School of Doctoral Studies in Biological Sciences

University of South Bohemia in České Budějovice
Faculty of Science

Gene targeting in Silkworm (*Bombyx mori*) by Engineered Endonucleases

Ph.D. Thesis

MSc. Suresh Sajwan

Supervisor: doc. RNDr. Michal Žurovec, CSc.

University of South Bohemia, Faculty of Science, Department of Genetics, České Budějovice
Biology Centre of the Academy of Sciences of the Czech Republic,
Institute of Entomology, České Budějovice

České Budějovice 2013

This thesis should be cited as:
Sajwan, S, 2013: Gene targeting in Silkworm (*Bombyx mori*) by Engineered Endonucleases. Ph.D. Thesis, University of South Bohemia, Faculty of Science, School of Doctoral Studies in Biological Sciences, České Budějovice, Czech Republic, 65 pp.

Annotation

This thesis describes the establishment of a precise gene targeting methodology in the silkworm *Bombyx mori* by technologies based on engineered endonucleases. Two classes of engineered endonucleases, ZFNs and full length TALENs were used for creating DSBs at specified sites in the colour marker genes (*BmBlos2* and *Bmwh3*). Direct embryo microinjection of engineered nucleases mRNA were performed and let the nuclease proteins to disrupt the functions of these marker genes by creating DSBs and inducing error prone NHEJ mechanism. These experiments showed that both ZFNs and TALENs could be used for targeted gene disruption in silkworms.

Declaration [in Czech]

Prohlašuji, že svoji disertační práci jsem vypracoval samostatně pouze s použitím pramenů a literatury uvedených v seznamu citované literatury.

Prohlašuji, že v souladu s § 47b zákona č. 111/1998 Sb. v platném znění souhlasím se zveřejněním své disertační práce, a to v úpravě vzniklé vypuštěním vyznačených částí archivovaných Přírodovědeckou fakultou elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejích internetových stránkách, a to se zachováním mého autorského práva k odevzdanému textu této kvalifikační práce. Souhlasím dále s tím, aby toutéž elektronickou cestou byly v souladu s uvedeným ustanovením zákona č. 111/1998 Sb. zveřejněny posudky školitele a oponentů práce i záznam o průběhu a výsledku obhajoby kvalifikační práce. Rovněž souhlasím s porovnáním textu mé kvalifikační práce s databází kvalifikačních prací Theses.cz provozovanou Národním registrem vysokoškolských kvalifikačních prací a systémem na odhalování plagiátů.

České Budějovice, 22. 2. 2013 ....

Suresh Sajwan

This thesis originated from a partnership of Faculty of Science, University of South Bohemia, and Institute of Entomology, Biology Centre of the ASCR, supporting doctoral studies in the Molecular and Cellular Biology and Genetics study programme. A substantial part of the research work was conducted at the Silk Material Research Unit, National Institute of Agrobiological Sciences, Tsukuba, Japan, under the supervision of Dr. Yoko Takasu and Dr. Toshiki Tamura.

Financial support

This research was supported by Grant P305/10/2406 from the Grant Agency of the Czech Republic, Grant LC06077 Research Center Program, Ministry of Education, Youth and Sports of the Czech Republic, and KAKENHI Grant Number 23580083 from Japan Society for the Promotion of Science.

Acknowledgements

I would like to thank with sense of gratitude to my supervisor Dr. Michal Zurovec for his valuable guidance. I express my heartfelt regards to our collaborators Drs. Yoko Takasu and Toshiki Tamura for their kind cooperation, sustained interest and constant encouragement during the course of this investigation. I am also thankful to my labmates for their humorous company and cooperation. My sincere appreciation goes to my family and friends for never ending love, affection, heartiest blessing, encouragement and patience. I am also highly thankful to Ministry of Education, Youth and Sports, Czech Republic for Doctoral study scholarship.

List of papers and author's contribution

The thesis is based on the following papers (listed chronologically):

I. Takasu Y, Kobayashi I, Beumer K, Uchino K, Sezutsu H, Sajwan S, Carroll D, Tamura T and Zurovec M. 2010. Targeted mutagenesis in the silkworm *Bombyx mori* using zinc finger nuclease mRNA injection. *Insect Biochem Mol Biol.* 40(10): 759-65. (IF = 3.246)
Suresh Sajwan participated in construction of pENTR-NLS-ZFN vectors for Gateway vector cloning and performed LR cloning to destination vector pCS2-DEST.

II. Sajwan S, Takasu Y, Tamura T, Uchino K, Sezutsu H and Zurovec M. 2013. Efficient disruption of endogenous *Bombyx* gene by TAL effector nucleases. *Insect Biochem Mol Biol.* 43(1): 17-23. (IF = 3.246)
Suresh Sajwan participated in construction of Yeast Gateway destination vectors, pDEST-YcPlac33 and pDEST-YcPlac11; performed yeast single strand annealing assays and in manuscript preparation.

Suresh Sajwan and Yoko Takasu contributed equally to this work.

"I agree to the above statements":

Michal Žurovec …………………………..

Contents

INTRODUCTION 1

REFERENCES 8

LIST OF ABBREVIATIONS 15

RESULTS 16

Part I - Efficient disruption of endogenous *Bombyx* gene by TAL effector nucleases 17
Supplementary informations 24

Part II - Targeted mutagenesis in the silkworm *Bombyx mori* using zinc finger nuclease mRNA injection 29
Supplementary informations 37

CONCLUSIONS 56

APPENDIX 57

Introduction

Gene targeting is a genetic method used for introducing changes in the endogenous gene sequence. These changes include various types of modifications ranging from gene disruption (point mutations, small or large deletions or insertions) to directed gene alteration (insertion of exogenous DNA at a specific gene location). Conventional gene targeting by virtue of homologous recombination with adequate selection procedures and effective screening schemes have been performed for many years in *Saccharomyces cerevisiae* (Orr and Szostak, 1983), murine embryonic stem cells (Mansour *et al*., 1988) and more recently in *Drosophila melanogaster* (Rong and Golic, 2000). Homologous recombination (HR) is utilized by eukaryotic somatic cells to accurately repair Double Stranded Breaks (DSBs) by using homologous chromosome as a template for repair (Szostak *et al*., 1983). Several fold increment of recombination rate between linear homologous DNA and genomic DNA was observed, when DNA double stranded breaks were generated in homologous regions of yeast (Rudin *et al*., 1989; Plessis *et al.,* 1992) and mammalian cells (Rouet *et al*., 1994; Choulika *et al*., 1995). Alternatively, DSBs are also repaired by error prone Non Homologous End Joining (NHEJ) repair pathway often resulting in DNA sequence disruption (insertions or deletions causing frameshift) (Moore and Haber, 1996.).

A long-standing goal of molecular biologists was to develop technologies for more efficient and precise gene targeting. Technologies based on engineered endonucleases have been developed and now engineered enzymes such as engineered meganucleases (Galetto *et al*., 2009), Zinc Finger Nucleases (Bibikova *et al*., 2001), TAL Effector Nucleases (Christian *et al*., 2010) and very recently, RNA guided CRISPRs approach (Cong *et al*., 2013) are available for creating precise DSBs in the genome of organisms of interest.

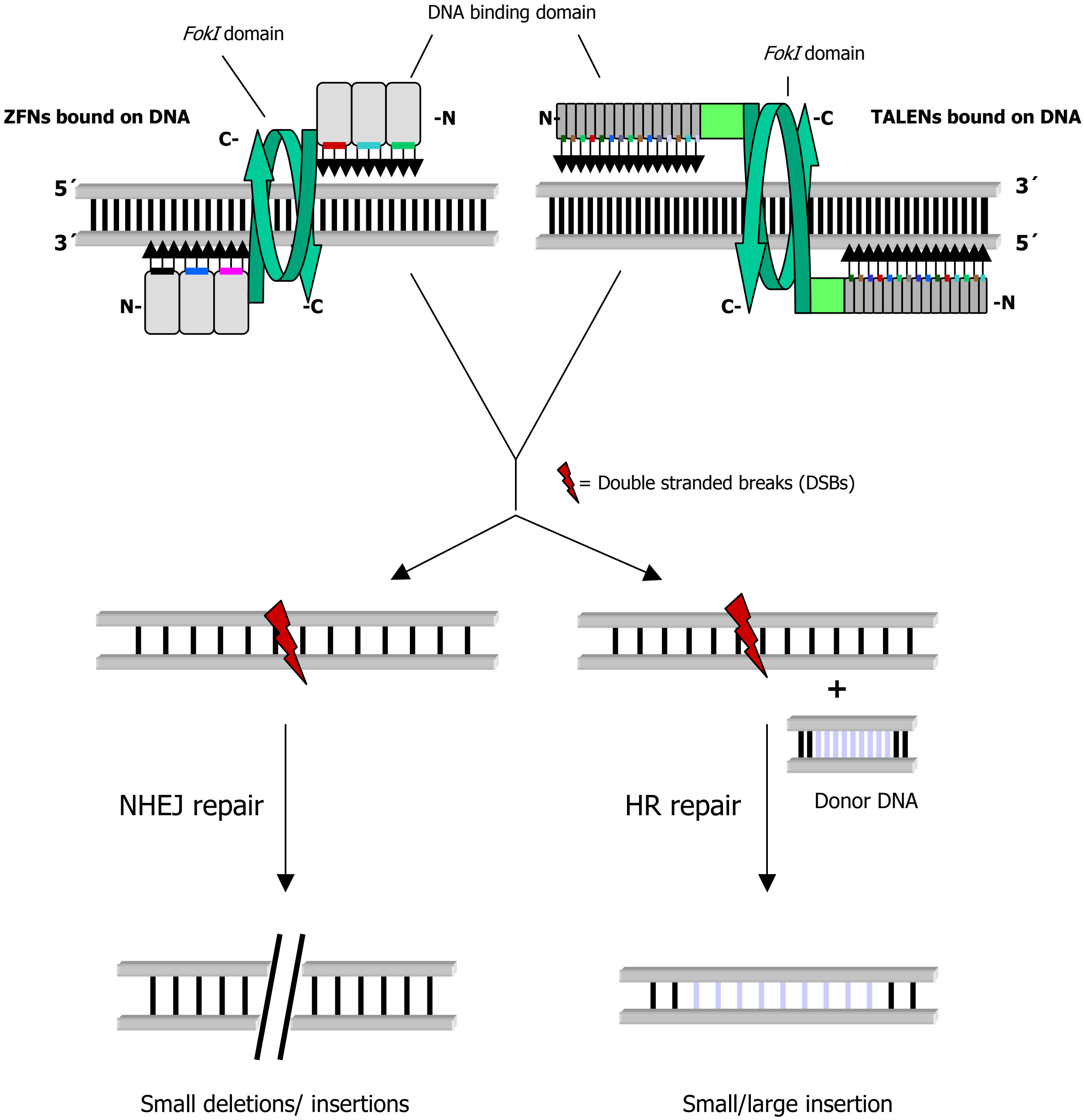
Among these engineered endonucleases, Zinc Finger Nucleases (ZFNs) and Transcription Activation-Like Effector Nucleases (TALENs), have been used widely for creating DNA double stranded breaks at user specified sites in the genome of several organisms including insects (Bibikova *et al*., 2001; Beumer *et al*., 2006; Beumer *et al*., 2008; Ma *et al*., 2012; Watanabe *et al*., 2012; Merlin *et al*., 2013), Xenopus (Lei *et al*., 2012), Zebrafish (Doyon *et al*., 2008; Bedell *et al*., 2012), plants (Townsend *et al*., 2009, Shukla *et al.,* 2009 ; Zhang *et al*., 2013), livestocks (Carlson *et al*., 2012) , rats and mouse (Cui *et al.,* 2011). These nucleases recognize, bind and perform cleavage at the specified target DNA sequence, which is repaired by well conserved repair machinery either NHEJ or HR pathways and have been harnessed for gene knockout or specific modifications. With the availability of

genome sequences of several organisms, these nucleases have become an invaluable tool in reverse genetic studies by disrupting or modifying a specific gene of interest to study its function.

