicated in the sixth column of Appendix 9. Interestingly, not all transgenic plants show positive PCR signals. Even after repeated attempts with varying amount of template DNA, 20% (19 out of 98) of transgenic samples showed no expected ZFN band. The depletion of template DNA in the PCR reaction mixture was not responsible for the negative results (data not shown). There was one plant with partial internal deletion in ZFN insert detected in this analysis.

3.2.2. GFP assay

In the expression plasmid (Figure 2), just downstream of the ZFN coding region lies a GFP gene as a screenable reporter gene. GFP gene is under

the control of stress (e.g. water-stress) responsive promoter *Wsi18* (Yi et al. 2010). Taking advantage of this, GFP assay was conducted for the harvested dry seeds. As shown in Figure 3, all the regenerated rice plants assayed had a capacity to express ZFNs as suggested by green fluorescence. The lack of fluorescence in RCc3::ZFN-7 is due to internal deletion in the integrated ZFN, a result in agreement with that in genomic PCR (data not shown).

3.2.3. Immunoblot analysis

To ensure the ZFN expression vectors were fully functional, immunoblot analysis was conducted using polyclonal anti-HA antibody as a primary antibody. As shown in Figure 4, all transgenic plants analyzed expressed high levels of ZFN proteins. In contrast, NT (non-transgenic) plant did not show any visible band corresponding to ZFN protein.

3.3. Analysis of genome editing events

The ability of the ZFN 42 to induce site-specific mutations was determined by a mismatch-based assay using mismatch-selective T7 endonuclease (T7E1). The principle is described in Figure 1.

A stretch of 300 bp region encompassing ZFN target site in *OsAP2* gene was PCR amplified using forward primer 5′-CCATCCTCTCCGACCTCATC-3′ and reverse primer 5′-CTAATCTCCCCACTGACCGC-3′. After denaturing and re-annealing PCR products allowing hetero-duplex formation, the resulting re-annealed PCR products were subjected to digestion with T7E1. As indicated in Figure 5,

hetero-duplex formation between wild-type and mutant DNA strands in the target region was identified in transgenic lines PGD1::ZFNRD-4, 5, 6, 7 and 8. The faint lower bands correspond to the expected size of T7E1 digestion products. The reason for lower bands being faint is that, in T_0 transgenic plants, ZFN-induced mutant cells are distributed as somatic mosaics with one or two mutant cells scattered among intact 100 or so cells. ZFN-induced somatic mutations in T_0 transgenic plants typically give rise to such somatic mosaics.

To further evaluate mutagenic events at the target site, the PCR amplified region was cloned from genomic DNA of PGD1::ZFNRD-6 and subjected to sequence analysis. Of 160 clones for which sequence information was available, 4 clones exhibited substitution mutations at the target site although no indel (insertion or deletion) was found (Figure 6). The total number of mutant bases were six, the mutation frequency well above the error rate of *Taq* polymerase (1 in 9,000 or 2.28×10^{-5}), suggesting the occurrence of targeted mutagenesis events at the *OsAP2* gene.

To see if the point mutations accompany amino acid substitutions, amino acid sequences were deduced as shown in Figure 7. Among the four clones that have substitution mutations at the ZFN target, amino acid substitutions were found in the three clones (the clones 180, 104 and 268). The clone 225 exhibited silent mutation (Pro→Pro), while the clone 180 conservative mutation (Ala→Val). The clone 104 is worthy of note because it has three amino acid subtitutions within a span of 17 bp, which is very unlikely to happen spontaneously.

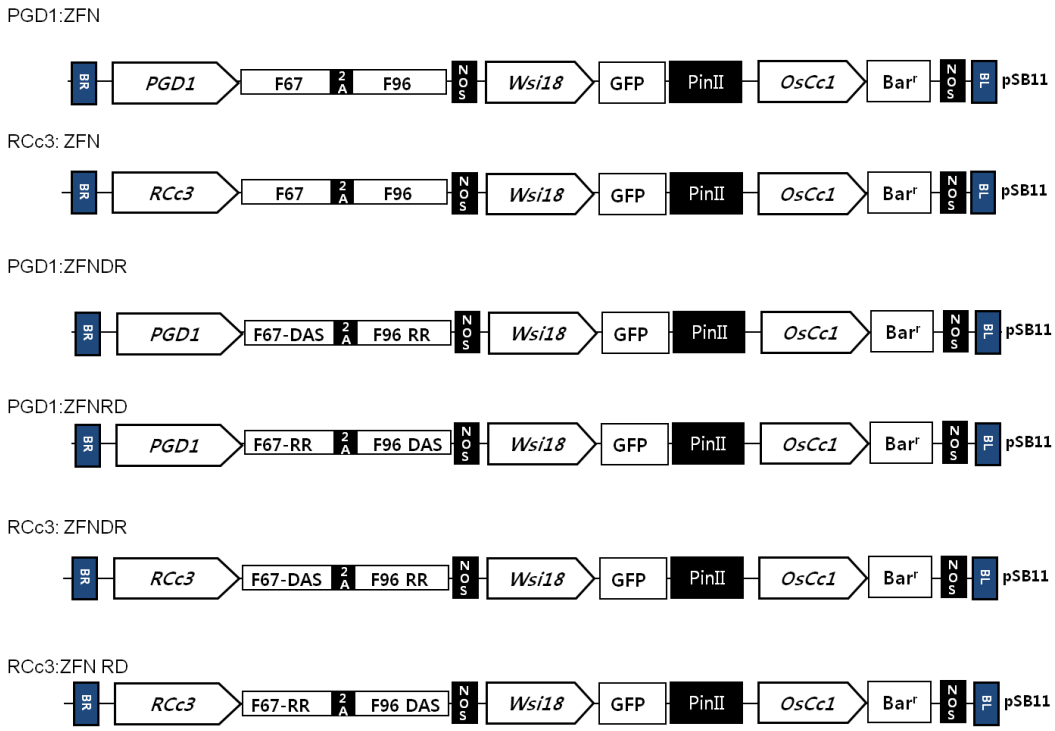


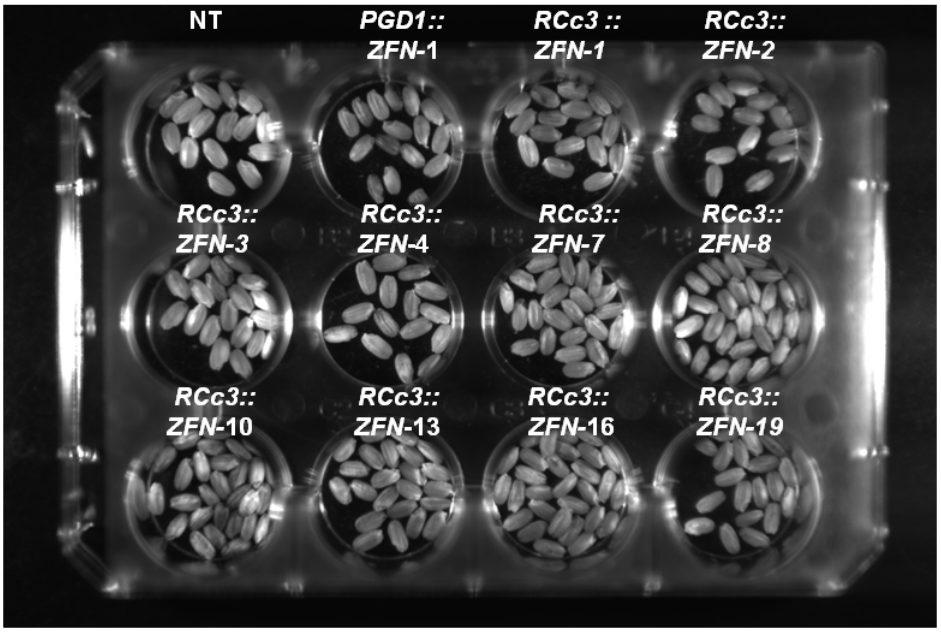
Figure 2. Ti-plasmid vector constructs for ZFN transformation. Two types of promoters were used to drive ZFN expression: *PGD1*, a constitutive promoter and *RCc3*, a root-specific promoter. Bar^r acts as a selectable marker and is under the control of a constitutive promoter, *OsCc1*. Also included is a visible marker GFP which is under the stress (dryness)-inducible promoter *Wsi18*. Nos and PinII are terminators. BR stands for right border; BL left border that are joined to the Ti plasmid pSB11. To enable the two ZFN monomers to express within a single plasmid the 2A sequence was incorporated. The 2A sequences flanking ZFN monomers allow the separation of the proteins by a ribosomal skipping mechanism, which leads to impairment of a peptide bond without stopping the translation

Table 1. Transgenic rice plants regenerated from ZFN transformed calli

Vector type	*Fok*I domain	No. of callus co-cultivated	Transgenic plant	Regeneration rate (%)[a]	Seed set[b]
PGD1::ZFN	F67/F96	720	1	0.14	1
RCc3:: ZFN	F67/F96	600	43	7.2	14
PGD1::ZFNDR	F67DAS/F96-RR	240	8	3.3	0
PGD1::ZFNRD	F67-RR/F96DAS	240	10	4.2	0
RCc3::ZFNDR	F67DAS/F96-RR	240	20	8.3	0
RCc3::ZFN RD	F67-RR/F96DAS	240	18	7.5	0
Total		2,280	100		15

[a](No. of transgenic plant/No. of callus co-cultivated)x100; [b]Number of plants. PGD1: constitutive promoter; RCc3: root-specific promoter; F67/F96: wild-type *Fok*I; F67DAS/F96-RR and F67-RR/F96DAS: *Fok*I variants. From 6 different vector constructs, a total of 100 transgenic plants were obtained. T_1 seeds were harvested from 15 T_0 plants.