ZFNs and TALENs are hybrid proteins composed of a nonspecific nuclease domain from the *FokI* restriction enzyme and a customized DNA binding domain that are connected by inter-domain linker region. In ZFNs, the DNA binding domain is usually made up of 3-4 Cys2His2 zinc fingers modules, which are derived from three-finger murine transcription factor Zif268 (Jamieson *et al*., 1994; Rebar and Pabo, 1994). Each ZF module has two beta strands and one alpha helix, amino acid residues of the alpha helix at positions –1 to +6, recognize and bind specifically to a DNA triplet by interacting with the major groove of the target DNA (Pavletich and Pabo, 1991). This part of alpha helix, known as the specificity determining residues region is interchangeable to recognize and bind different DNA triplets. Due to the works of Carlos Barbas laboratory of The Scripps Research Institute (Segal *et al.*, 1999; Dreier *et al*., 2001; Dreier *et al*., 2005) and commercial companies e.g. Sangamo Biosciences (Liu *et al.,* 2002) and ToolGen (Bae *et al.,* 2003), large pool of zinc finger modules are available that can recognize most of the 64 triplets. ZFNs with 3-4 zinc fingers specifically recognize and bind consecutive DNA triplets (Pabo *et al.,* 2001) in a linear fashion in the 3' to 5' direction (Figure 1).

The DNA binding domain of TALENs with a simpler DNA binding code was described recently (Christian *et al*., 2010; Bogdanove *et al*., 2010; Miller *et al*., 2011). This DNA-binding domain originates from bacterial transcription factors of the Transcription Activator-Like (TAL) Effector family (Boch *et al*., 2009; Moscou and Bogdanove, 2009). These are highly conserved proteins secreted by a number of plant pathogenic bacteria of the genus *Xanthomonas* into plant cells as a part of the virulence process. The proteins bind to promoter sequences in the host plant and activate specific genes that cause disease or trigger a host resistance response. TAL effectors consist of a C-terminus bearing nuclear localization signal together with a transcription activation domain, a central repetitive domain responsible for DNA binding and the N-terminus encoding sequences needed for secretion (Szurek *et al*., 2002; Marois *et al*., 2002). The central sequence contains a varying number of modular repeated units of 34 (sometimes also 33 or 35) amino acid residues called TAL effector repeat domains. These domains are largely conserved except for amino acids 12 and 13, so called “repeated variable di-residues” (RVDs), in each repeat. The last TAL effector repeat is shorter (a “half-repeat”). The RVDs are crucial for recognition of a single DNA base and display the

Figure 1. Diagram of a pair of ZFNs and TALENs bound to their target sites. N and C denote the N terminus and C terminus of each protein. The bindings of ZFNs and TALENs relative to their target sites have been shown. ZFNs bind in the 3' to 5' direction whereas TALENs binding occur in the 5' to 3' direction. Coloured bar represents the variable region in each modular repeat of DNA binding domain. DSBs activate NHEJ or HR repair pathways that can be utilized for one of the several outcomes (small insertion or deletions, directed small or large insertions guided by exogenous donor DNA).



following preferences: HD binds specifically to C, NG to T, NI to A and NN to G (Boch *et al*., 2009; Moscou and Bogdanove, 2009). The TAL effector repeat domains can be used to assemble 13-28 repeat units in any predetermined order (Boch and Bonas, 2010). These tandem repeat units in the DNA binding domain recognize and bind each DNA base specifically in one to one fashion in the 5' to 3' direction (Figure 1). In addition, the activity of a TAL effector requires a T nucleotide, which precedes the nucleotide sequence specified by the RVDs (Boch *et al.,* 2009; Moscou and Bogdanove, 2009).

TALENs are analogous to ZFNs, in that the DNA-binding domain of the TAL effector is fused to the catalytic domain of *FokI* nuclease. Requirement of the dimerization of *FokI* nuclease domain to create DSBs makes these engineered nucleases to be designed and constructed always in pairs (Smith *et al.,* 2000; Mani *et al.,* 2005). Each nuclease pair binds the DNA target sites adjacent of a spacer region, leaving the *FokI* domains to dimerize and cleave the DNA in the spacer region. The optimal spacer region for ZFNs is 5-6 bp (Bibikova *et al.,* 2001; Urnov *et al.,* 2005) whereas for TALENs the spacer region ranges from 12-30 bp because of the longer amino acid sequence in the inter-domain linker region (Christian *et al.,* 2010).

Several software tools and engineering platforms are easily accessible for constructing DNA binding domains of these engineered endonucleases to induce cleavage in a user defined genomic sequence. The websites of Carlos Barbas tool (zincfingertool.org)(Mandell and Barbas, 2006) and Zinc finger consortium (http://zifit.partners.org/ZiFiT/)(Sander *et al.,* 2010) help in locating a suitable target sites and for designing corresponding DNA binding domains of a ZFNs, whereas methods like Modular Assembly method (Carroll *et al*., 2006; Wright *et al*., 2006), Oligomerised Pool Engineering method (OPEN)(Maeder *et al*., 2009) and Context dependent Assembly method (CoDA)(Sander *et al*., 2011a) are used for constructing the DNA binding domain of ZFNs. For TALENs, Zinc finger consortium website (http://zifit.partners.org/ZiFiT/Disclaimer.aspx) (Reyon *et al*., 2012a) and TAL Effector Nucleotide Targeter 2.0 (http://tale-nt.cac.cornell.edu/) (Doyle *et al.,* 2012) are used for target finding and designing the DNA binding domain. In addition procedures like Golden gate assembly methods (Li *et al*., 2011; Cermak *et al.,* 2011; Zhang *et al*., 2011; Sanjana *et al*., 2012) Standard cloning methods (Huang *et al*., 2011 and Sander *et al*., 2011b) and Solid phase assembly methods (Reyon *et al*., 2012b and Briggs *et al*., 2012) are available for construction. Some commercial vendors also provide ready to use engineered nucleases e.g. custom ZFNs from Sangamo Biosciences (Richmond, CA, USA) and custom TALENs from

Cellectis Bioresearch (Paris, France) and Transposagen Biopharmaceuticals (Lexington, KY, USA) but at considerable cost.

In principle, since each zinc finger recognizes a DNA triplet, a ZFN pair with three zinc fingers should specifically recognize and bind 23-24 bp of DNA target sequence together with spacer region, long enough to ensure specificity in any genome. However, published studies have reported that not all engineered DNA binding domains worked well as ZFNs (Miller *et al*., 2007; Cornu *et al*., 2008; Porteus and Baltimore, 2008). Whereas the widespread use of Modular Assembly approach has been hampered by low success rate (Ramirez *et al*., 2008), target site limitations, labour intensive and lengthy procedures of OPEN and CoDA methods for active ZFN selection hinder their use in research community. In spite of the progress in ZFN construction some of these ZFNs had been observed with complete absence of cleavage activity while others were associated with either low or high genomic toxicity by creating DSBs at closely related sequences. These failures are likely caused by each of specificity i.e. discrimination between closely related DNA sequences and cleavage activity of ZFNs bound to DNA in a eukaryotic environment.

TALEN proteins represent a better gene-targeting reagent than ZFNs as they have superior targeting range and are much simpler to construct with higher specificity and nuclease activity. There are very loose DNA sequence constraints for TALENs binding; therefore, TAL binding sites are much more frequent than ZFN targets and occur at least once in every 35 bp (Cermak *et al*., 2011) thus has superior targeting range to ZFN which cuts once in every 500 bp (Sander *et al*., 2011a). TALENs are more specific as they bind to longer target sequence than ZFNs. In addition, the individual RVD module associates with individual nucleotide independently, free from target site overlap or context dependent effects that together impart high specificity to TALENs in any complex genome. The ability to discriminate among closely related DNA sequences in plants suggests that TALENs are highly specific (Zhang *et al*., 2013). It also allows a simple, fast and inexpensive modular assembly strategy for TALEN construction, which usually takes around 5 days (Cermak *et al*., 2011) and can be performed in most molecular biology laboratories. The nuclease activity of TALENs was reported to be dramatically dependent on the optimal truncation of TAL-effector part of the molecule (Hockemeyer *et al*., 2011; Miller *et al*., 2011; Wood *et al*., 2011). The N-terminal TALEN truncations seem to be limited to 152 N-terminal amino acid residues (NΔ152) preceding the TAL-effector repeats, which were shown by (Miller *et al*., 2011) to be essential for full DNA binding activity. The C-terminal TALEN truncations showed high activity up to 18 amino acid residues (C+18) following the TAL-effector repeats

(Miller *et al*., 2011; Mussolino *et al*., 2011). The C-terminal residues form a linker between TAL-effector DNA binding domain and *FokI* nuclease domain. The TALENs with shorter linkers were shown to have optimal efficiency for DNA targets with half-sites separated with shorter spacers (Christian *et al*., 2012). The success rate of functional TALENs is more than 90% (Cermak *et al.*, 2011) and have been proven less toxic to cells than ZFN even when fused with wild-type *Fok1* nuclease domain (Mussolino *et al*., 2011).

Silkworm (*Bombyx mori*), being an economically important insect for silk production for ages, is fast becoming a model organism for Lepidopteran insects and has been used to address fundamental questions on genetics (Goldsmith *et al*., 2005), developmental physiology (Li *et al*., 2006) and developmental biology (Zhong *et al*., 2005; Sato *et al*., 2008). Complete genomic sequence of *B. mori* (Mita *et al*., 2004; Xia *et al*., 2004; Silkworm Genome Consortium, 2008) has also started the possibility of producing genetically engineered silkworms to produce recombinant silk proteins for biomedical and industrial applications.