A


NT
PGD1::
ZFN-1
RCc3 ::
ZFN-1
RCc3::
ZFN-2
RCc3::
ZFN-3
RCc3::
ZFN-4
RCc3::
ZFN-7
RCc3::
ZFN-8
RCc3::
ZFN-10
RCc3::
ZFN-13
RCc3::
ZFN-16
RCc3::
ZFN-19

B

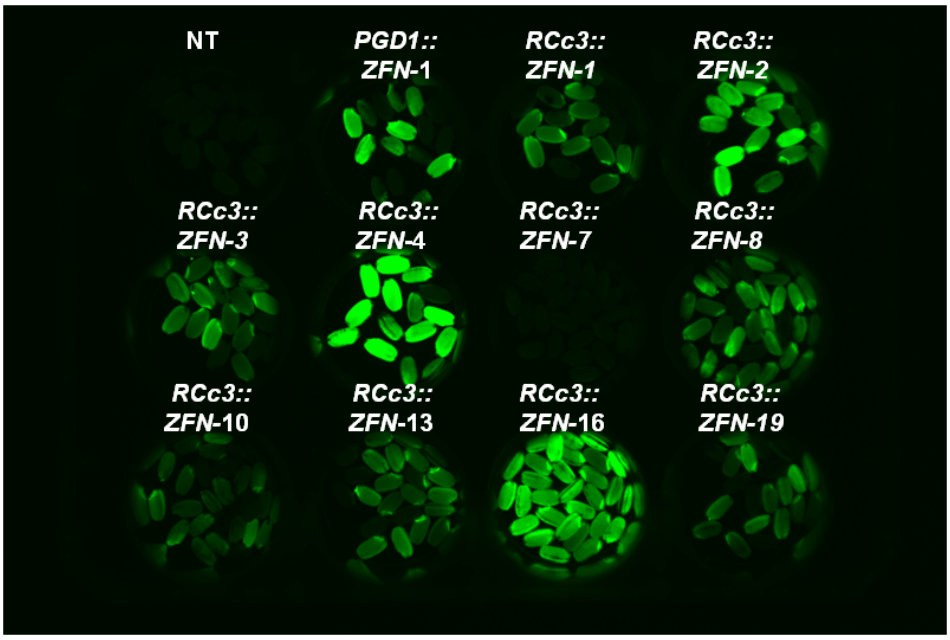
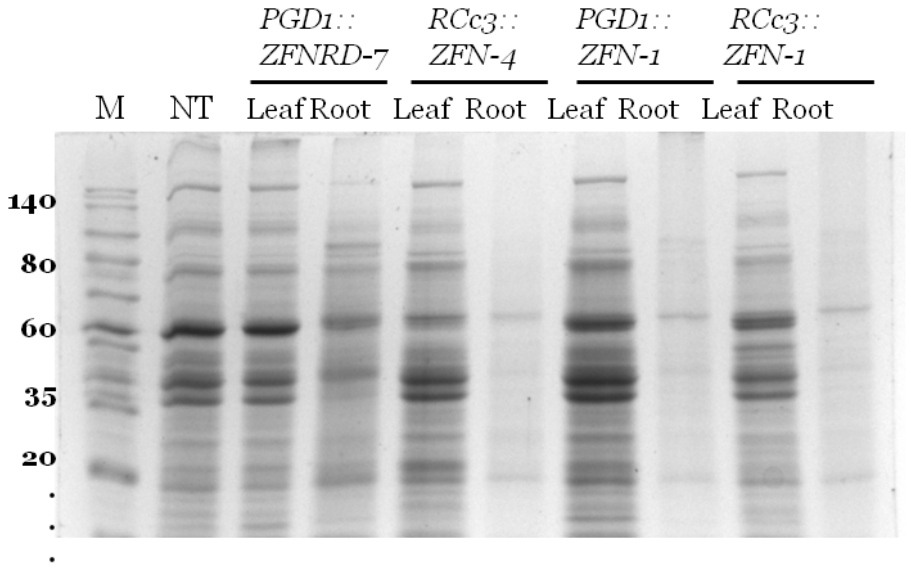

NT
PGD1::
ZFN-1
RCc3::
ZFN-1
RCc3::
ZFN-2
RCc3::
ZFN-3
RCc3::
ZFN-4
RCc3::
ZFN-7
RCc3::
ZFN-8
RCc3::
ZFN-10
RCc3::
ZFN-13
RCc3::
ZFN-16
RCc3::
ZFN-19

Figure 3. GFP assay for the dried rice seeds. (A) Photo image. The photo of non-transgenic and transgenic rice seed samples was taken under white light. The ID of each transgenic line is denoted above each sample. **(B) GFP image.** The photo of the same sample was taken under blue light to illuminate green fluorescence from GFP proteins (see Methods section). NT: non-transgenic. The lack of fluorescence in RCc3::ZFN-7 is due to impaired ZFN insert, consistent with the genomic PCR result showing internal deletion in ZFN insert.

A


PGD1::
ZFNRD-7
RCc3::
ZFN-4
PGD1::
ZFN-1
RCc3::
ZFN-1
M
NT
Leaf
Root
Leaf
Root
Leaf
Root
Leaf
Root
140
80
60
35
20

B

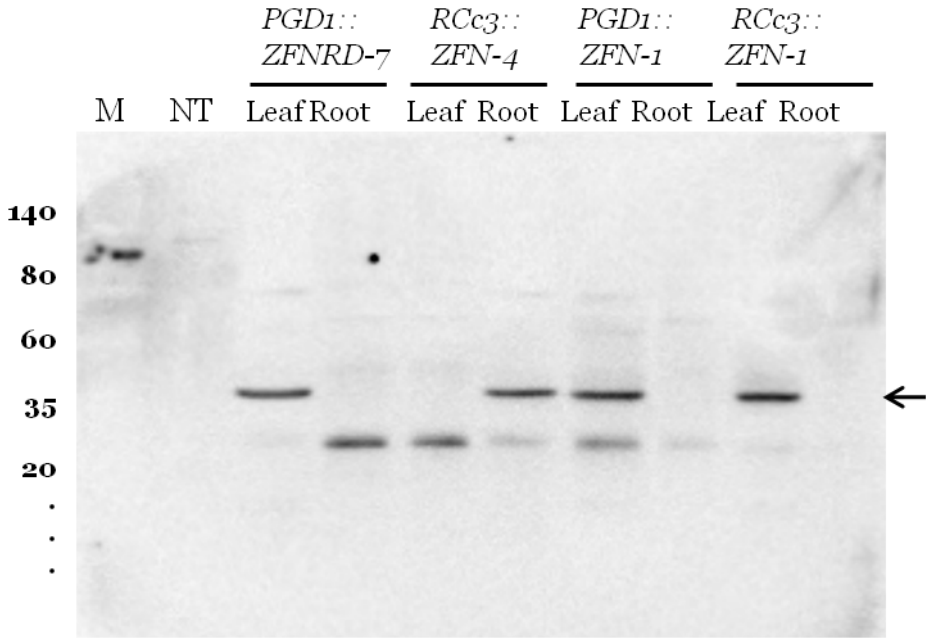

PGD1::
ZFNRD-7
RCc3::
ZFN-4
PGD1::
ZFN-1
RCc3::
ZFN-1
M
NT
Leaf
Root
Leaf
Root
Leaf
Root
Leaf
Root
140
80
60
35
20

Figure 4. Immunoblot analysis for ZFN expression in the transgenic rice plants. (A) Protein gel photograph. Total protein from rice leaf and root tissue was separated by SDS-PAGE as described (Laemmli 1970) and stained with Coomassie blue. **(B) Immunoblot gel photograph.** The types promoters, *Fok*I domains, and tissues are indicated on the top of the gel. RD indicates ZFN with a variant *Fok*I. *PGD1*: a constitutive promoter; *RCc3*: a root-specific promoter. NT: non-transgenic (negative control); M: Molecular size marker. The immnunoblot gel photo was taken after exposing 160 sec. The arrow indicates bands corresponding expressed ZFN proteins.

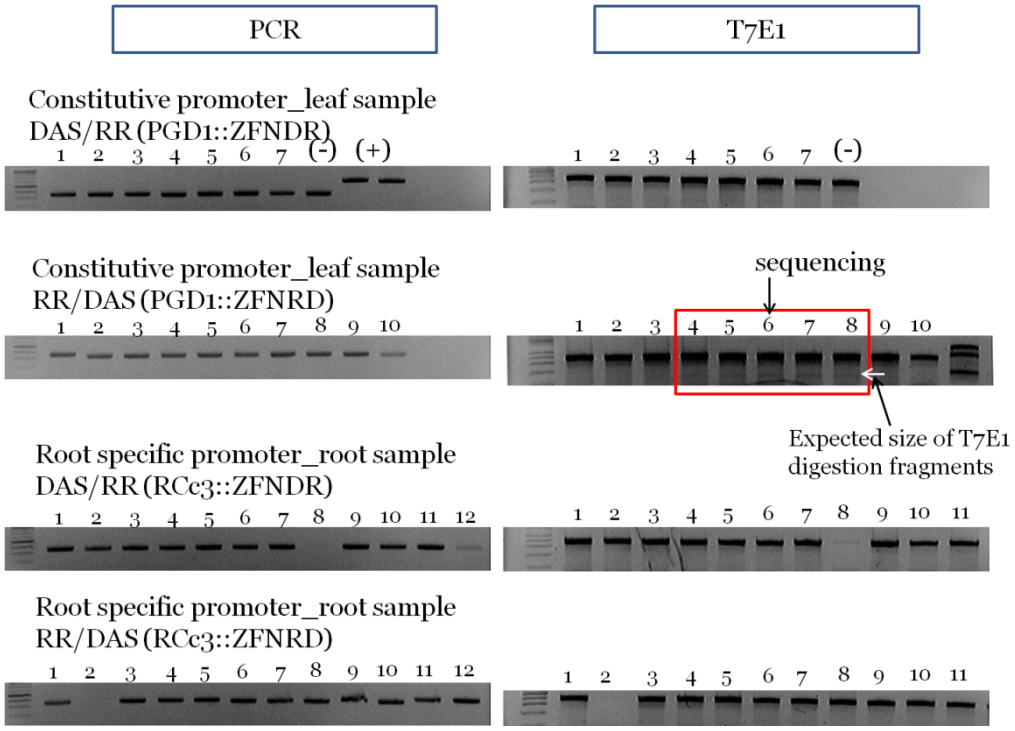


Figure 5. A mismatch-based assay using mismatch-selective T7 endonuclease. Hetero-duplex formation between wild and mutant type DNA strands in the target genomic region was detected in the transformed lines, PGD1::RD-4,5,6,7 and 8. The faint lower bands correspond to the expected size of T7E1 digestion products. PCR amplification was carried out for the 300 bp region spanning the target site using forward primer 5′-CCATCCTCTCCGACCTCATC-3′ and reverse primer 5′-CTAATCTCCCCACTGACCGC-3′. PGD1::ZFNRD-6 was chosen and subjected to further sequence analysis for the target site. The numbers above each gel photo indicate transgenic line numbers; (-): non-transgenic plant.