The first attempt for gene targeting in silkworm had been done by using baculovirus, where the chimeric *L-chain–GFP* gene was stably integrated and transmitted into the genome by homologous recombination (Yamao *et al*., 1999). A revolutionary approach to generate stable germline transformation of the silkworm was done by using a *piggyBac* transposon-derived vector (Tamura *et al*., 2000). Since than, many systems were developed for generating transgenic silkworms based on *piggyBac* transposons. The *piggyBac* transposon facilitated strategies like GAL4/UAS (Imamura *et al*., 2003), enhancer trapping (Uchino *et al*., 2008), and FLP Recombinase-Mediated Site-Specific Recombination (Long *et al*., 2012) systems for spatiotemporal expression studies, gene identification and genome manipulation, respectively. RNAi (Quan *et al*., 2002 and Tomita *et al*., 2009) was also used for gene silencing. Low transforming frequency, limited carrying capacity, random integration of *piggyBac* transposons (Fraser, 2012) and incomplete gene silencing by RNAi, limit the broad applicability of these gene-targeting reagents in silkworm.

In this research, we therefore established a precise gene targeting method in silkworm by using ZFNs and TALENs. In the first part, ZFNs were used to assess its application for gene targeting in silkworm. ZFNs construction by Modular assembly method, *in vitro* transcription and direct embryo microinjection of customized ZFNs mRNA are described to disrupt functions of two epidermal colour marker genes (*BmBlos2 and Bmwh3*). Genetic analysis of mutation spectrum caused by NHEJ is closely discussed. Although the success rate

of ZFNs have been reported low, the somatic as well as germline mutations recovered showed that ZFNs can be utilized for generating gene specific mutations in *Bombyx* germline.

In the second part, the potential of using full length TALENs for site-specific mutagenesis in silkworm is discussed. Construction of full length TALEN architecture by Golden gate assembly is described. Advantage of using TALENs over ZFNs and genetic analysis of mutation spectrum caused by NHEJ are discussed. A yeast assay is also described for the selection of functional engineered nucleases with high cleavage activity and low toxicity. All TALENs were able to generate somatic and germline mutations albeit at low frequency.

References

Bae KH, Kwon YD, Shin HC, Hwang MS, Ryu EH, Park KS *et al*., 2003. Human zinc fingers as building blocks in the construction of artificial transcription factors. *Nat Biotechnol*. 21(3):275-80.

Bibikova M, Carroll D, Segal DJ, Trautman JK, Smith J, Kim YG and Chandrasegaran S. 2001.Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Mol Cell Biol.* 21(1):289-97.

Bedell VM, Wang Y, Campbell JM, Poshusta TL, Starker CG, Krug RG *et al*., 2012. In vivo genome editing using a high-efficiency TALEN system. *Nature*. 491(7422):114-8.

Beumer K, Bhattacharyya G, Bibikova M, Trautman JK and Carroll D. 2006. Efficient Gene Targeting in *Drosophila* With Zinc-Finger Nucleases. *Genetics.* 172(4): 2391-403.

Beumer K, Trautman JK, Bozas A, Liu JL, Rutter J, Gall JG and Carroll D. 2008. Efficient gene targeting in *Drosophila* by direct embryo injection with zinc-finger nucleases. *Proc Natl Acad Sci U S A*. 105(50):19821-6.

Boch J and Bonas U. 2010. Xanthomonas AvrBs3 family-type III effectors: discovery and function. *Annu Rev Phytopathol.* 48:419-36.

Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A and Bonas U. 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science.* 326(5959):1509-12.

Bogdanove AJ, Schornack S and Lahaye T. 2010. TAL effectors: finding plant genes for disease and defense. *Curr Opin Plant Biol.* 13(4):394-401.

Briggs AW, Rios X, Chari R, Yang L, Zhang F, Mali P and Church GM. 2012. Iterative capped assembly: rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers. *Nucleic Acids Res*. 40(15):e117.

Carlson DF, Tan W, Lillico SG, Stverakova D, Proudfoot C *et al*., 2012. Efficient TALEN-mediated gene knockout in livestock. *Proc Natl Acad Sci U S A*.109 (43): 17382-7.

Carroll D, Morton JJ, Beumer KJ and Segal DJ. 2006. Design, construction and in vitro testing of zinc finger nucleases. *Nature Protoc.* 1(3): 1329-1341.

Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C *et al*., 2011. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* 39(12): e82.

Choulika A., Perrin A., Dujon B. and Nicolas J.-F. 1995. Induction of homologous recombination in mammalian chromosomes by using the I-*Sce*I system of *Saccharomyces cerevisiae. Mol. Cell. Biol.* 15(4): 1968–1973.

Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A, Bogdanove AJ and Voytas DF. 2010. Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics.* 186(2):757-61.

Christian ML, Demorest ZL, Starker CG, Osborn MJ, Nyquist MD, Zhang Y *et al.*, 2012. Targeting G with TAL effectors: a comparison of activities of TALENs constructed with NN and NK repeat variable di-residues. *PLoS One*. 7(9):e45383.

Cong L, Ran FA, Cox D, Lin S, Barretto R *et al.*, 2013. Multiplex Genome Engineering using CRISPR/Cas Systems. *Science*. [Epub ahead of print]

Cornu TI, Thibodeau-Beganny S, Guhl E, Alwin S, Eichtinger M, Joung JK and Cathomen T. 2008.DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases. *Mol Ther.* 16(2): 352-8.

Cui X, Ji D, Fisher DA, Wu Y, Briner DM and Weinstein EJ. 2011. Targeted integration in rat and mouse embryos with zinc-finger nucleases. *Nat Biotechnol*. 29(1): 64-7.

Doyle E.L., Booher N.J., Standage D.S., Voytas D.F., Brendel V.P., VanDyk, J.K., and Bogdanove, A.J. 2012. TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. *Nucleic Acids Res*. 40 (Web Server issue):W117-22.

Doyon Y, McCammon JM, Miller JC, Faraji F, Ngo C, Katibah GE *et al.*, 2008. Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. *Nat Biotechnol.* 26(6): 702-8.

Dreier B, Beerli RR, Segal DJ, Flippin JD and Barbas CF 3rd. 2001. Development of zinc finger domains for recognition of the 5'-ANN-3' family of DNA sequences and their use in the construction of artificial transcription factors. *J Biol Chem*. 276(31):29466-78.

Dreier B, Fuller RP, Segal DJ, Lund CV, Blancafort P, Huber A, Koksch B and Barbas CF 3rd. 2005. Development of zinc finger domains for recognition of the 5'-CNN-3' family DNA sequences and their use in the construction of artificial transcription factors. *J Biol Chem*. 280(42): 35588-97.

Fraser MJ Jr. 2012. Insect transgenesis: current applications and future prospects. *Annu Rev Entomol*. 57:267-89.

Galetto R, Duchateau P and Pâques F. 2009. Targeted approaches for gene therapy and the emergence of engineered meganucleases. *Expert Opin Biol Ther.* 9(10):1289-303.

Goldsmith MR, Shimada T and Abe H. 2005. The genetics and genomics of the silkworm, *Bombyx mori*. *Annu Rev Entomol*. 50:71-100.

Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP *et al*., 2011. Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotechnol.* 29(8): 731-4.

Huang P, Xiao A, Zhou M, Zhu Z, Lin S and Zhang B. 2011. Heritable gene targeting in zebrafish using customized TALENs. *Nat Biotechnol*. 29(8):699-700.

Imamura, M., Nakai, J., Inoue, S., Quan, G.X., Kanda, T. and Tamura, T. 2003. Targeted gene expression using the GAL4/UAS system in the silkworm *Bombyx mori*. *Genetics.* 165(3): 1329-1340.

International Silkworm Genome Consortium, 2008. The genome of a lepidopteran model insect, the silkworm *Bombyx mori*. *Insect Biochem Mol Biol*. 38(12): 1036-45.

Jamieson AC, Kim SH and Wells JA.1994. In vitro selection of zinc fingers with altered DNA-binding specificity. *Biochemistry*. 33(19): 5689-95.

Kim YG, Cha J and Chandrasegaran S.1996. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc Natl Acad Sci U S A*. 93(3): 1156-60.

Lei Y, Guo X, Liu Y, Cao Y, Deng Y, Chen X, Cheng CH, Dawid IB, Chen Y and Zhao H. 2012. Efficient targeted gene disruption in *Xenopus* embryos using engineered transcription activator-like effector nucleases (TALENs). *Proc Natl Acad Sci USA*. 109(43): 17484-9.

Li T, Huang S, Jiang WZ, Wright D, Spalding MH, Weeks DP and Yang B. 2011. TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Res.* 39(1): 359-372.

Li XH, Wu XF, Yue WF, Liu JM, Li GL and Miao YG. 2006. Proteomic analysis of the silkworm (*Bombyx mori* L.) hemolymph during developmental stage. *J Proteome Res.* 5(10): 2809-14.

Liu J, Li C, Yu Z, Huang P, Wu H, Wei C, Zhu N, Shen Y, Chen Y, Zhang B, Deng WM and Jiao R. 2012. Efficient and Specific Modifications of the *Drosophila* Genome by means of an Easy TALEN Strategy. *J Genet Genomics*. 39(5):209-15.

Liu Q, Xia Z, Zhong X and Case CC. 2002. Validated zinc finger protein designs for all 16 GNN DNA triplet targets. *J Biol Chem*. 277(6): 3850-6.

Long DP, Zhao AC, Chen XJ, Zhang Y, Lu WJ, Guo Q, Handler AM and Xiang ZH.2012. FLP recombinase-mediated site-specific recombination in silkworm, *Bombyx mori*. *PLoS One*. 7(6): e40150.

Mandell JG and Barbas CF 3rd. 2006. Zinc Finger Tools: custom DNA-binding domains for transcription factors and nucleases. *Nucleic Acids Res*. 34(Web Server issue): W516-23.

Ma S, Zhang S, Wang F, Liu Y, Liu Y, Xu H, Liu C, Lin Y, Zhao P and Xia Q. 2012. Highly efficient and specific genome editing in silkworm using custom TALENs. *PloSOne*. 7(9): e45035.

Maeder ML, Thibodeau-Beganny S, Sander JD, Voytas DF and Joung JK. 2009. Oligomerized Pool ENgineering (OPEN): An "Open-Source" Protocol for Making Customized Zinc Finger Arrays. *Nature Protocols.* 4(10): 1471-1501.

Mani M, Smith J, Kandavelou K, Berg JM and Chandrasegaran S. 2005. Binding of two zinc finger nuclease monomers to two specific sites is required for effective double-strand DNA cleavage. *Biochem Biophys Res Commun*. 334(4): 1191-7.

Mansour SL, Thomas KR and Capecchi MR. 1988. Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. *Nature*. 336(6197): 348-52.

Marois E, Van den Ackerveken G and Bonas U. 2002. The xanthomonas type III effector protein AvrBs3 modulates plant gene expression and induces cell hypertrophy in the susceptible host. *Mol Plant Microbe Interact.* 15(7): 637-46.

Merlin C, Beaver LE, Taylor OR, Wolfe SA and Reppert S.M. 2013. Efficient targeted mutagenesis in the monarch butterfly using zinc finger nucleases. *Genome Res.* 23(1):159-68.

Miller JC, Holmes MC, Wang J, Guschin DY, Lee YL, Rupniewski I *et al.* 2007. An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol*. 25(7): 778-85.

Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, Meng X *et al.,*2011. A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol*. 29(2): 143-8.

Mita K, Kasahara M, Sasaki S, Nagayasu Y, Yamada T, Kanamori H *et al*., 2004. The genome sequence of silkworm, *Bombyx mori. DNA Res*.11 (1): 27–35.

Moscou MJ and Bogdanove AJ. 2009. A simple cipher governs DNA recognition by TAL effectors. *Science.* 326(5959): 1501.

Moore JK and Haber JE. 1996. Cell cycle and genetic requirements of two pathways of nonhomologous end-joining repair of double-strand breaks in *Saccharomyces cerevisiae*. *Mol Cell Biol.* 16(5): 2164-73.

Mussolino C, Morbitzer R, Lütge F, Dannemann N, Lahaye and T Cathomen T. 2011. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res.* 39(21): 9283-93.

Pabo CO, Peisach E and Grant RA. 2001. Design and selection of novel Cys2His2 zinc finger proteins. *Annu Rev Biochem*. 70: 313-40.

Pavletich NP and Pabo CO. 1991. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. *Science*. 252(5007): 809-17.

Plessis A., Perrin A., Haber J. E. and Dujon B. 1992.Site-specific recombination determined by I-*SceI*, a mitochondrial group I intron-encoded endonuclease expressed in the yeast nucleus. *Genetics.* 130(3): 451–460.

Porteus MH and Baltimore D.2003. Chimeric Nucleases Stimulate Gene Targeting in human cells. *Science*. 300(5620): 763.

Orr, W. T. and Szostak, J. W. 1983. Yeast recombination: the association between double-strand gap repair and crossingover. *Proc. Natl Acad. Sci. USA* **80**, 4417–4421.

Quan, G.X., Kim, I., Komoto, N., Sezutsu, H., Ote, M., Shimada, T *et al*., 2002. Characterization of the kynurenine 3-monooxygenase gene corresponding to the white egg 1 mutant in the silkworm *Bombyx mori*. *Mol. Genet. Genomics.* 267(1): 1-9.

Ramirez CL, Foley JE, Wright DA, Müller-Lerch F, Rahman SH *et al*., 2008. Unexpected failure rates for modular assembly of engineered zinc fingers.*Nat Methods*. 5(5): 374-5.

Rebar EJ and Pabo CO.1994. Zinc finger phage: affinity selection of fingers with new DNA-binding specificities. *Science*. 263(5147): 671-3.

Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD and Joung JK. 2012b. FLASH assembly of TALENs for high-throughput genome editing. *Nat Biotechnol*. 30(5): 460-5.

Reyon D, Khayter C, Regan MR, Joung JK, Sander JD. 2012a. Engineering designer transcription activator-like effector nucleases (TALENs) by REAL or REAL-Fast assembly. *Curr Protoc Mol Biol*. Chapter 12:Unit 12.15.

Rong YS and Golic KG. 2000. Gene targeting by homologous recombination in Drosophila. *Science*. 288(5473): 2013-8.

Rouet P., Smih F. and Jasin M. 1994. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol. Cell. Biol.* 14(12): 8096–8106.

Rudin N., Sugarman E. and Haber J. E. 1989. Genetic and physical analysis of double-strand break repair and recombination in *Saccharomyces cerevisiae. Genetics* 122(3): 519–534.

Sander JD, Maeder ML, Reyon D, Voytas DF, Joung JK and Dobbs D. 2010. ZiFiT (Zinc Finger Targeter): an updated zinc finger engineering tool. *Nucleic Acids Res*. 38(Web Server issue):W462-8.

Sander JD, Dahlborg EJ, Goodwin MJ, Cade L, Zhang F, Cifuentes D *et al.,* 2011a. Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). *Nat Methods.* 8(1): 67–69.

Sander JD, Cade L, Khayter C, Reyon D, Peterson RT, Joung JK and Yeh JR. 2011b. Targeted gene disruption in somatic zebrafish cells using engineered TALENs. *Nat Biotechnol*. 29(8):697-8.

Sanjana NE, Cong L, Zhou Y, Cunniff MM, Feng G and Zhang F. 2012. A transcription activator-like effector toolbox for genome engineering. *Nature Protoc*.7(1):171-92.

Sato K, Matsunaga TM, Futahashi R, Kojima T, Mita K, Banno Y and Fujiwara H. 2008. Positional cloning of a *Bombyx* wingless locus flugellos (fl) reveals a crucial role for fringe that is specific for wing morphogenesis. *Genetics*. 179(2): 875-85.

Segal DJ, Dreier B, Beerli RR, and Barbas CF 3rd. 1999. Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5'-GNN-3' DNA target sequences. *Proc Natl Acad Sci U S A*. 96(6): 2758-63.

Shukla VK, Doyon Y, Miller JC, DeKelver RC, Moehle EA et al., 2009. Precise genome modification in the crop species Zea mays using zinc-finger nucleases. *Nature*. 459(7245):437-41.

Smith J, Bibikova M, Whitby FG, Reddy AR, Chandrasegaran S and Carroll D. 2000. Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. *Nucleic Acids Res*. 28(17): 3361-9.

Szostak, J. W., Orr, W. T., Rothstein, R. J. and Stahl, F. W.1983 The double-strand-break repair model for recombination. *Cell.* 33: 25–35.

Szurek B, Rossier O, Hause G and Bonas U. 2002. Type III-dependent translocation of the Xanthomonas AvrBs3 protein into the plant cell. *Mol Microbiol*. 46(1): 13-23.

Tamura T, Thibert C, Royer C, Kanda T, Abraham E, Kamba M *et al.,* 2000. Germline transformation of the silkworm *Bombyx mori* L. using a *piggyBac* transposon-derived vector. *Nat. Biotechnol*. 18(1): 81–84.

Tomita, S. and Kikuchi, A. 2009. Abd-B suppresses lepidopteran proleg development in posterior abdomen. *Dev. Biol*. 328(2): 403-9.

Townsend JA, Wright DA, Winfrey RJ, Fu F, Maeder ML, Joung JK and Voytas DF. 2009. High-frequency modification of plant genes using engineered zinc-finger nucleases. *Nature*. 459(7245):442-5.

Uchino K, Sezutsu H, Imamura M, Kobayashi I, Tatematsu K, Iizuka T, Yonemura N, Mita K and Tamura T. 2008. Construction of a piggyBac-based enhancer trap system for the analysis of gene function in silkworm *Bombyx mori*. *Insect Biochem Mol Biol.* 38(12):1165-73.

Urnov FD, Miller JC, Lee YL, Beausejour CM, Rock JM, Augustus S *et al*., 2005. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature*. 435(7042): 646-51.

Watanabe T, Ochiai H, Sakuma T, Horch HW, Hamaguchi N, Nakamura T *et al*., 2012. Non-transgenic genome modifications in a hemimetabolous insect using zinc-finger and TAL effector nucleases. *Nat Commun*. 3: 1017.

Wood AJ, Lo TW, Zeitler B, Pickle CS, Ralston EJ, Lee AH *et al*., 2011. Targeted genome editing across species using ZFNs and TALENs. *Science.* 333(6040): 307.

Wright DA, Thibodeau-Beganny S, Sander JD, Winfrey RJ, Hirsh AS *et al*., 2006. Standardized reagents and protocols for engineering zinc finger nucleases by modular assembly. *Nature Protoc.* 1(3): 1637 – 1652.

Xia Q, Zhou Z, Lu C, Cheng D, Dai F, Li B, Zhao P, Zha X, Cheng T *et al*., 2004. A draft sequence for the genome of the domesticated silkworm (*Bombyx mori*). *Science*. 306(5703): 1937–1940.

Yamao M, Katayama N, Nakazawa H, Yamakawa M, Hayashi Y, Hara S, Kamei K and Mori H. 1999. Gene targeting in the silkworm by use of a baculovirus. *Genes & Dev.* 13(5): 511-516.

Zhang F, Cong L, Lodato S, Kosuri S, Church GM and Arlotta P. 2011. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat Biotechnol*. 29(2): 149-153.

Zhang Y, Zhang F, Li X, Baller JA, Qi Y, Starker CG, Bogdanove AJ and Voytas DF. 2013. Transcription activator-like effector nucleases enable efficient plant genome engineering. *Plant Physiol.* 161(1): 20-7.

Zhong BX, Li JK, Lin JR, Liang JS, Su SK, Xu HS, Yan HY, Zhang PB and Fujii H. 2005. Possible effect of 30K proteins in embryonic development of silkworm *Bombyx mori*. *Acta Biochim Biophys Sin (Shanghai).* 37(5): 355-61.

LIST OF ABBREVIATIONS

BmBlos2	-	*Bombyx mori* homologue of the human BLOS2 *gene*
Bmwh3	-	*Bombyx mori* white gene
CoDA	-	Context Dependent Assembly method
CRISPRs	-	Clustered Regularly Interspaced Short Palindromic Repeats
DSBs	-	Double Stranded Breaks
HR	-	Homologous Recombination
NHEJ	-	Non Homologous End Joining
OPEN	-	Oligomerized Pool ENgineering method
RNAi	-	RNA interference
RVDs	-	Repeated Variable Di-Residues
TAL Effector	-	Transcription Activator-Like (TAL) Effector
TALENs	-	Transcription Activation-Like Effector Nucleases
ZFNs	-	Zinc Finger Nucleases

RESULTS

Part I

Efficient disruption of endogenous *Bombyx* gene by TAL effector nucleases

Insect Biochemistry and Molecular Biology 43 (2013) 17–23

Contents lists available at SciVerse ScienceDirect

Insect Biochemistry and Molecular Biology

journal homepage: www.elsevier.com/locate/ibmb

Efficient disruption of endogenous *Bombyx* gene by TAL effector nucleases

Suresh Sajwan[a,1], Yoko Takasu[b,1], Toshiki Tamura[b], Keiro Uchino[b], Hideki Sezutsu[b], Michal Zurovec[a,*]

[a] *Institute of Entomology, Biology Centre ASCR, and Faculty of Science, University of South Bohemia, 37005 Ceske Budejovice, Czech Republic*
[b] *National Institute of Agrobiological Sciences, 1-2 Owashi, Tsukuba, Ibaraki 305-8634, Japan*