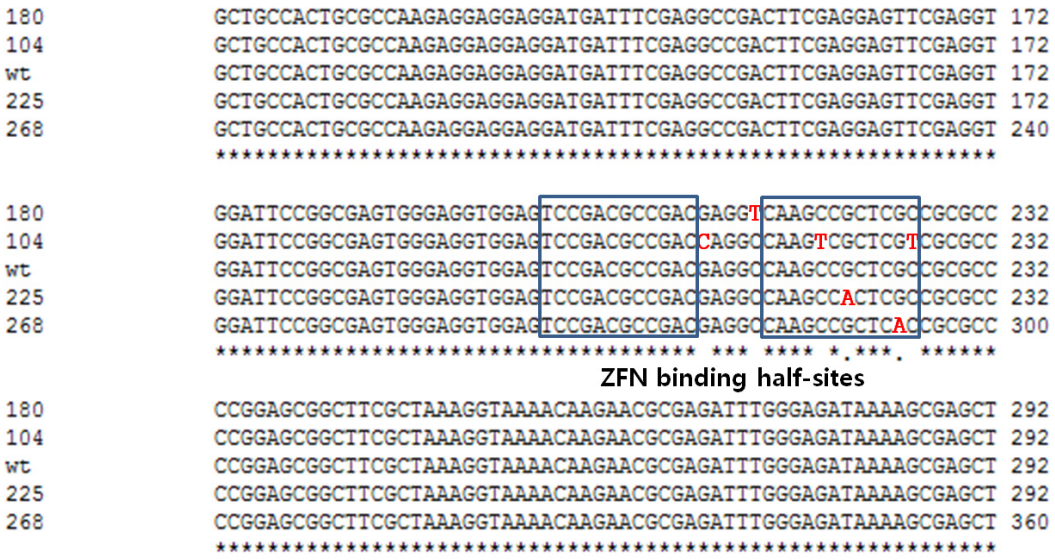


Figure 6. Detection of substitution mutations at the expected target site of *OsAP2* gene in rice. To further analyze mutagenesis events at the target site, the transgenic line PGD1::ZFNRD-6 was chosen and subjected to further sequence analysis for the target site. The 300 bp target region amplified by PCR was cloned and sequenced. Of 160 clones that have been sequenced, 4 clones show substitution mutations at the ZFN target site. The boxes represent ZFN binding sites; nucleotides highlighted in red represent substituted nucleotides. The size of PCR products is about 300 bp. The numbers to the left identify the individual clones that had been subjected to DNA sequencing. Wt: wild-type.

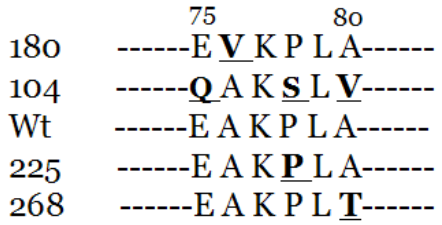


Figure 7. Amino acid sequence alignment for substitution mutations described in Figure 6. The numbers at the top of the amino acid sequences indicate the numbers of amino acids from the start codon. The mutated amino acids are highlighted by both bold letters and underlines. The numbers to the left identify the individual clones that had been subjected to DNA sequencing. Wt: wild-type.

4. Discussion

To my knowledge this is the first reported attempt at ZFN-guided mutagenesis for a gene with no selectable or screenable phenotype in plants. All previous reports dealing with ZFN-directed mutagenesis in plants have targeted genes with either selectable or screenable phenotypes (Cai et al. 2009; Curtin et al. 2011; de Pater et al. 2009; Even-Faitelson 2011; Lloyd et al. 2005; Marton et al. 2010; Osakabe et al. 2010; Petolino et al. 2010; Shukla et al. 2009; Tovkach et al. 2009; Townsend et al. 2009; Tzfira et al. 2012; Wright et al. 2005; Zhang et al. 2010). Obviously, targeting genes with selectable or screenable phenotype will make experiments much easier and more informative.

The analysis of targeted mutagenesis events by ZFN in my study showed the occurrence of 6 substitution mutations at the expected target region within sequences of 160 300-bp clones (Figure 6). If the 6 substitution mutations had arisen purely by random errors, the number of substitution mutations should have been much less, suggesting the occurrence of ZFN-directed mutagenesis events at the target *OsAP2*. The logic supporting this is as follows: The total number of base pairs sequenced is 48,000 bp (160 clones x 300 bp/clone). The error rate of *Taq* polymerase is 1/9,000 (Wikipedia) or 2.28×10^{-5} (Thermo Scientific et al.) and so the expected base errors by *Taq* polymerase within a span of 48,000 bp sequence is estimated to be 5.3 bp (48,000 bp x 1/9,000) by the error rate of Wikipedia; 1 bp (48,000 bp x 2.28×10^{-5}) by the error rate of Thermo Scientific et al. Moreover, the base errors arisen from *Taq* polymerase must be distributed randomly throughout the length of 48,000

bp sequence. As a result, it is extremely unlikely that 6 mutations had occurred in such concentrated region as in Figure 6 purely by chance. Rather, it is more likely that the six mutations had been resulted from ZFN-directed events.

However, indel mutation, a signature mutation usually resulting from non-homologous end joining (NHEJ), was not detected in my assay at the expected target site. A number of speculations could be made for now.

First, it is possible that continued overexpression of ZFNs has evoked toxic effect on the cell, resulting in selective apoptosis or cell death and leaving only undamaged or little damaged cells to survive. Cytotoxicity has long been keen issue in the ZFN community (Alwin et al. 2005; Bibikova et al. 2002, Cornu et al. 2008; Kim et al. 2009; Miller et al. 2007; Osiak et al. 2011; Pattanayak et al. 2011; Ramalingam et al. 2011; Szczepek et al. 2007). Homodimerization of *Fok*I domain leading to off-target cleavage has been one of the major causes of cytotoxicity (Carroll 2011; Durai et al. 2005; Ramalingam et al. 2011; Urnov et al. 2010). Indeed, the variant *Fok*I domain designed to exert obligate heterodimerization has demonstrated reduced toxicity and as a result, enhanced targeting efficiency (Doyon et al. 2011; Guo et al. 2010; Miller et al. 2007; Szczepek et al. 2007). Unusually low rate of regeneration of transgenic rice plants particularly from wild type FokI-ZFN transformed calli (Table 1) and the absence of detectable indel mutations from the transgenic plants (Figure 6) strongly suggest the widespread occurrence of cytotoxicity in my experiment. This implication is further supported by the significantly increased rate of rice regeneration from the calli transformed

with obligate heterodimeric variant *Fok*I-ZFNs (Table 1). The toxic effect of wild-type *Fok*I-ZFNs was particularly pronounced when expressed under the constitutive promoter (*PGD1*). This is probably due to the fact that, in the presence of ZFN expression, actively dividing cells in calli were more severely damaged to ZFN cleavage because of loose chromosomes. In contrast, calli containing ZFNs driven by root-specific promoter (*RCc3*) might have been protected from ZFN off-targeting because of negligible ZFN expression in calli.

Second, it is likely that high levels of continuous ZFN expression may have contributed negatively to the cytotoxic effect of *Fok*I nuclease (Kim et al. 2009, Carlson et al. 2012). Although there is a report claiming plants can tolerate strong ZFN expression (de Pater 2009), there are numerous reports suggesting toxic effect of continued expression of ZFN (Alwin et al. 2005; Carroll 2008; Durai et al. 2005). My data also support this argument. The greater extent of reduced regeneration rate in constitutive promoter relative to root-specific promoter in case of wild-type *Fok*I ZFNs could be explained by toxicity resulting from continued strong expression of ZFN (Table 1). This suggests that it is desirable that ZFNs be expressed either transiently or under the direction of weak promoters (Alwin at al. 2005; Ramalingam 2011). Indeed inducible promoters have been used successfully (Osakebe et al. 2010; Townsend et al. 2009; Shukla et al. 2009; Zhang et al. 2010). If the above reasoning is correct, one might ask a question why my transgenic rice plants were so vigorous (data not shown). It is likely that cells that carry ZFN-induced mutations were growth-impaired and thus outgrown by unmodified cells (Kim at al.

2009). Another possibility might reside in the nature of *OsAP2* expression pattern. *OsAP2* gene encodes a putative transcription factor and presumably controls flower and seed development in rice (Okamuro 1997; Oh et al. 2009). So during the most of vegetative growth periods, the *OsAP2* gene was inactive and locked in a compact chromatin structure, making ZFN nucleases inaccessible. In the meantime, the potential off-target sites are likely to be accessible by still active ZFNs. Dividing cells are known to be more susceptible to DNA-cleavage activities in general. Although progresses are being made at an accelerated pace, engineered ZFN technology is still in the stage of development. Cytotoxicity and targeting efficiency are closely linked. Many of the causes of toxicity are known and tackling toxicity problem would lead to give rise to better and improved ZFNs.

Third, the modular assembly followed by subsequent selection of active ZFNs for the sequence *OsAP2* via yeast SSA spotting assay (Chapter I) may not have produced ZFNs with optimal specificity and affinity for *OsAP2* target sequence *in vivo* because yeast SSA assay does not take plant chromosomal context into consideration. It is also possible that ZFNs were inaccessible to the target due to compact chromatin structure, DNA methylation, or other factors (Carroll 2011).

There is an implication of *OsAP2* gene being involved in early flower development like AP2 genes in other plants. Thus it is possible that bi-allelic knockout of *OsAP2* should abolish flower development, leading to sterility. In my study, it is too early to draw any conclusion because the majority (85%) of the transgenic plants is still at vegetative stage.

Genotyping throughout T_1 and T_2 generations would help address this important question.

One lesson from my studies is that the efficacy and safety of ZFNs in the cells are largely unpredictable and should be determined empirically. Thus it will be desirable to develop a quick and easy assay system to test the delivery method and the efficiency and toxicity of ZFNs in the cell beforehand. Alternatively enriching mutant cells using surrogate reporters would be a clever option (Kim et al. 2011). The ZFN field is still in its early stages especially in plant community. The more experiences go through, the better understanding we would gain of the parameters that are most critical and the procedures that are most effective (Carroll 2008).

Recently, a DNA-binding motif that has been identified in proteins known as transcription activator-like effectors (TALEs) has been attached to the *Fok*I nuclease domain to yield novel fusion proteins, TALEN (Baker 2012a). TALENs, like ZFNs, consist of assembled DNA-binding motifs coupled to *Fok*I nuclease. The DNA-binding motifs, TALEs, originally derived from plant pathogens belonging to *Xanthomonas* are made up of 'repeats', or domains of 33-35 conserved amino acids. In the middle of each repeat are two variable amino acids that determine which nucleotide a repeat recognizes. While a zinc finger unit recognizes 3 nucleotides, repetitive units of the TALEs interact with specific nucleotides in a 1:1 ratio (Baker 2012b; Carlson et al. 2012). Since TALENs are easy to construct, reliable in function and rely on plant pathogen-derived DNA-binding domain, adoption of TALENs for plant breeding purpose will be a worthwhile option for ZFNs.

5. References

Alwin S, Gere MB, Guhl E, Effertz K, Barbas CF III, Segal DJ, Weitzman MD, Cathomen T. (2005) Custom zinc-finger nucleases for use in human cells. Mol Ther. 12: 610-607.

Baker M. (2012a) Gene-editing nucleases. Nat Methods 9: 23-26.

Baker M. (2012b) How TAL effectors bind DNA. Nat Methods 9: 222.

Bibikova M, Golic M, Golic KG, Carroll D. (2002) Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. Genetics 161: 1169-1175.