ARTICLE INFO

Article history:
Received 31 August 2012
Received in revised form
18 October 2012
Accepted 19 October 2012

Keywords:
Gene targeting
ZFN
Nonhomologous end joining
MMEJ
BmBLOS2
Chimeric nucleases

ABSTRACT

Engineered nucleases are proteins that are able to cleave DNA at specified sites in the genome. These proteins have recently been used for gene targeting in a number of organisms. We showed earlier that zinc finger nucleases (ZFNs) can be used for generating gene-specific mutations in *Bombyx mori* by an error-prone DNA repair process of non-homologous end joining (NHEJ). Here we test the utility of another type of chimeric nuclease based on bacterial TAL effector proteins in order to induce targeted mutations in silkworm DNA. We designed three TAL effector nucleases (TALENs) against the genomic locus *BmBLOS2*, previously targeted by ZFNs. All three TALENs were able to induce mutations in silkworm germline cells suggesting a higher success rate of this type of chimeric enzyme. The efficiency of two of the tested TALENs was slightly higher than of the successful ZFN used previously. Simple design, high frequency of candidate targeting sites and comparable efficiency of induction of NHEJ mutations make TALENs an important alternative to ZFNs.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Bombyx mori is one of the most economically important insects. Its large size, possibility of mass rearing and inability to survive without human care endow *Bombyx* with great potential for biotechnological applications representing a unique protein expression system for pharmacological proteins, modified silk and other value-added products. With well established genetics, genomics, and physiology *Bombyx* has become an important model for functional studies of lepidopteran species. The silkworm was the first insect other than *Drosophila* in which gene targeting by zinc finger nucleases was shown to work (Takasu et al., 2010). Further progress in the research of *Bombyx* is dependent on the development of additional efficient methods of functional gene analysis.

Zinc finger nucleases are artificial chimeric proteins allowing targeted gene inactivation in a wide range of organisms. They are composed of a Zinc finger transcription factor DNA binding domain and a nonspecific Fok I nuclease domain (Kim et al., 1996). The DNA binding domain usually comprises an array of 3–4 zinc fingers (ZFs), each of which recognizes 3 base pairs in DNA targets. The ZFNs function in pairs which are held together by the dimerization of Fok I nuclease domains and bind opposing targets on DNA separated by a 5–6 bp spacer (Alwin et al., 2005). A major technical advantage is the ZFs can be custom-modified to build nucleases with novel cleavage specificity; however, ZF modules are not equally reliable or efficient in building effective nucleases, and for a number of nucleotide triplets there are no known ZF modules (Carroll et al., 2006). In addition, 'context-dependent' interactions between individual fingers in a zinc-finger array complicate the use of a simple modular strategy for the production of novel sequence-specific ZFNs (Mussolino et al., 2011).

Several novel methods which take into account the ZF context significantly increased the success rate of newly designed ZFNs. They start with randomized or partially randomized zinc finger domains and use various selection schemes in phages, bacteria or yeasts to maximize their DNA binding specificity and affinity (Gupta et al., 2012; Maeder et al., 2008; Pearson, 2008; Sander et al., 2011). Such selection of new ZFNs is, however, expensive and time consuming. A proprietary technology is also available for developing ZFNs which may give a higher success rate (Pearson, 2008; Urnov et al., 2010). Despite the progress in designing new ZFNs, there are still serious limitations in the diversity of target DNA. The availability of targetable DNA sites for ZFNs for publicly available ZFN platforms is around 1 per 500 nt (Sander et al., 2011). It is therefore hard to use ZFNs especially for smaller genes or coding sequences with lower GC content (Isalan et al., 1997).

Abbreviations: NHEJ, nonhomologous end joining; SD-MMEJ, synthesis-dependent microhomology-mediated end joining; HR, homologous recombination; ssa, single-strand annealing; ZF, zinc finger; ZFN, zinc finger nuclease; TALEN, TAL effector nucleases; RVD, repeated variable di-residue; BLT-1, TALEN target 1 in *Bombyx BmBLOS2* gene.

* Corresponding author. Tel.: +420 387775283; fax: +420 385310354.
E-mail address: zurovec@entu.cas.cz (M. Zurovec).

[1] These authors contributed equally to this work.

0965-1748/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ibmb.2012.10.011

Another type of chimeric nuclease containing a more versatile DNA binding domain with a simpler DNA binding code was described recently (Christian et al., 2010; Bogdanove et al., 2010; Miller et al., 2011). This DNA-binding domain originates from bacterial transcription factors of the Transcription Activator-Like (TAL) Effector family (Boch et al., 2009; Moscou and Bogdanove, 2009). These are highly conserved proteins secreted by a number of plant pathogenic bacteria of the genus *Xanthomonas* into plant cells as a part of the virulence process. The proteins bind to promoter sequences in the host plant and activate specific genes that cause disease or trigger a host resistance response. TAL effectors consist of a C-terminus bearing nuclear localization signal together with a transcription activation domain, a central repetitive domain responsible for DNA binding and the N-terminus encoding sequences needed for secretion (Szurek et al., 2002; Marois et al., 2002). The central sequence contains a varying number of modular repeated units of 34 (sometimes also 33 or 35) amino acid residues called TAL effector repeat domains. These domains are largely conserved except for amino acids 12 and 13, so called "repeated variable di-residues" (RVDs), in each repeat and the last TAL effector repeat domain which is shorter (a "half-repeat"). The RVDs are crucial for recognition of a single DNA base and display the following preferences: HD binds specifically to C, NG to T, NI to A and NN to G (Boch et al., 2009; Moscou and Bogdanove, 2009). The TAL effector repeat domains can be used to assemble 13–28 repeat units in any predetermined order (Boch and Bonas, 2010). In addition, the activity of a TAL effector requires a T nucleotide which precedes the nucleotide sequence specified by the RVDs (Boch et al., 2009; Moscou and Bogdanove, 2009).

TAL effector nucleases (TALENs) are analogous to ZFNs, in that the DNA-binding domain of the TAL effector is fused to the catalytic domain of Fok I nuclease. Similar to ZFNs, TALENs function in pairs via the dimerization of the Fok I nuclease domain and bind two half sites on the DNA separated by a spacer. The spacers in TALEN targets are longer than for ZFNs (between 12 and 30 bp) because of the longer amino acid sequence between the DNA-binding domain and FokI nuclease subunit (Christian et al., 2010). Moreover, there are very loose DNA sequence constraints for TALENs binding; therefore, TAL binding sites are much more frequent than ZFN targets and occur at least once in every 35 bp (Cermak et al., 2011). In addition, the individual RVD modules are associated with individual nucleotides independently allowing a simple, fast and inexpensive modular assembly strategy for TALEN construction. TALEN constructs are also commercially available from several vendors.

In our earlier study we mutagenized endogenous *B. mori* epidermal color loci *BmBLOS2* and *Bmwh3* by direct microinjection of custom-designed ZFN mRNAs into embryos (Takasu et al., 2010). The frequency with which the designed ZFNs functioned as mutagenic agents was quite low since successful mutagenesis was observed only in one of the three selected targets. In this study we examined the utility of TALENs for targeted gene modification in the silkworm. We designed three TALENs against target sites in the *BmBLOS2* genomic locus previously targeted by ZFNs (Takasu et al., 2010), microinjected TALEN-encoding RNAs into *Bombyx* embryos and examined their efficiency in gene disruption. We also adapted a simple functional assay for the evaluation of TALEN cleavage activity in yeast cells.

2. Material and methods

2.1. Silkworm strains

A nondiapausing strain, *w1-pnd* (*white egg 1*), which is *wild-type* (*wt*) for the *BmBLOS2* gene was used in all of the experiments. The strain is maintained at the Transgenic Silkworm Research Unit (National Institute of Agrobiological Sciences, Tsukuba, Japan). The larvae were reared on an artificial diet (Nihon Nosanko, Yokohama, Japan) at 25 °C.

2.2. DNA target site search

Targets for TALENs were searched for using an older version of a web program TALEN targeter (Doyle et al., 2012) (https://boglab.plp.iastate.edu/) in the 158 bp region of the *BmBLOS2* gene (Fujii et al., 2008) including the whole third exon with flanking 24 and 22 bp sequences in the second and third introns, respectively. Three different TALEN targets, BLT-1-3, were selected (Fig. 1) so that they would overlap with the previously used BL-1 ZFN target (Takasu et al., 2010). All three sites were found using the following search conditions: spacer length 15–30 bp, and 15–30 modular repeats. In addition, target BLT-3 was obtained after relaxing the condition for target base composition. The distances of BLT-1, BLT-2 and BLT-3 putative cleavage sites were 20 bp upstream, 19 bp downstream and 3 bp upstream from the BL-1 ZFN cleavage site, respectively. The target sites of selected TALEN pairs contained 19–22 bp spacer sequences (Fig. 1).

2.3. Golden Gate assembly of TALEN sequences

First we constructed a *pENT-TAL* vector from *pTAL3*, a backbone Golden Gate assembly plasmid, and a *pENT-NLS-G-Fok I* Gateway entry vector (Beumer et al., 2008) for the LR Clonase reaction (Invitrogen, Carlsbad, CA, USA). The PCR fragment (2660 bp) was amplified from the plasmid using primers TAL3F (5′-AAAGTCGACGGTTCAACAATGGCTTCCTC-3′) and TAL3R (5′-AAGACTAGTGGGGAAATTCGAGCTCCT-3′), digested by restriction enzymes Sal I and Spe I and inserted into the Sal I-Xba I site of the Gateway entry vector *pENT-NLS-G-Fok I* (Beumer et al., 2008).

The TALEN sequences were designed *in silico* by simple combination of repeat segments containing the appropriate RVDs. The segment arrays were prepared by employing Golden Gate assembly without the use of PCR (Cermak et al., 2011) using the Golden Gate TALEN Kit purchased from Addgene (Cambridge, MA, USA). The assemblies of monomeric modules were cloned between the Esp3I sites of the *pENT-TAL* plasmid (Fig. S1) (Cermak et al., 2011). The correct TALEN assembly was verified by sequencing. The DNA binding domains of TALEN pairs of the first TAL-target (BLT-1) contained 28 and 24 modular repeated units, the TALEN pair specific for the second TAL-target (BLT-2) contained 22 and 24 repeats, and the TALEN pair for the third target (BLT-3) contained 21 and 20 repeats (Fig. 1).

Completed TALEN coding sequences were then transferred into the Gateway destination plasmid *pCS2* containing the SP6 promoter for *in vitro* transcription, or to yeast expression vectors *pDEST-YcPlac33* and *pDEST-YcPlac111* described below using a Gateway LR Clonase reaction (Invitrogen, Carlsbad, CA, USA). (Fig. S1).

2.4. In vitro mRNA synthesis

TALEN constructs in the *pCS2 e*xpression vector were purified with a HiSpeed plasmid midi kit (Qiagene, Dusseldorf, Germany), linearized by digesting with Kpn I, treated with proteinase K (Nakarai, Japan), and subjected to *in vitro* transcription using an mMESSAGE mMACHINE kit with an SP6 promoter (Ambion, Carlsbad, CA, USA) as described previously (Takasu et al., 2010). The mRNA was precipitated with LiCl, washed with 70% ethanol three times, and dried in a vacuum centrifuge.

2.5. Silkworm embryo microinjection

The mRNAs encoding chimeric nucleases were dissolved in 0.5 mM phosphate buffer (pH 7.0) at a concentration of 0.2 μg/μl for each mRNA. A total of 3–5 nl of RNA solution was injected into

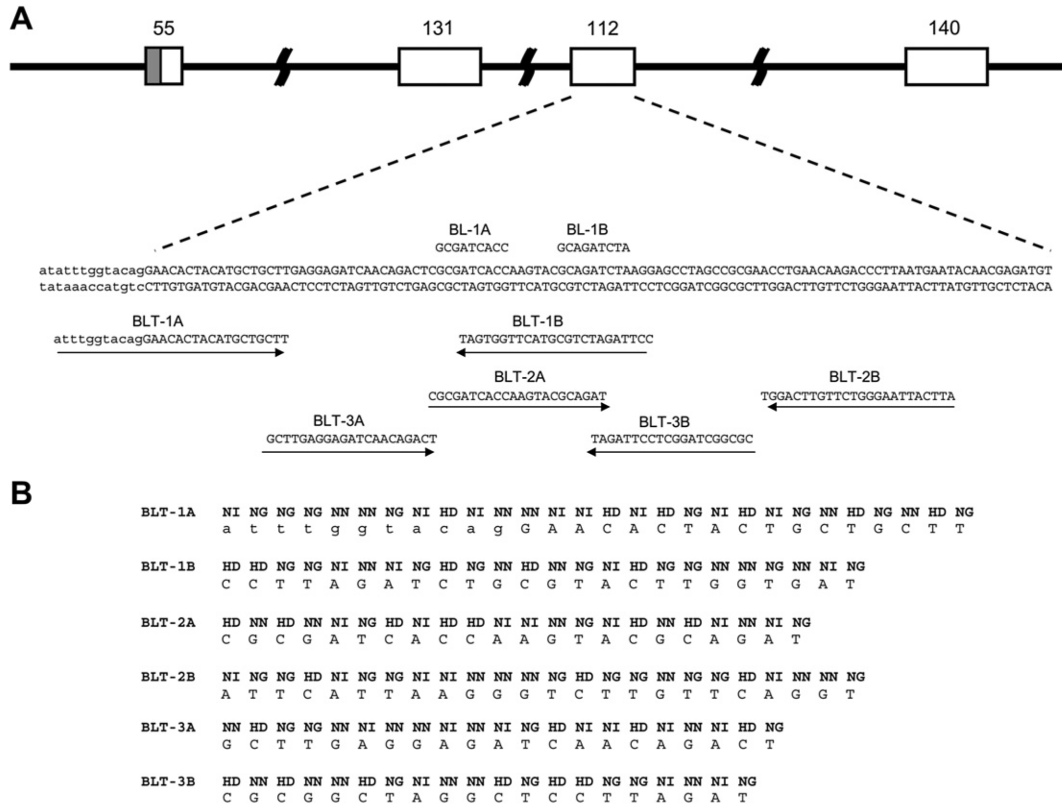


Fig. 1. Structure of the *BmBLOS2* gene, TALEN and BL-1 ZFN target positions and TALEN RVDs. Open boxes and numbers above represent ORF and length in bp. Filled box represents the 5′ UTR. The sequence below the line indicates the third exon (capital letters) with the adjacent partial sequence of the second intron (small letters). ZFN target sites are indicated above the sequence of exon 3; TALEN targets are shown below (A). RVDs (bold letters) and predicted target bases (lower rows of normal letters) of selected TALENs (B).

syncytial preblastoderm stage silkworm embryos through the chorion 4–8 h after oviposition as described previously (Tamura et al., 2000; Takasu et al., 2010). The slit in the chorion was sealed with instant glue (Aron Alpha, Konishi Co, Osaka) and the embryos were incubated at 25 °C in a humidified atmosphere.

2.6. The crossing scheme and screening strategy

The phenotype of *BmBLOS2* mutations involves defects of urate granule deposition in the epidermis causing its transparency ("*oily* skin"). Somatic mosaicism visible as patches of darker epidermis was screened at the fifth larval instar in G_0 chimeras (Fig. 2A, B). Germline mutations were detected during the first larval instar of G_1 larvae (Fig. 2C). The location of the *BmBLOS2* gene on the *Z* chromosome allowed screening for G_1 hemizygous females. Mutagenized G_0 male moths were crossed to mutagenized or *wt* females (Fig. S2), and the G_1 larvae were checked for the *oily* phenotype. Mutations were confirmed by PCR and sequencing.

2.7. Sequence analysis

Genomic DNA of G_1 mutants was extracted using a Blood & Tissue Genomic DNA Extraction Miniprep System (Viogene, Sijhih, Taiwan) according to the manufacturer's instructions. About 30 mg of tissue was used for each DNA sample. DNA fragments containing the targeted site of interest were amplified by PCR using primers from flanking regions F1 (5′-CTTCCAATTTGAGGGCAATG-3′) and R1 (5′-AATTTCACCACCTCATTCAACT-3′) as used previously with BL-1 ZFNs. The PCR products were gel purified and sequenced using the same primers.

2.8. Construction of yeast Gateway destination vectors

The construction started from yeast shuttle vectors, *YCplac33* (*Ura3, ARS1/CEN4, AmpR*) and *YCplac111* (*Leu2, ARS1/CEN4, AmpR*) which were converted into Gateway destination vectors by using efficient yeast *in vivo* gap repair cloning (Fig. S3). The shuttle vectors were digested by Sma I in the polylinker region and cotransformed into the BY4742 yeast cells together with a Pfu polymerase amplified "*Dest*" cassette (*M-PGal1–T7-attR1-CmR-ccdB-attR2-V5-6XHis-CYC1pA-CS*) from a *pYES-DEST52* plasmid (Invitrogen) using the method developed by Gietz and Woods (Gietz and Woods, 2002). The amplification primers were as follows:

GalRexUp – CCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCCGCGCGTTGGCCGATTCA.

GalRexDwn – CGTTGTAAAACGACGGCCAGTGAATTCGAGCTCGGTACCCACGGATTAGAAGCCGCCGA.

The primers contained 40 bp overhangs (underlined region), which consequently created a terminal homology at both ends of the Sma I-linearised shuttle vectors. Yeast cells containing gap-repaired plasmids were selected on plates containing glucose and lacking uracil (for *YCplac33* plasmid) or leucine (for *YCplac111* plasmid). Total DNA was isolated from yeast transformants and used for the transformation of competent *E. coli* DH5α cells. The *pDEST-YCplac33* and *pDEST-YCplac111* plasmids were isolated from *E. coli* and verified by sequencing.

2.9. Yeast single-strand annealing assays

We used a modified yeast-based strategy for testing ZFNs originally developed by Doyon et al. (2008). The yeast tester strain

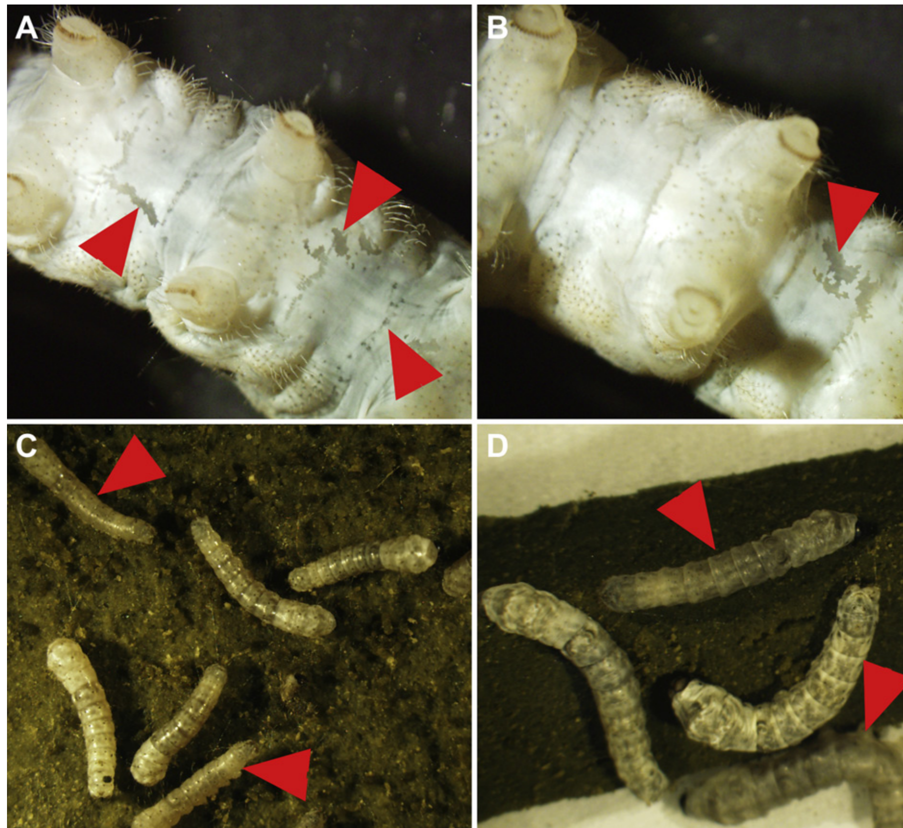


Fig. 2. Somatic and germline mutations in silkworm larvae. Ventral sides of fifth instar G_0 larvae showing mosaic areas of *BmBLOS2* mutant tissue (A, B). First instar (C) and third instar (D) of G_1 larvae with mutant individuals indicated by arrow heads.

BY4741 (*his3Δ1; leu2Δ0; met15Δ0; ura3Δ0*) was obtained from Dr. Z. Palkova (Charles University, Prague). The reporter plasmid was obtained from Dr. F. Urnov (Doyon et al., 2008). It contained a divided yeast *MEL1* gene bearing a 452 bp direct duplication separated by a 1.48 kb insert sequence containing a polylinker (Fig. S3). The synthetic DNA fragment containing tandem copies of the candidate target sequences was synthesized by GenScript Corp. (Piscataway, New Jersey). The fragments were subcloned using restriction sites Bam HI and Asc I into the reporter construct containing *HO-poly-KanMX4-HO* (Voth et al., 2001) and the completed construct was stably integrated into the *HO* locus of yeast chromosome *IV*. The yeast selections for neo/kan resistance gene were made on YPD agar plates with G418 (Sigma–Aldrich, St Louis, USA).