Cai CQ, Doyon Y, Ainley WM, Miller JC, Dekelver RC, Moehle EA, Rock JM, Lee YL, Garrison R, Schulenberg L, Blue R, Worden A, Baker L, Faraji F, Zhang L, Holmes MC, Rebar EJ, Collingwood TN, Rubin-Wilson B, Gregory PD, Urnov FD, Petolino JF. (2009) Targeted transgene integration in plant cells using designed zinc finger nucleases. Plant Mol. Biol. 69: 699-709.

Carlson DF, Fahrenkrug SC, Hackett PB. (2012) Targeting DNA with fingers and TALENs. Molecular Therapy — Nucleic Acids 1: e3.

Carroll D. (2008) Progress and prospects: Zinc-finger nucleases as gene therapy agents. Gene Ther. 15: 1463-1468.

Carroll D. (2011) Genome engineering with zinc-finger nucleases. Genetics 188: 773-782.

Cornu TI, Thibodeau-Beganny S, Guhl E, Alwin S, Eichtinger M, Joung JK, Cathomen T. (2008) DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases. Mol Ther. 16: 352-358.

Curtin SJ, Zhang F, Sander JD, Haun WJ, Starker C, Baltes NJ, Reyon D, Dahlborg EJ, Goodwin MJ, Coffman AP, Dobbs D, Joung JK, Voytas DF, Stupar RM. (2011) Targeted mutagenesis of duplicated genes in soybean with zinc-finger nucleases. Plant Physiol. 156: 466-473.

de Pater S, Neuteboom LW, Pinas JE, Hooykaas PJ, van der Zaal BJ. (2009) ZFN-induced mutagenesis and gene-targeting in *Arabidopsis* through *Agrobacterium*-mediated floral dip transformation. Plant Biotechnol J. 7: 821-835.

Doyon Y, Vo TD, Mendel MC, Greenberg SG, Wang J, Xia DF, Miller JC, Urnov FD, Gregory PD, Holmes MC. (2011) Enhancing zinc-finger-nuclease activity with improved obligate heterodimer architectures. Nat Methods 8: 74-79.

Durai S, Mani M, Kandavelou K, Wu J, Porteus MH, Chandrasegaran S. (2005) Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells. Nucleic Acids Res. 33: 5978-5990.

Eisenstein, M. (2005) Human gene repair at your fingertips. Nat Methods 2: 405.

Even-Faitelson L, Samach A, Melamed-Bessudo C, Avivi-Ragolsky N, Levy AA. (2011) Localized egg-cell expression of effector proteins for targeted modification of the *Arabidopsis* genome. Plant J. 68: 929-937.

Guo J, Gaj T, Barbas CF III. (2010) Directed evolution of an enhanced and highly efficient *FokI* cleavage domain for zinc finger nucleases. J Mol Biol. 400: 96-107.

Ha SH, Liang YS, Jung H, Ahn MJ, Suh SC, Kweon SJ, Kim DH, Kim

YM, Kim JK. (2010) Application of two bicistronic systems involving 2A and IRES sequences to the biosynthesis of carotenoids in rice endosperm. Plant Biotechnol J. 8: 928-938.

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 T-DNA. Plant J. 6: 271-282.

Jang IC, Choi WB, Lee KH, Song SI, Nahm BH, Kim JK. (2002) High-level and ubiquitous expression of the rice cytochrome *c* gene *OsCc1* and its promoter activity in transgenic plants provides a useful promoter for transgenesis of monocots. Plant Physiol. 129: 1473-1481.

Jang IC, Nahm BH, Kim JK. (1999) Subcellular targeting of green fluorescent protein to plastids in transgenic rice plants provides a high-level expression system. Molecular Breeding 5: 453-461.

Kim H, Um E, Cho SR, Jung C, Kim H, Kim JS. (2011) Surrogate reporters for enrichment of cells with nuclease-induced mutations. Nat Methods 8: 941-943.

Kim HJ, Lee HJ, Kim H, Cho SW, Kim JS. (2009) Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. Genome Res. 19: 1279-1288.

Kim S, Kim EJ, Kim JS. (2010) Construction of combinatorial libraries that encode zinc finger-based transcription factors. Methods Mol Biol. 649: 133-147.

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. 19: 165-174.

Laemmli UK. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680-685.

Lloyd A, Plaisier CL, Carroll D, Drews GN. (2005) Targeted mutagenesis using zinc-finger nucleases in *Arabidopsis*. Proc Natl Acad Sci USA. 102: 2232–2237.

Marton I, Zuker A, Shklarman E, Zeevi V, Tovkach A, Roffe S, Ovadis M, Tzfira T, Vainstein A. (2010) Nontransgenic genome modification in plant cells. Plant Physiol. 154: 1079-1087.

Miller JC, Holmes MC, Wang J, Guschin DY, Lee YL, Rupniewski I, Beausejour CM, Waite AJ, Wang NS, Kim KA, Gregory PD, Pabo CO, Rebar EJ. (2007) An improved zinc-finger nuclease architecture for highly specific genome cleavage. Nat Biotechnol. 25: 778-785.

Oh SJ, Kim YS, Kwon CW, Park HK, Jeong JS, Kim JK. (2009) Overexpression of the transcription factor *AP37* in rice improves grain yield under drought conditions. Plant Physiol. 150: 1368-1379.

Okamuro JK, Caster B, Villarroel R, Van Montagu M, Jofuku KD. (1997) The AP2 domain of APETALA2 defines a large new family of DNA binding proteins in *Arabidopsis*. Proc Natl Acad Sci USA. 94: 7076-7081.

Osakabe K, Osakabe Y, Toki S. (2010) Site-directed mutagenesis in *Arabidopsis* using custom-designed zinc finger nucleases. Proc Natl Acad Sci USA. 107: 12034-12039.

Osiak A, Radecke F, Guhl E, Radecke S, Dannemann N, Lütge F, Glage S, Rudolph C, Cantz T, Schwarz K, Heilbronn R, Cathomen T. (2011)

Selection-independent generation of gene knockout mouse embryonic stem cells using zinc-finger nucleases. PLoS One. 6: e28911.

Pattanayak V, Ramirez CL, Joung JK, Liu DR. (2011) Revealing off-target cleavage specificities of zinc-finger nucleases by *in vitro* selection. Nat Methods 8: 765-770.

Park SH, Yi N, Kim YS, Jeong MH, Bang SW, Choi YD, Kim JK. (2010) Analysis of five novel putative constitutive gene promoters in transgenic rice plants. J. Exp. Bot. 61: 2459-2467.

Petolino JF, Worden A, Curlee K, Connell J, Strange Moynahan TL, Larsen C, Russell S. (2010) Zinc finger nuclease-mediated transgene deletion. Plant Mol Biol. 73: 617-628.

Perez EE, Wang J, Miller JC, Jouvenot Y, Kim KA, Liu O, Wang N, Lee G, Bartsevich VV, Lee YL, Guschin DY, Rupniewski I, Waite AJ, Carpenito C, Carroll RG, Orange JS, Urnov FD, Rebar EJ, Ando D, Gregory PD, Riley JL, Holmes MC, June CH. (2008) Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. Nat Biotechnol. 26: 808-816.

Ramalingam S, Kandavelou K, Rajenderan R, Chandrasegaran S. (2011) Creating designed zinc-finger nucleases with minimal cytotoxicity. J Mol Biol. 405: 630-641.

Sambrook J, Russell DW. (2001) Molecular cloning: a laboratory manual. 3rd ed. Cold Spring Harbor Press, N.Y.

Shukla VK, Doyon Y, Miller JC, DeKelver RC, Moehle EA, Worden SE, Mitchell JC, Arnold NL, Gopalan S, Meng X, Choi VM, Rock JM, Wu YY, Katibah GE,Zhifang G, McCaskill D, Simpson MA, Blakeslee B,

Greenwalt SA, Butler HJ, Hinkley SJ, Zhang L, Rebar EJ, Gregory PD, Urnov FD. (2009) Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. Nature 459: 437-441.

Szczepek M, Brondani V, Büchel J, Serrano L, Segal DJ, Cathomen T. (2007) Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. Nat Biotechnol. 25: 786-793.

Tovkach A, Zeevi V, Tzfira T. (2009) A toolbox and procedural notes for characterizing novel zinc finger nucleases for genome editing in plant cells. Plant J. 57: 747-757.

Townsend JA, Wright DA, Winfrey RJ, Fu F, Maeder ML, Joung JK, Voytas DF. (2009) High-frequency modification of plant genes using engineered zinc-finger nucleases. Nature 459: 442-445.

Tzfira T, Weinthal D, Marton I, Zeevi V, Zuker A, Vainstein A. (2012) Genome modifications in plant cells by custom-made restriction enzymes. Plant Biotechnol J. 10: 373-389.

Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD. (2010) Genome editing with engineered zinc finger nucleases. Nat Rev Genet. 11: 636-646.

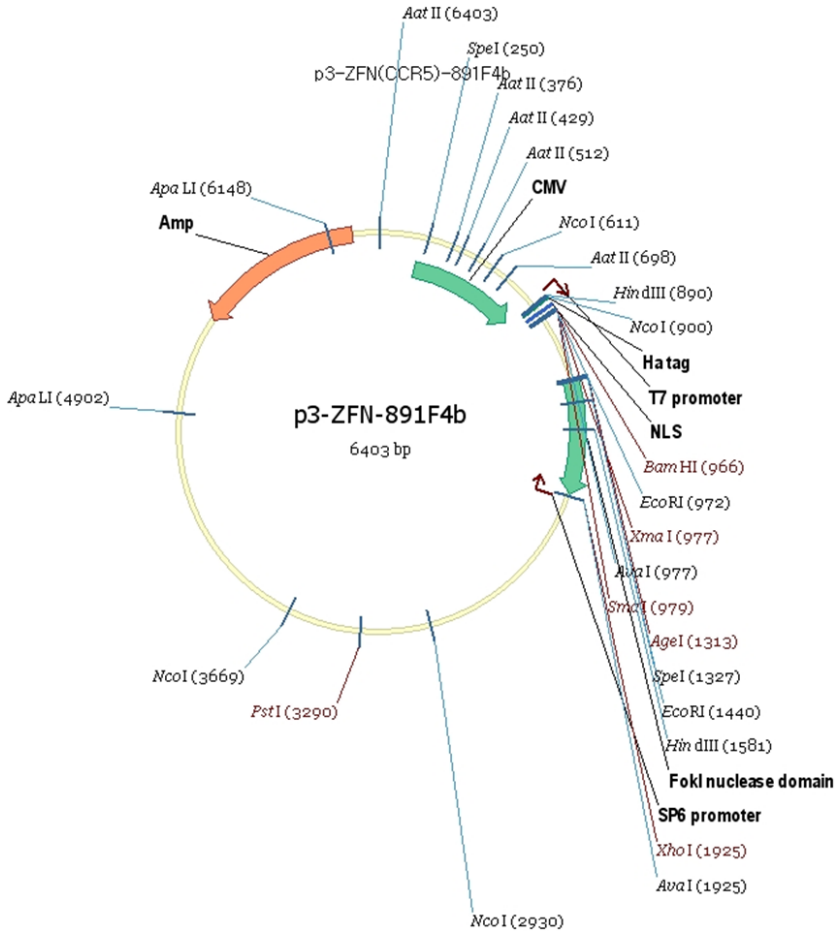
Wright DA, Townsend JA, Winfrey RJ Jr, Irwin PA, Rajagopal J, Lonosky PM, Hall BD, Jondle MD, Voytas DF. (2005) High-frequency homologous recombination in plants mediated by zinc-finger nucleases. Plant J. 44: 693-705.

Yi N, Oh SJ, Kim YS, Jang HJ, Park SH, Jeong JS, Song SI, Choi YD, Kim JK. (2011) Analysis of the *Wsi18*, a stress-inducible promoter that is active in the whole grain of transgenic rice. Tansgenic Res. 20:

153-163.

Zhang F, Maeder ML, Unger-Wallace E, Hoshaw JP, Reyon D, Christian M, Li X, Pierick CJ, Dobbs D, Peterson T, Joung JK, Voytas DF. (2010) High frequency targeted mutagenesis in *Arabidopsis thaliana* using zinc finger nucleases. Proc Natl Acad Sci USA. 107: 12028-12033.

APPENDIX 1

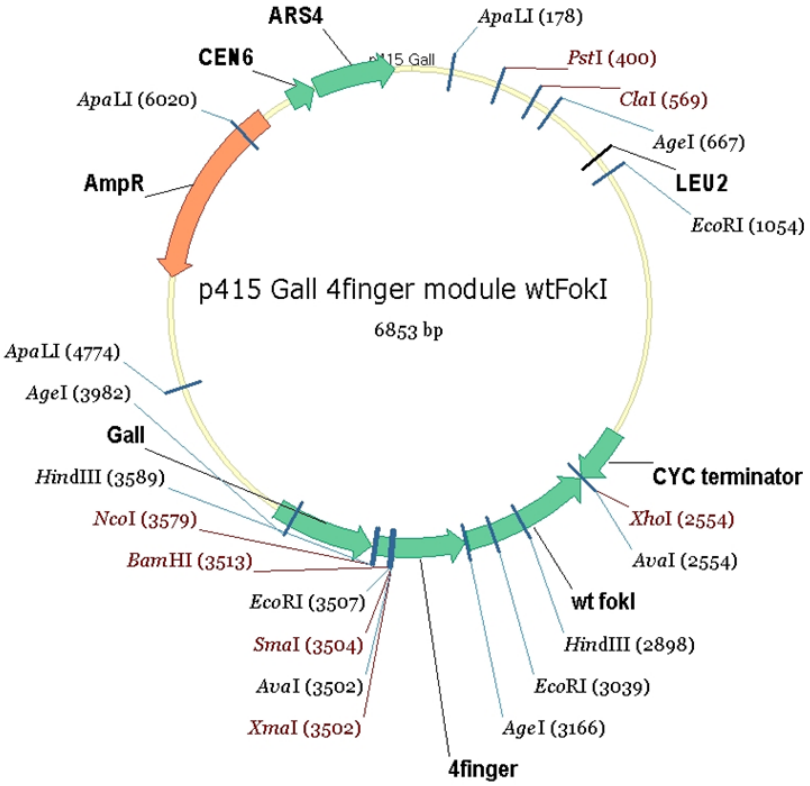


The p3 backbone plasmid carrying a 4-finger ZFN insert, ZFN-891F4b. The p3 vector, served as a backbone vector, is a derivative of pcDNA3.0 plasmid (Invitrogen). The p3 has a bacterial origin of replication, ColE1, and ampicillin ORF that serves as a selection marker. Downstream of CMV promoter lies a multiple cloning site (MSC) between T7 promoter (forward direction with respect to P_{CMV}) and SP6 promoter (backward direction). The MSC region is highlighted in Appendix 2.

APPENDIX 2

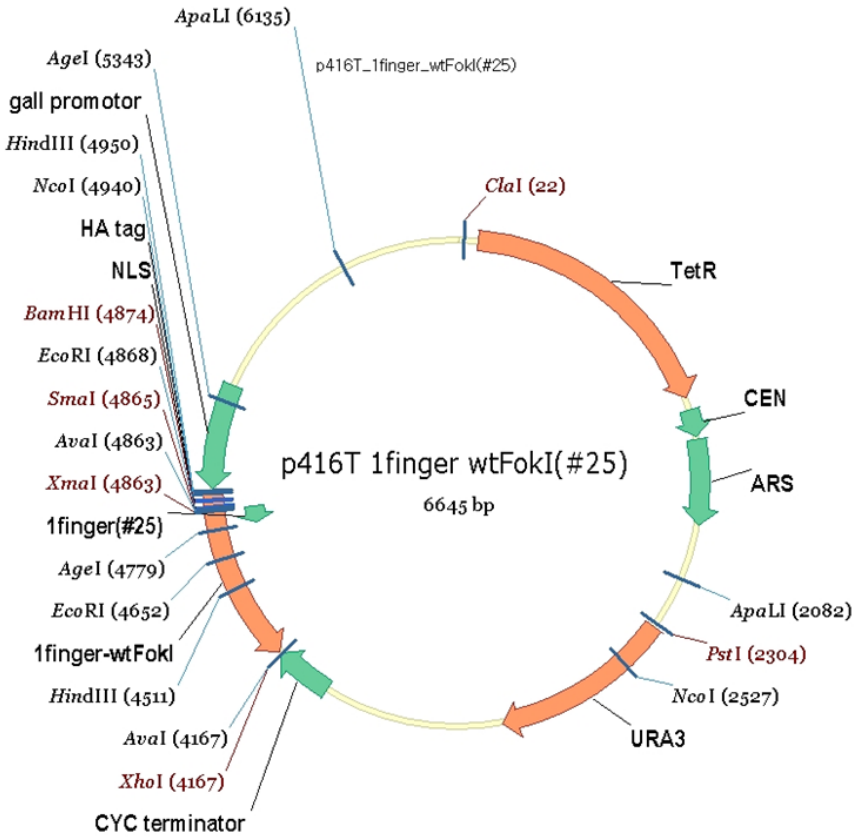
Nucleotide sequence of the MCS of the p3 expression plasmid housing ZF107 (Zinc Finger Consortium Collection, Addgene). Amino acid residues important for DNA recognition specificity are indicated by small squares. The keys in the figure are referred to Section 2.1 in Chapter I (Appendix 2 is taken from Kim et al. 2010).

APPENDIX 3



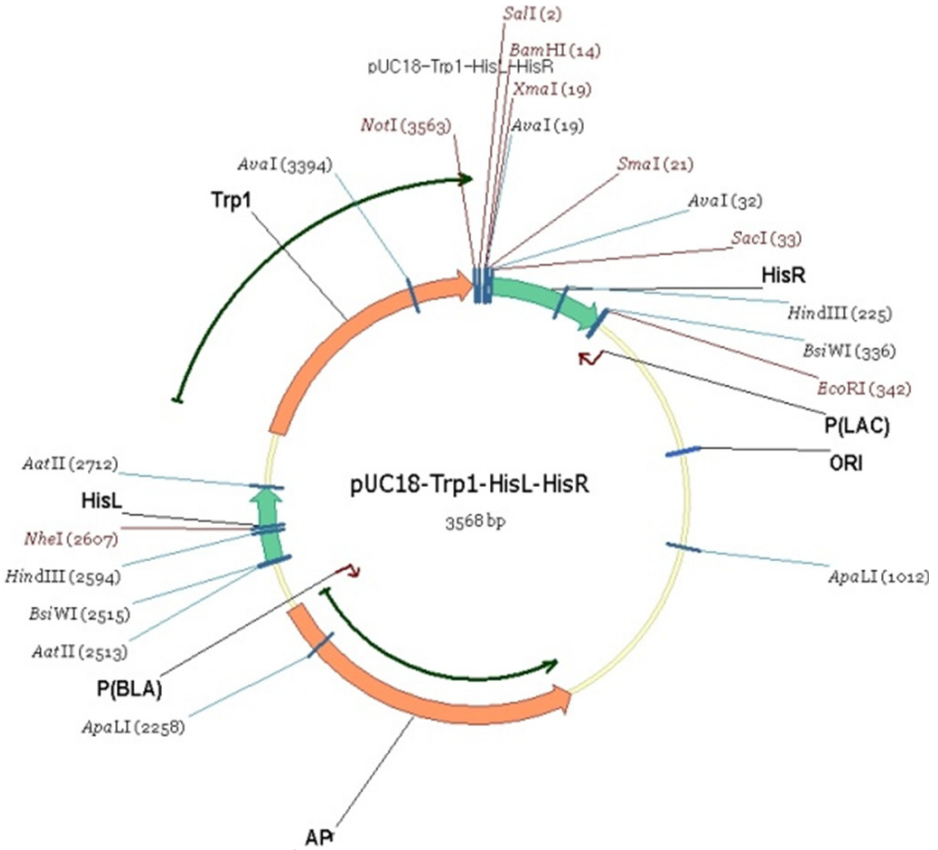
p415 yeast expression vector. The p415 vector served as a backbone vector of yeast ZFN expression vectors that express ZFN-L. It is a galactose inducible yeast expression vector and carries Amp^R for selection. *Bam*HI (3513) and *Xho*I (2554) served as cloning sites for 4F-ZFNs which carry identical restriction sites at both ends.

APPENDIX 4



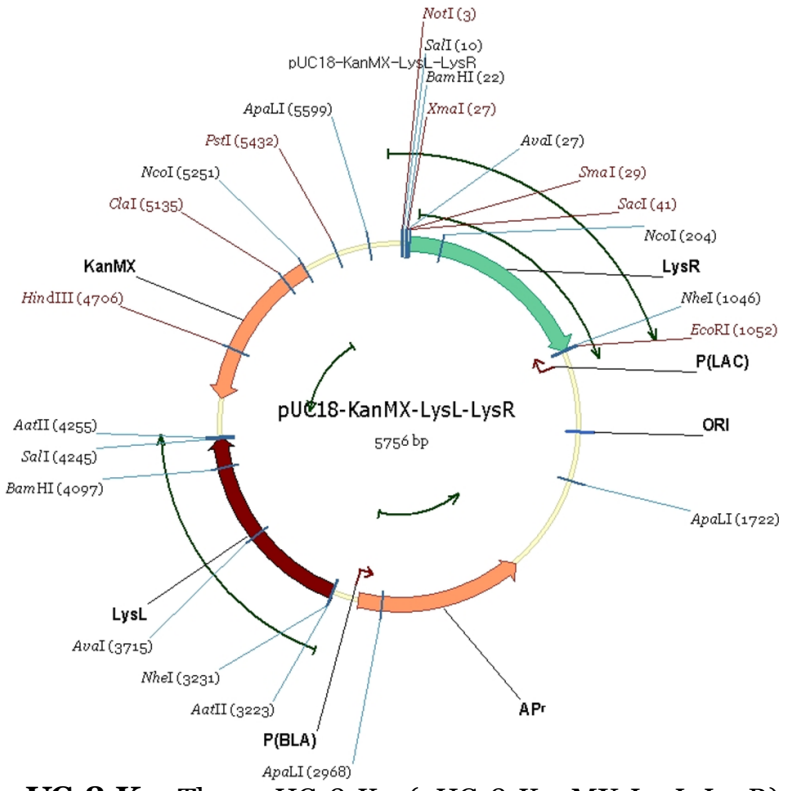
p416T yeast expression vector. The p416T vector served as a backbone vector of yeast ZFN expression vectors that express ZFN-R. It is a galactose inducible yeast expression vector and carries Amp^R for selection. *Bam*HI (4874) and *Xho*I (4167) served as cloning sites for 4F-ZFNs which carry identical restriction sites at both ends.