Such a tester yeast strain carrying an integrated reporter construct is used for transformation with pairs of TALEN-expressing plasmids *pDEST-YCplac33* and *pDEST-YCplac111.* Upon digestion by a TALEN the functional *MEL1* gene is restored by single-strand annealing (ssa) of the duplicated regions (Fig. S3) and the resulting α-galactosidase activity is measured by spectrophotometry at OD_{405} as described previously (Doyon et al., 2008).

3. Results

3.1. Target site selection and assembly of TALEN open reading frames

In order to evaluate the utility of TALENs for mutagenesis in *Bombyx*, we chose to target an endogenous locus, *BmBLOS2*, which has previously been successfully mutated with ZFNs. The *BmBLOS2* mutant shows an "*oily*" phenotype of translucent larval integument and the gene product is involved in urate granule accumulation in larval epidermis, rendering *wt* opaque (Fujii et al., 2008). As described in Material and methods, the on-line sequence analysis software TAL Effector Nucleotide Targeter (Doyle et al., 2012) was used to search for TALEN targets. The software predicted 17 targetable positions within the 112 bp sequence of the third exon of *BmBLOS2*. To avoid differences in the activity of custom targeting nucleases caused by DNA site-specific constraints (e.g. chromatin states) we selected TALEN target sites very close to or overlapping with the previously used ZFN target, *BL-1,* located in the third exon (Takasu et al., 2010). The sequence polymorphism of the third exon in the *Bombyx pnd* strain is relatively low and mutations in this region, including an 18 bp in-frame deletion, were previously shown to cause a null phenotype (Takasu et al., 2010).

TALEN coding regions were constructed by simple combination of RVD modules using Golden Gate assembly as described in Material and methods. The completed TALEN coding sequences were subcloned into the Gateway destination vectors for *in vitro* transcription or yeast expression (Fig. S1).

3.2. Yeast assays of TALEN activity

In order to identify the active TALENs and evaluate their potential toxicity we adapted a yeast (*Saccharomyces cerevisiae*) ssa assay which uses reporter activity of the secreted form of α-galactosidase (encoded by the yeast *MEL1* gene) as an indicator of nuclease function (Doyon et al., 2008). The yeast tester strain carrying the reporter construct inserted into the genome was transformed with appropriate pairs of TALEN expressing constructs. The OD_{405} corresponding to the α-galactosidase activity was assessed as described in Material and methods. The results showed that all three newly designed TALENs had detectable reporter

activity in yeast, suggesting that all three nucleases were functional enzymes (Fig. S4). All three TALENs were therefore taken for further *in vivo* experiments with *Bombyx* embryos.

3.3. In vivo experiments with silkworm embryos

We prepared RNAs encoding corresponding TALEN pairs using *in vitro* transcription and injected them into silkworm eggs at the syncytial preblastoderm stage as described in Material and methods. Approximately 480 preblastoderm stage embryos were injected with each pair of TALENs. The resulting G_0 larvae were inspected for patches of mosaic epidermis corresponding to *BmBLOS2* mutations in some somatic cells. As shown in Table 2, somatic mosaic G_0 larvae occurred with all three TALENs in relatively low frequency (less than 16%). The TALENs displayed the following rank order of activity in the induction of somatic mutations: BLT-1 > BLT-2 > BLT-3 (see Table 2). We also noticed that the patches of translucent epidermis were significantly smaller than those observed earlier with BL-1 ZFN mutagenesis (Takasu et al., 2010) (Fig. 2A, B).

The germline mutants were detected by "*oily*" phenotypes as described in Material and methods (Fig. 2C, D). Larvae from different broods were cultured separately. The number of observed mutants suggested that TALENs have the same order of activity in the germ line as in somatic mosaics: BLT-1 > BLT-2 > BLT-3.

As shown in Table 1, the survival rate of the eggs microinjected with individual TALEN pairs differed. The survival rate of embryos microinjected with the BLT-1 TALEN construct was 36.7% and only 14% (69 individuals) became fertile adults. In all we obtained 27 *oily* mutants.

The silkworm eggs microinjected with the BLT-2 TALEN construct had the lowest hatchability (20.9%) and only 36 of the 478 G_0 individuals became fertile adults. These produced 16 sibling groups of which one produced 18 *oily* mutants.

The embryos microinjected with the BLT-3 TALEN had the highest survival rate and their hatchability reached 48.6%. A total of 98 G_0 individuals out of 479 microinjected eggs became fertile adults. From these we obtained 52 broods and found 4 *oily* mutants in one of them.

3.4. Sequencing of representative mutations

To confirm the newly induced mutants, we amplified the DNA fragments flanking the target region using PCR and analyzed the sequences. Fig. 3 shows an alignment of DNA target sequences from 31 representative *BmBLOS2* mutants which were obtained in experiments with microinjections of RNA encoding BLT-1 and BLT-3 TALENs. We were unable to amplify the DNA from BLT-2 mutants suggesting that a larger deletion occurred.

As shown in Fig. 3A sequencing analysis of 27 BLT-1 TALEN mutants from 3 sibling groups revealed at least 4 types of mutations. The sequenced junctions contained small 2–15 bp deletions or insertions. Sibling group number 7 included two types of mutations, 4 and 15 bp deletions, confirming that two independent mutations could occur in germline cells of a single individual. All mutants from sibling group 10 contained a 15 bp deletion, whereas sibling group 12 contained a deletion-insertion mutation.

Four mutants of a single sibling group produced by BLT-3 TALEN were analyzed; they all contained the same 2 bp insertion (Fig. 3A).

Table 1
Survival of microinjected *B. mori* embryos.

Experiment	Embryos used	% hatched	% fertile adults	All broods	All G_1 progeny
BLT-1	480	37	14	28	3895
BLT-2	478	21	8	16	3831
BLT-3	479	49	20	49	8804
BL-1[a]	480	51	36	55	16,350

[a] Takasu et al. (2010).

Table 2
The efficiency of chimeric nuclease mutagenesis with direct embryo injection in *Bombyx mori*.

Experiment	% somatic mosaics	Yielders	Germline mutants	% germline mutants
BLT-1	15	3	27	0.69
BLT-2	10	1	18	0.47
BLT-3	6	1	4	0.05
BL-1[a]	72	5–9	46	0.28

[a] Takasu et al. (2010).

4. Discussion

BmBLOS2 is one of the two epidermal marker genes which we previously targeted by ZFNs (Takasu et al., 2010). In our previous study it was difficult to find a suitable targetable sequence for ZFN in the *BmBLOS2* locus; consequently, we had to choose ZFN targets containing two non GNN- triplets. These generally display a low success rate, and one out of the three previously examined ZFNs yielded germ line mutations. In the present study, we chose three potential TALEN target sites within the third exon of the *BmBLOS2* gene which overlapped with the BL-1 ZFN target. All three sites were successfully mutagenized by the three TALENs we designed, confirming a high success rate for this type of chimeric enzyme. Recent reports consistently describe the success rate of effective TALEN pairs tested on a larger scale to be greater than 70% (Cade et al., 2012).

The use of TALENs for NHEJ mutagenesis of the *BmBLOS2* gene solved the problems with the shortage of targetable sequences we experienced previously for ZFNs (Takasu et al., 2010). TALEN candidate cleavage sites occur much more often than those for ZFNs. The approximate TALEN target frequency is at least one targetable site per every 35 bp (Cermak et al., 2011) and it may be higher in coding regions. The older version of TARGETER prediction software using the TALEN platform of Cermak et al. (2011) identified at least 17 targetable positions within the 112 bp third exon sequence of *BmBLOS2*. The new version of this software (Doyle et al., 2012) predicted even more (28) targetable positions. Such flexibility in target selection will likely make it possible to find targets in almost any region of a gene including the translational start site, or to design targets that contain a restriction site within a spacer area to simplify mutation detection.

According to some reports by other researchers, TALENs cause a wider diversity of DNA mutations than ZFNs (Moore et al., 2012). In our experiments sequence analysis of TALEN mutated sites showed that most mutants contained small 2–15 bp deletions or insertions/deletions. A class of 4 bp deletions was found in 3 out of 31 mutants analyzed. This does not seem to result from the simple blunt ending of cut overhangs since the sites of the deletion were slightly shifted from the predicted double stranded break. Instead, this deletion is defined by a 2 bp microhomology (AC). Similarly, both 15 bp deletions found in broods 7 and 10 also seem to contain microhomologies (GATCA and CA, respectively). The mutation from brood 12 showed a 2 bp insertion so that the junction sequence CGATC is homologous to the template nearby, which is a typical example of synthesis-dependent microhomology-mediated end joining (SD-MMEJ) (Yu and McVey, 2010) (Fig. 3A). The mutation from brood 16 was a 2 bp insertion at the cleavage site. Such short inserts are the most frequently observed group of mutations in ZFN mutagenesis at the BL-1 target (Takasu et al., 2010).

```
A
(WT)  TGCTTGAGGAGATCAACAGACTCGCGATCACCAAGTACGCAGATCTAAGGAGCCTAGCCGCGAACCTGAACAA
      BLT-1 TALEN
(B7)  TGCTTGAGGAGATCA---------------CCAAGTACGCAGATCTAAGGAGCCTAGCCGCGAACCTGAACAA     (6)
(B7)  TGCTTGAGGAGATCAAC----TCGCGATCACCAAGTACGCAGATCTAAGGAGCCTAGCCGCGAACCTGAACAA     (3)
(B10) TGCTTGAGGAGATCAACA---------------AGTACGCAGATCTAAGGAGCCTAGCCGCGAACCTGAACAA     (10)
(B12) TGCTTGAGGAGATCAAC-GA-TCGCGATCACCAAGTACGCAGATCTAAGGAGCCTAGCCGCGAACCTGAACAA     (8)

      BLT-3 TALEN
(B16) TGCTTGAGGAGATCAACAGACTCGCGATCACCAAGTGTACGCAGATCTAAGGAGCCTAGCCGCGAACCTGAACAA   (4)
B
(WT)  TGCTTGAGGAGATCAACAGACTCGCGATCACCAAGTACGCAGATCTAAGGAGCCTAGCCGCGAACCTGAACAA

      PUBLISHED BL-1 ZFN MUTATIONS
(1)   TGCTTGAGGAGATCAACAGACTCGCGATCACCA--------GATCTAAGGAGCCTAGCCGCGAACCTGAACAA     (20)
(2)   TGCTTGAGGAGATCAACAGACTCGCGATCACCAAGTAACGCAGATCTAAGGAGCCTAGCCGCGAACCTGAACAA   (2+5)*
(3)   TGCTTGAGGAGATCAACAGACTCGCGATC------G-CGCAGATCTAAGGAGCCTAGCCGCGAACCTGAACAA    (8)
(4)   ATTACTATCTTGAAATTTCGAAATCCAGTATTTTAGTTGCAGATCTAAGGAGCCTAGCCGCGAACCTGAACAA    (1)
(5)   TGCTTGAGGAGATCAACAGACT------------------AGATCTAAGGAGCCTAGCCGCGAACCTGAACAA     (1)
(6)   TGCTTGAGGAGATCAACAGACTCGCGATCACCAAAGTACGCAGATCTAAGGAGCCTAGCCGCGAACCTGAACAA   (3)

      NEWLY OBSERVED BL-1 ZFN MUTATIONS
(7)   TGCTTGAGGAGATCAACAGACTCGCGATCACCAAGAGTACGCAGATCTAGGGAGCCTATCCGCGAACCTGAACAA  (2)
(8)   TGCTTGAGGAGATCAACAGACTCGCGATCACCA-GTACGCAGATCTAAGGAGCCTAGCCGCGAACCTGAACAA     (1)
(9)   TGCTTGAGGAGATCAACAGACTCGCGATCACCAAGTAAGTACGCAGATCTAAGGAGCCTATCCGCGAACCTGAACAA (5+29)*
(10)  TGCTTGAGGAGATCAACAGACTCGCGATCACCAAGTAACGCAGATCTAAGGAGCCTAGCCGCGAACCTGAACAA   (1+3+1)*
(11)  TGCTTGAGGAGATCAACAGACTCGCGATCAC----T---CAGATCTAAGGAGCCTAGCCGCGAACCTGAACAA    (1)
(12)  TGCTTGAGGAGATCAACAGACTCGCGATCACCA--------GATCTAAGGAGCCTAGCCGCGAACCTGAACAA     (39)
(13)  TGCTTGAGGAGATCAACAGACTCGCGATCACCAA--------ATCTAAGGAGCCTAGCCGCGAACCTGAACAA     (1)
(14)  TGCTTGAGGAGATCAACAGACTCGCGATCACCA-GATCGCAGATCTAAGGAGCCTAGCCGCGAACCTGAACAA    (46)
(15)  deletion 82 bp + 3 bp insertion ---GATCAGATCTAAGGAGCCTAGCCGCGAACCTGAACAA      (16)
(16)  TGCGTGAGGAGATCAACATACTCGCGATCACCA--TACGCAGATCTAAGGAGCCTATCCGCGAAGCTGAACAA    (7)
(17)  TGCTTGAGGAGATCAACAGACTCGCGATCAC---G--CGCAGATCTAAGGAGCCTAGCCGCGAACCTGAACAA    (7)
(18)  TGCTTGAGGAGATCAACAGACTCGCGATCACCGCATACCGCAGATCTAAGGAGCCTAGCCGCGAACCTGAACAA   (1)
(19)  TGCTTGAGGAGATCAACAGACTCGCGATCAC-------GCAGATCTAAGGAGCCTAGCCGCGAACCTGAACAA    (3)
(20)  deletion 232 bp  ------------------------------------------------            (6)
(21)  deletion 71 bp + insertion 814 bp
```