APPENDIX 5



Map of pUC18-H. The pUC18-H (pUC18-Trp1-HisL-HisR) contains overlapping truncated copies of yeast His gene (HisL and HisR) with a partial duplication of coding sequence. The sequence duplication flanks a Trp1 gene, a selection marker and a polylinker for cloning the target sequence being tested. After cleavage of the target sequence by the ZFN, recombination between the His sequence duplication results in loss of Trp1, conferring Trp^- phenotype. *OsAP2* ORF sequence inserted into the MCS was delivered to the yeast genome of *Mat* **a** strain via *Bsi*WI fragments. HisL and HisR facilitate integration of *OsAP2* ORF sequence into the yeast genome by stimulating homologous recombination with genomic HIS locus.

APPENDIX 6



Map of pUC18-K. The pUC18-K (pUC18-KanMX-LysL-LysR) contains overlapping truncated copies of yeast Lys gene (LysL and LysR) with a partial duplication of coding sequence. The sequence duplication flanks a KanMX gene, a selection marker and a polylinker for cloning the target sequence being tested. After cleavage of the target sequence by the ZFN, recombination between the Lys sequence duplication results in loss of KanMX, conferring Kan$^-$ phenotype). *OsAP2* ORF sequence inserted into the MCS was delivered to the yeast genome of *Mat* **α** strain via *Nhe*I fragments. LysL and LysR facilitate integration of *OsAP2* ORF sequence into the yeast genome by stimulating homologous recombination with genomic LYS locus.

APPENDIX 7

a) Wild-type *Fok*I cleavage domain
QLVKSELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYG
YRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVEE
NQTRNKHINPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNHITNCNGA
VLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINF

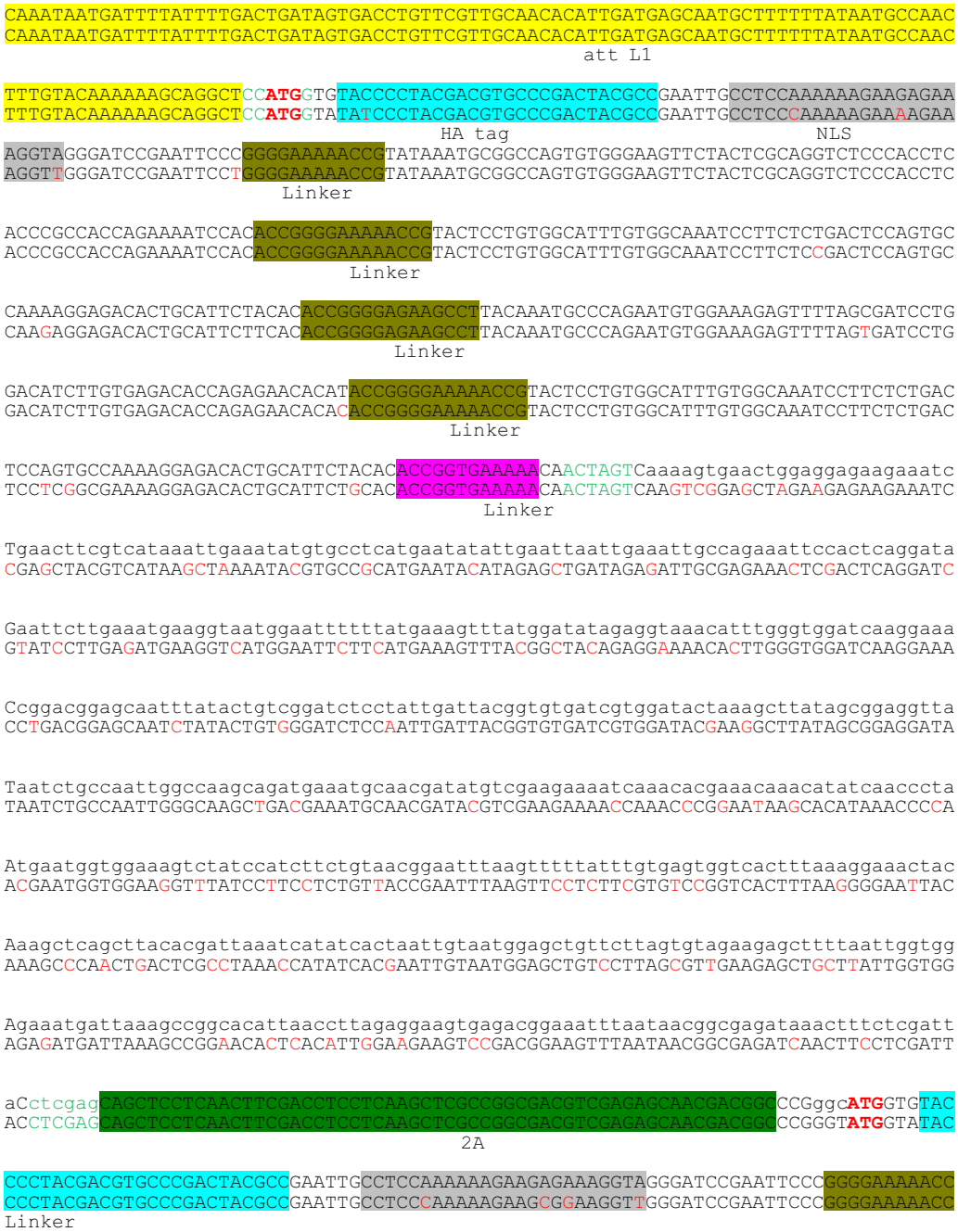
b) FCD_{R18-28} (*Sharkey'*)
QLVKSELEEKKSELRHKLKYVPHEYIELIEIARNPTQDRILEMKVMEFLMKVYG
YRGEHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGHADEMQRYVEE
NQTRNKHINPNEWWKVYPSSVTEFKFLFVSGYFKGDYKAQLTRLNHITNCNGA
VLSVEELLIGGEMIQAGTLTLEEVRRKFNNGEINF

c) *Sharkey'* D483R (RR variant)
QLVKSELEEKKSELRHKLKYVPHEYIELIEIARNPTQDRILEMKVMEFLMKVYG
YRGEHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGHARREMQRYVEE
NQTRNKHINPNEWWKVYPSSVTEFKFLFVSGYFKGDYKAQLTRLNHITNCNGA
VLSVEELLIGGEMIQAGTLTLEEVRRKFNNGEINF

d) *Sharkey'* DAMQS (DAS variant)
QLVKSELEEKKSELRHKLKYVPHEYIELIEIARNPTQDRILEMKVMEFLMKVYG
YRGEHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGHADAMQSYVEE
NQTRNKHINPNEWWKVYPSSVTEFKFLFVSGYFKGDYKAQLTRLNHITNCNGA
VLSVEELLIGGEMIQAGTLTLEEVRRKFNNGEIN

Zinc finer nuclease construct sequences. The complete amino acid sequences of (a) the *Fok*I cleavage domain, (b) *Sharkey*', (c) *Sharkey*' D483R and (d) *Sharkey'* DAMQS. Amino acids 384 to 579 of the full-length *Fok*I were used as the cleavage domain. Differences between wild type and selected FCD variants are underlined. Differences in *Sharkey* variants relative to wild-type are in red. Mutations unique to heterodimers are in blue (Appendix 7 was taken from Guo et al. 2010).

APPENDIX 8


att L1
HA tag
NLS
Linker
Linker
Linker
Linker
Linker
2A
Linker

GTACAAGTGCCCCGAGTGCGGCAAGAGCTTCAGCCACACCGGCCACCTGCTGGAGCACCAGCGGACCCACACCGGGGAAA
GTACAAGTGCCCGGAGTGCGGCAAGAGTTTCAGCCACACCGGGCACCTGCTGGAGCACCAGCGGACACACACGGGGAAA
Linker

AACCGTATGAGTGTAATTACTGTGGAAAAACCTTTAGTGTGAGCTCAACCCTTATTAGACATCAGAGAATCCACACCGGG
AACCGTATGAGTGTAATTACTGTGGAAAAACCTTTAGTGTGAGCTCAACCCTTATTAGACATCAGAGAATCCACACCGGG
Linker

GAAAAACCGTACACATGCAGTGACTGTGGGAAGGCTTTCAGAGATAAATCATGTCTCAACAGACATCGGAGAACTCATAC
GAAAAACCTTACACATGCAGTGACTGTGGGAAGGCTTTCAGAGATAAATCATGCCTGAACAGACATCGGAGAACTCATAC
 Linker

CGGGGAAAAACCGTACAAATGTGACGAATGTGGAAAAAACTTTACCCAGTCCTCCAACCTTATTGTACATAAGAGAATTC
CGGGGAAAAACCGTACAAATGTGACGAATGTGGAAAGAACTTTACCCAGTCCTCCAACCTAATAGTCCATAAGAGAATCC
  Linker

ATACCGGTGAAAAACATCTAGACaaaagtgaactggaggagaagaaatctgaacttcgtcataaattgaaatatgtgcct
ATACCGGTGAAAAGCATCTAGACAAATCCGAACTGGAGGAGAAGAAATCTGAGCTTCGTCATAAACTGAAATATGTGCCT
     Linker

Catgaatatattgaattaattgaaattgccagaaattccactcaggatagaattcttgaaatgaaggtaatggaatttttt
CATGAATATATTGAGCTCATTGAGATTGCCAGAAATTCCACTCAGGATAGAATTCTTGAAATGAAAGTGATGGAATTCTT

Tatgaaagtttatggatatagaggtaaacatttgggtggatcaaggaaaccggacggagcaatttatactgtcggatctc
TATGAAAGTCTATGGATATCGTGGTAAGCATTTGGGTGGATCAAGGAAACCGGACGGAGCAATTTATACTGTGGGAAGCC

Ctattgattacggtgtgatcgtggatactaaagcttatagcggaggttataatctgccaattggccaagcagatgaaatg
CTATTGATTACGGTGTGATCGTGGATACTAAAGCTTATAGCGGAGGTTATAATCTGCCAATTGGCCAAGCGGATGAAATG

Caacgatatgtcgaagaaaatcaaacacgaaacaaacatatcaaccctaatgaatggtggaaagtctatccatcttctgt
CAACGATATGTCGAAGAAAATCAAACAAGGAACAAACATATCAACCCTAATGAATGGTGGAAAGTCTACCCATCTTCTGT

Aacggaatttaagtttttatttgtgagtggtcactttaaaggaaactacaaagctcagcttacacgattaaatcatatca
CACGGAATTTAAGTTTCTGTTTGTGAGCGGTCACTTTAAGGGAAACTATAAAGCTCAGCTTACACGTCTGAATCACATCA

Ctaattgtaatggagctgttcttagtgtagaagagcttttaattggtggagaaatgattaaagccggcacattaacctta
CTAATTGTAATGGAGCTGTTCTTAGTGTTGAAGAGCTTCTGATTGGTGGAGAAATGATCAAAGCCGGCACACTGACTTTG

gaggaagtgagacggaaatttaataacggcgagataaactttctcgattagtagGTCGACCGCGGCCGCCAAATAATGAT
GAGGAAGTGAGACGGAAATTTAATAACGGCGAGATAAACTTTCTCGATTAGTAGGTCGACCGCGGCCGCCAAATAATGAT

TTTATTTTGACTGATAGTGACCTGTTCGTTGCAACACATTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAA
TTTATTTTGACTGATAGTGACCTGTTCGTTGCAACACATTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAA
                                                  att L2

AAGCAGGCT
AAGCAGGCT

APPENDIX 9

<List of ZFN-Transformed Rice Plants>

No.	Plant ID	*Fok*Itype	Transfer date to Greenhouse	No. of Seeds Harvested	PCR detection of ZFN integration	GFP assay for dry seeds	Dates of Cocultivation*
1	PGD1-ZFN1	WT (F67/F96)	20110706	101	X	Positive	2011-3/3,3/10,3/17,
							3/24,3/31,7/14(총6회)
2	RCc3-ZFN1	WT (F67/F96)	20110624	880	X	Positive	2011-3/28,3/31,4/4,
							4/7,4/18 (총5회)
3	RCc3-ZFN2	WT (F67/F96)	20110706	191	X	Positive	"
4	RCc3-ZFN3	WT (F67/F96)	20110706	230	X	Positive	"
5	RCc3-ZFN4	WT (F67/F96)	20110713	204	O	Positive	"
6	RCc3-ZFN5	WT (F67/F96)	20110713				
7	RCc3-ZFN6	WT (F67/F96)	20110721	0	△		
8	RCc3-ZFN7	WT (F67/F96)	20110727	323	□(Deletion)	Negative	"
9	RCc3-ZFN8	WT (F67/F96)	20110727	127	O	Positive	"
10	RCc3-ZFN9	WT (F67/F96)	20110727	VG	O	ND	"
11	RCc3-ZFN10	WT (F67/F96)	20110727	116	O	Positive	"
12	RCc3-ZFN11	WT (F67/F96)	20110727	VG	O	ND	"
13	RCc3-ZFN12	WT (F67/F96)	20110727	VG	O	ND	"
14	RCc3-ZFN13	WT (F67/F96)	20110727	229	O	Positive	"
15	RCc3-ZFN14	WT (F67/F96)	20110804	24	O	ND	"
16	RCc3-ZFN15	WT (F67/F96)	20110804	VG	O	ND	"
17	RCc3-ZFN16	WT (F67/F96)	20110804	127	O	Positive	"
18	RCc3-ZFN17	WT (F67/F96)	20110804	24	O	ND	"
19	RCc3-ZFN18	WT (F67/F96)	20110804	VG	△	ND	"
20	RCc3-ZFN19	WT (F67/F96)	20110804	56	X	Positive	"
21	RCc3-ZFN20	WT (F67/F96)	20110811	VG	O	ND	"
22	RCc3-ZFN21	WT (F67/F96)	20110811	VG	O	ND	"
23	RCc3-ZFN22	WT (F67/F96)	20110824	7	X	ND	"
24	RCc3-ZFN23	WT (F67/F96)	20110824	VG	O	ND	"
25	RCc3-ZFN24	WT (F67/F96)	20110824	3	O	ND	"
26	RCc3-ZFN25	WT (F67/F96)	20110824	VG	O	ND	"
27	RCc3-ZFN26	WT (F67/F96)	20110907	VG	O	ND	"
28	RCc3-ZFN27	WT (F67/F96)	20110907	VG	O	ND	"
29	RCc3-ZFN28	WT (F67/F96)	20110907	VG	O	ND	"

No.	Plant ID	*FokI*type	Transfer date to Greenhouse	No. of Seeds Harvested	PCR detection of ZFN integration	GFP assay for dry seeds	Dates of Cocultivation*
30	RCc3-ZFN29	WT (F67/F96)	20110907	VG	O	ND	"
31	RCc3-ZFN30	WT (F67/F96)	20110907	VG	O	ND	"
32	RCc3-ZFN31	WT (F67/F96)	20110907	VG	O	ND	"
33	RCc3-ZFN32	WT (F67/F96)	20110907	VG	O	ND	"
34	RCc3-ZFN33	WT (F67/F96)	20110907	VG	X	ND	"
35	RCc3-ZFN34	WT (F67/F96)	20110907	VG	O	ND	"
36	RCc3-ZFN35	WT (F67/F96)	20110907	VG	O	ND	"
37	RCc3-ZFN36	WT (F67/F96)	20110907	VG	O	ND	"
38	RCc3-ZFN37	WT (F67/F96)	20110907	VG	O	ND	"
39	RCc3-ZFN38	WT (F67/F96)	20110921	VG	O	ND	"
40	RCc3-ZFN39	WT (F67/F96)	20110928	VG	△	ND	"
41	RCc3-ZFN40	WT (F67/F96)	20110928	VG	O	ND	"
42	RCc3-ZFN41	WT (F67/F96)	20110928	VG	O	ND	"
43	RCc3-ZFN42	WT (F67/F96)	20111011	VG	O	ND	"
44	RCc3-ZFN43	WT (F67/F96)	20111026	VG	O	ND	"
45	RCc3-ZFN44	WT (F67/F96)	20111110	VG	O	ND	
46	PGD1-ZFNDR1	F67-DAS/F96-RR	20110921	VG	O	ND	2011-6/16, 6/23
47	PGD1-ZFNDR2	F67-DAS/F96-RR	20110927	VG	O	ND	"
48	PGD1-ZFNDR3	F67-DAS/F96-RR	20111025	VG	O	ND	"
49	PGD1-ZFNDR4	F67-DAS/F96-RR	20111025	VG	X	ND	"
50	PGD1-ZFNDR5	F67-DAS/F96-RR	2111110	VG	O	ND	"
51	PGD1-ZFNDR6	F67-DAS/F96-RR	2111110	VG	O	ND	"
52	PGD1-ZFNDR7	F67-DAS/F96-RR	20111123	VG	O	ND	"
53	PGD1-ZFNDR8	F67-DAS/F96-RR	20111221	VG	O	ND	
54	PGD1-ZFNRD1	F67-RR/F96-DAS	20110928	VG	O	ND	2011-6/16, 6/23
55	PGD1-ZFNRD2	F67-RR/F96-DAS	20110928	VG	O	ND	"
56	PGD1-ZFNRD3	F67-RR/F97-DAS	20111011	VG	O	ND	"
57	PGD1-ZFNRD4	F67-RR/F97-DAS	20111011	VG	X	ND	
58	PGD1-ZFNRD5	F67-RR/F97-DAS	20111011	VG	O	ND	
59	PGD1-ZFNRD6	F67-RR/F98-DAS	20111025	VG	O	ND	"
60	PGD1-ZFNRD7	F67-RR/F99-DAS	20111025	VG	O	ND	"
61	PGD1-ZFNRD8	F67-RR/F100-DAS	20111110	VG	O	ND	"
62	PGD1-ZFNRD9	F67-RR/F101-DAS	20111123	VG	△	ND	"
63	PGD1-ZFNRD10	F67-RR/F102-DAS	20111123	VG	O	ND	"

No.	Plant ID	*FokI*type	Transfer date to Greenhouse	No. of Seeds Harvested	PCR detection of ZFN integration	GFP assay for dry seeds	Dates of Cocultivation*
64	RCc3-ZFNDR1	F67-DAS/F96-RR	20110907	VG	O	ND	2011-6/16, 6/23
65	RCc3-ZFNDR2	F67-DAS/F96-RR	20110907	VG	O	ND	"
66	RCc3-ZFNDR3	F67-DAS/F96-RR	20110907	VG	O	ND	"
67	RCc3-ZFNDR4	F67-DAS/F96-RR	20110921	VG	△	ND	"
68	RCc3-ZFNDR5	F67-DAS/F96-RR	20110928	VG	O	ND	"
69	RCc3-ZFNDR6	F67-DAS/F96-RR	20110928	VG	O	ND	"
70	RCc3-ZFNDR7	F67-DAS/F96-RR	20110928	VG	O	ND	"
71	RCc3-ZFNDR8	F67-DAS/F96-RR	20110928	VG	O	ND	"
72	RCc3-ZFNDR9	F67-DAS/F96-RR	20110928	VG	O	ND	"
73	RCc3-ZFNDR10	F67-DAS/F96-RR	20110928	VG	O	ND	"
74	RCc3-ZFNDR11	F67-DAS/F96-RR	20111011	VG	X	ND	"
75	RCc3-ZFNDR12	F67-DAS/F96-RR	20111011	VG	O	ND	"
76	RCc3-ZFNDR13	F67-DAS/F96-RR	20111025	VG	O	ND	"
77	RCc3-ZFNDR14	F67-DAS/F96-RR	20111025	VG	O	ND	"
78	RCc3-ZFNDR15	F67-DAS/F96-RR	20111025	VG	O	ND	"
79	RCc3-ZFNDR16	F67-DAS/F96-RR	20111025	VG	O	ND	"
80	RCc3-ZFNDR17	F67-DAS/F96-RR	20111110	VG	△	ND	"
81	RCc3-ZFNDR18	F67-DAS/F96-RR	20111110	VG	ND	ND	"
82	RCc3-ZFNDR19	F67-DAS/F96-RR	20111123	VG	O	ND	"
83	RCc3-ZFNDR20	F67-DAS/F96-RR	20111123	VG	O	ND	"
84	RCc3-ZFNRD1	F67-RR/F96-DAS	20110824	VG	O	ND	2011-6/16, 6/23
85	RCc3-ZFNRD2	F67-RR/F96-DAS	20110824	VG	O	ND	"
86	RCc3-ZFNRD3	F67-RR/F96-DAS	20110824	VG	O	ND	"
87	RCc3-ZFNRD4	F67-RR/F96-DAS	20110907	VG	O	ND	"
88	RCc3-ZFNRD5	F67-RR/F96-DAS	20110907	VG	O	ND	"
89	RCc3-ZFNRD6	F67-RR/F96-DAS	20110907	VG	O	ND	"
90	RCc3-ZFNRD7	F67-RR/F96-DAS	20110907	VG	O	ND	"
91	RCc3-ZFNRD8	F67-RR/F96-DAS	20110907	VG	X	ND	"
92	RCc3-ZFNRD9	F67-RR/F96-DAS	20110921	VG	O	ND	"
93	RCc3-ZFNRD10	F67-RR/F96-DAS	20110921	VG	X	ND	"
94	RCc3-ZFNRD11	F67-RR/F96-DAS	20110921	VG	O	ND	"
95	RCc3-ZFNRD12	F67-RR/F96-DAS	20110921	VG	X	ND	"
96	RCc3-ZFNRD13	F67-RR/F96-DAS	20110928	VG	O	ND	"
97	RCc3-ZFNRD14	F67-RR/F96-DAS	20110928	VG	O	ND	"

No.	Plant ID	*Fok*Itype	Transfer date to Greenhouse	No. of Seeds Harvested	PCR detection of ZFN integration	GFP assay for dry seeds	Dates of Cocultivation*
98	RCc3-ZFNRD15	F67-RR/F96-DAS	20111110	VG	O	ND	"
99	RCc3-ZFNRD16	F67-RR/F96-DAS	20111110	VG	O	ND	"
100	RCc3-ZFNRD17	F67-RR/F96-DAS	20111110	VG	X	ND	"
101	RCc3-ZFNRD18	F67-RR/F96-DAS	20111123	VG	ND	ND	"
102	NT				X	Negative	

PGD1: constututive promoter; RCc3: root-specific promoter; NT: non-transgenic negative control; WT: wild-type; VG: vegetativegrowth phase; X: no detection; O: positive signal; △: hardly detectable; □: partial internal deletion in the ZFN integrate; ND: not determined.

*For each plate about 120 embryogenic calli were co-cultivated with *Agrobacterium*.

요 약

징크핑거뉴클레아제 (zinc finger nuclease; ZFN)는 최근 개발된 강력한 표적돌연변이(targeted mutagenesis) 유발 도구로서 동물시스템에서는 그 유용성이 이미 확인되어 광범위하게 활용되고 있지만, 식물에서는 쌍자엽식물 모델식물인 *Arabidopsis*와 작물인 담배, 옥수수, 콩에서만 그 적용이 보고되어 있을 뿐, 단자엽식물 모델식물인 벼에 대해서는 아직 보고된 바가 없다. ZFN을 이용한 표적특이적 돌연변이 유발과 같은 효과적이고 안정적인 유전자 특이적 돌연변이 유도 방법이 본 연구를 통해 벼에 대해서도 적용될 수 있음을 보여준다면 이는 대부분의 식량작물이 속한 단자엽식물의 육종 및 유전자 기능 연구에 있어서 획기적인 진전이 될 것이다. 한편, 벼의 *OsAP2* 유전자는 AP2 domain을 갖고 있는 전사인자로서, 특히, 뿌리특이적 프로모터에 연결하여 벼에서 발현시켰을 때 생식성장기의 가뭄저항성이 증가하여 등숙률 및 알곡수가 증가하고 그 결과 수확량이 15-25% 증가하는 특성을 보인다 (Oh et al. 2009). 본 연구에서는 단자엽식물의 모델식물인 벼의 *OsAP2* 유전자를 표적돌연변이시킬 목적으로 네 개의 징크핑거 모듈을 가진 ZFN을 모듈어셈블리법(modular assembly method)에 의해 61쌍 합성하였고, 효모 single-strand annealing (SSA) 법을 사용하여 합성된 ZFN 중에서 표적 DNA에 대한 절단 활성이 좋은 4쌍의 ZFN을 선발하였다. 이 중 가장 절단 활성이 좋은 한 쌍의 ZFN에서 징크핑거단백질 도메인을 추출하여 여기에 야생형 *FokI* (F67/F96) 또는 변이형 *FokI* (F67-DAS/F96-RR 및 F67-RR/F96-DAS) 절단도메인과 결합하고, 이들 각각에 전신발현용 프로모터와 뿌리특이적 프로모터를 각각 연결하여 형질전환 벼를 만들어 ZFN을 발현시켰다. 그 결과, 변이형 *Fok*I 도메인 및 전신발현용 프로모터를 장착한 ZFN을 발현하는 형질전환 벼에서 *OsAP2* 유전자의 표적 위치에 *Taq* 중합효소의 오류율 (error rate)을 훨씬 상회하는 빈도의 치환돌연변이가 유발됨을 확인하였다. 이 실험에 사용된 ZFN 유전자는 벼의 코돈

에 최적화시키기 위해 독일소재 DNA합성회사 (Entelechon GmBH)에 주문 합성하였는데, 이 때 한 쌍의 ZFN 유전자를 한 개의 프로모터 작용으로 동시에 같은 양 발현시킬 수 있도록 두 유전자 사이에 2A sequence를 연결하였고, 이렇게 고안 합성된 ZFN 유전자 쌍을 서로 다른 세 쌍의 *FokI* nuclease domain (F67/F96, F67-DAS/F96-RR, F67-RR/F96-DAS)에 연결하였다 (사정상, *FokI* domain 중 F67/F96만 코돈 최적화함). 합성한 세 쌍의 ZFN을 각각 2종의 프로모터 즉, 전신 발현용 프로모터 PGD1 또는 뿌리특이적 프로모터 RCc3가 장착된 벼 형질전환용 운반체 내로 도입시켰다. 이렇게 하여 제작 완료된 총 6종의 운반체를 *Agrobacterium*을 이용하여 벼 캘러스에 형질전환시켰고 형질전환세포로부터 총 100개 식물체를 재분화시켰으며, 이 중 15 개체로부터 T_1종자를 채종하였다. 총 98개 형질전환 벼 식물체에 대해 genomic PCR 분석을 한 결과 형질전환 벼 중 약 80%에서 원하는 ZFN 밴드가 나타났고 다만, 이 중 한 식물체에서 ZFN 유전자 일부에 결실이 있는 것으로 관찰되었다. 일부 식물체로부터 수확한 건조 종자에 대해 GFP 형광단백질 발현을 분석한 결과, PCR 실험에서 부분 결실된 것으로 확인된 식물체 1개를 제외한 분석한 모든 형질전환 벼에서 형광단백질이 발현되었다. 형광단백질 유전자는 ZFN 유전자와 2 kb 거리를 두고 연관되어 있다. 또 immunoblot 분석 결과, 분석한 형질전환 시료 모두에서 ZFN 단백질이 강하게 발현되고 있음을 확인하였다. *OsAO2*유전자 부위에서 실제로 ZFN에 의한 표적 돌연변이가 유발되는지 확인하기 위해 T7 endonuclease (T7E1) assay를 수행한 결과 분석한 시료 중 5개체에서 표적돌연변이가 유발된 증거를 확보하였고, 이 중 한 개체를 선택하여 예상되는 표적부위를 포함하는 약 300 bp를 cloning한 후 염기서열을 분석한 결과 표적 부위에 다수 (모두 6개)의 치환돌연변이가 집중되어 있음을 확인하였다. 이 치환돌연변이 빈도를 *Taq* 중합효소의 오류율과 비교한 결과, *Taq* 중합효소의 내재적 오류율을 훨씬 상회하는 것으로 나타났고, 이는 표적 돌연변이가 실제로 일어났음을 강력히 시사한다고 볼 수 있다. 다만, indel (삽입

/결실)이 관찰되지 않은 것은 off-targeting으로 인한 세포독성으로 인해 선택적 세포 사멸이 일어난 결과로 해석되지만, 좀 더 면밀한 분석과 indel 발견을 위해 T_1및 T_2종자에 대한 유전분석을 할 필요가 있다고 판단된다.

주요어 : 유전자 가위, 징크핑거뉴클레아제, 표적돌연변이, 표적돌연변이 유발, 유전자 편집, 게놈 편집.

학 번 : 8425-3313

감사의 글

이 학위논문은 수많은 사람들의 도움의 결과물입니다.
먼저, 자유로이 실험 설비를 사용하고 소속 연구원들의 도움을 받을 수 있도록 해 주신 서울대 자연과학대학 화학부의 김진수 교수님과 명지대 생명과학정보학부의 김주곤 교수님의 배려가 본 논문의 완성에 결정적이었음을 밝힙니다. 하찮은 질문에도 인내심으로 답해주고 꼼꼼히 실험을 지도해 준, 김진수 교수님 연구실의 조승우 박사과정 학생과 김주곤 교수님 연구실의 정진서 박사님, 그리고 두 연구실의 제게 도움을 준 연구원들 모두에게 깊이 감사드립니다.
재입학하기 전에 박사과정 지도교수님으로 모셨던 박효근 교수님이 정년퇴직하신 후라 난감한 상황이 될 뻔 했던 순간에 지도교수가 되길 기꺼이 승낙하셨던 서울대 농업생명과학대학 식물생산학부 강병철 교수님에게도 깊은 감사드립니다.
30년만에 쓰는 논문이라 어설프기 짝이 없는 논문이었는데도 배려와 인내심으로 최종 출판본과 같이 업그레이드할 수 있도록 꼼꼼히 지도해 주신 Dissertation committee의 최도일 교수님, 강병철 교수님, 김주곤 교수님, 김진수 교수님, 허진회 교수님께 다시 한번 감사드립니다.
이미 고인이 되셨거나 은퇴하셔서 이제는 뵙기 힘든 분들이 되셨지만, 저의 학부 때부터 지도해주시고 음양으로 많은 도움을 주셨던 故 표현구 교수님, 故 염도의 교수님, 대학원 석사, 박사과정의 지도교수님이신 박효근 교수님, 그리고 찾아 뵐 때마다 바쁘신 중에도 소중한 시간을 내어 좋은 말씀을 들려주시곤 했던 고광출 교수님, 이병일 교수님, 김병동 교수님께 마음 깊이 감사드립니다.
새로운 캠퍼스에 재입학하여 서먹해 하는 저를 따뜻하게 맞아주시던 모습이 지금도 생생한 이승구 교수님, 김기선 교수님 그리고 짧지 않은 시간을 버틸 수 있도록 조언과 편달을 아끼지 않았던 대학 동기 및 후배 교수님들에게도 깊이 감사드립니다.
희생과 인내심으로 이 나이되도록 묵묵히 지켜봐 주신 가족, 그리고 경제적으로 어려움을 겪을 때 도움을 주었던 친구들과 선후배님들, 특히 충남대학교의 임용표 교수님과 B & K 특허사무소의 백승남 변리사님께 다시 한번 마음 속 깊이 감사드립니다.