Fig. 3. Sequence comparison of *BmBLOS2* alleles produced by TALENs and ZFNs. TALEN and ZFN target sites are highlighted in yellow. Microhomologies are highlighted in grey and inserted bases are shown in boxes. Dashes indicate deleted bases. WT, wild-type DNA sequence; B7, brood number 7; B10, brood number 10; B12, brood number 12; B16, brood number 16. The number of individuals sharing a certain repair junction is indicated in parenthesis on the right. "*" indicates mutants that appeared independently in different broods.

Sequence analysis showed that 209 published and novel BL-1 ZFN mutant DNA junctions could be assigned to 21 classes (Fig. 3B). The most frequently occurring mutations were 1–4 bp insertions homologous to the neighboring nucleotide(s) at the cleavage site including mutation types 2, 6, 7, 9 and 10, which arose as a result of nine independent events. Simple deletions based on MMEJ (mutation types 1, 12, 19 and 20) were also observed frequently, together with combinations of deletions and insertions based on SD-MMEJ (mutation types 3, 11, 14 and 17) and other types of mutations that have microhomologies at the junctions (mutation types 15 and 18). On the other hand, we could find only 4 independent mutations that occurred, most probably, through blunt end joining without microhomology (mutation classes 5, 8, 13 and 16).

Our results of TALEN and ZFN mutagenesis suggest that the major repair mechanisms of double strand breaks in the silkworm are MMEJ, SD-MMEJ, and blunt end joining with 1–4 bp insertion or short deletions. This appears to be similar to *Drosophila melanogaster* (Bozas et al., 2009). For wider applications of TALEN mutagenesis in genome editing it is important to establish the conditions of homologous recombination (HR). HR can be used to modify the targeted sites by introducing a dominant marker gene, which would allow easy screening and genotyping. Beumer et al. (2008) succeeded to increase the efficiency of HR with co-injected donor plasmid by using a DNA ligase IV deficient fruit fly strain. These results suggest that knockout of DNA ligase IV may also be effective to promote HR in *Bombyx*.

The survival rate of microinjected *Bombyx* embryos and larvae was somewhat lower for TALENs than for ZFNs in previous experiments. The larval mortality was slightly higher and only about 50% of hatched larvae became fertile adults (Table 1). The BLT-2 microinjected individuals displayed the lowest survival rate and only 8% of them became fertile adults. The egg hatchability for BLT-2 was lower than the standard range of 30–70% (Takasu et al., 2010). Although we cannot exclude the possibility that slight toxicity of BLT-2 contributes to the lower egg hatchability, the lower survival rate of BLT-2-injected larvae was observed only in last iterations. It suggests that this variation occurred due to the lower purity of the last BLT-2 RNA preps or similar experimental variable.

We assume that experimental variations also significantly influence the efficiency of mutagenesis. The efficiency of mutation induction by the tested TALENs in the silkworm germ line described in this study was 0.69% and 0.47% for BLT-1 and BLT-2, respectively. The previously reported efficiency of BL-1 ZFN mutagenesis was 0.28% (Takasu et al., 2010). In addition, further rounds of BL-1 ZFN microinjections mostly resulted in slightly higher efficiency with mutation rates ranging from 0 to 2% (Takasu, unpublished). We expect that larger samples of microinjected embryos will bring greater statistical precision and increased efficiency.

Because the *BmBLOS2* gene resides on the sex chromosome (Z) and the gene knockouts led to recessive mutations, we were able to detect only 1/2 of the mutations induced in the G_0 embryos by phenotypic screening (see Fig. S2). This means that the rate of NHEJ-induced mutations (shown in Table 2) was greater than 1% for BLT-1.

The TALEN technology is quite new and further research and sequence optimization are in progress. The TALEN constructs we used were based on architecture described by Cermak et al. (2011) and contained full-length N and C-termini of TAL proteins. Several recent articles have shown that different platforms for building TALENs with the N- and C-terminal truncation of TAL sequences may have a big impact on cleavage efficiency (Cade et al., 2012; Miller et al., 2011, 2012; Mussolino et al., 2011). Direct comparison of the activity of truncated and full-length TALENs is complicated

because the truncation of the C-terminal sequences also requires the use of a shorter spacer sequence between the binding half sites (Sun et al., 2012). Further optimization of TALEN sequences may increase their function and subsequent success rate, as well as the number of targetable sites in given DNA sequence.

5. Conclusion

Our results show that TALENs can be successfully used for gene mutagenesis in *Bombyx* and confirm the utility of TALENs especially at loci lacking canonical ZFN target sites. Despite the fact that ZFNs are of a smaller size and their amino acid sequences are less repetitive in nature, which may give them an advantage for certain applications, we did not observe any disadvantages of TALENs concerning these parameters. We found that the design and construction of TALENs was very simple and fast and the success rate was better than with ZFNs since all three TALENs examined yielded germline mutants. We also showed that TALEN activity can be preexamined in yeast, further increasing the potential efficacy of the more difficult and time-consuming steps required for germline introduction and screening for mutations in the host. The efficiency of mutagenesis we observed in testing only three TALENs should be sufficient for detection of mutant alleles based solely on molecular methods. Our results represent an important step towards the routine use of gene targeting in the silkworm.

Acknowledgment

We thank Fyodor D. Urnov and Michal Cap for the help with the establishment of yeast assays, Dr. Zdena Palkova (Charles University, Prague) for providing yeast tester strain *BY4741*, Mr. Kaoru Nakamura and Mr. Kohji Hashimoto for embryo microinjection and silkworm rearing, and Dr. Qiang Zhang for technical assistance. We also thank Dr. Marian R. Goldsmith for critical reading of the manuscript. This research was supported by grant KJB501410801 from the Grant Agency of the Academy of Sciences of Czech Republic and KAKENHI Grant Number 23580083 from Japan Society for the Promotion of Science.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.ibmb.2012.10.011.

References

Alwin, S., Gere, M.B., Guhl, E., Effertz, K., Barbas 3rd, C.F., Segal, D.J., Weitzman, M.D., Cathomen, T., 2005. Custom zinc-finger nucleases for use in human cells. Mol. Ther. 12, 610–617.
Beumer, K.J., Trautman, J.K., Bozas, A., Liu, J.L., Rutter, J., Gall, J.G., Carroll, D., 2008. Efficient gene targeting in drosophila by direct embryo injection with zinc-finger nucleases. Proc. Natl. Acad. Sci. U S A 105, 19821–19826.
Boch, J., Bonas, U., 2010. Xanthomonas AvrBs3 family-type III effectors: discovery and function. Annu. Rev. Phytopathol. 48, 419–436.
Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., Bonas, U., 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326, 1509–1512.
Bogdanove, A.J., Schornack, S., Lahaye, T., 2010. TAL effectors: finding plant genes for disease and defense. Curr. Opin. Plant Biol. 13, 394–401.
Bozas, A., Beumer, K.J., Trautman, J.K., Carroll, D., 2009. Genetic analysis of zinc-finger nuclease-induced gene targeting in Drosophila. Genetics 182, 641–651.
Cade, L., Reyon, D., Hwang, W.Y., Tsai, S.Q., Patel, S., Khayter, C., Joung, J.K., Sander, J.D., Peterson, R.T., Yeh, J.R., 2012. Highly efficient generation of heritable zebrafish gene mutations using homo- and heterodimeric TALENs. Nucleic Acids Res.
Carroll, D., Morton, J.J., Beumer, K.J., Segal, D.J., 2006. Design, construction and *in vitro* testing of zinc finger nucleases. Nat. Protoc. 1, 1329–1341.
Cermak, T., Doyle, E.L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J.A., Somia, N.V., Bogdanove, A.J., Voytas, D.F., 2011. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res. 39, e82.
Christian, M., Cermak, T., Doyle, E.L., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A.J., Voytas, D.F., 2010. Targeting DNA double-strand breaks with TAL effector nucleases. Genetics 186, 757–761.
Doyle, E.L., Booher, N.J., Standage, D.S., Voytas, D.F., Brendel, V.P., Vandyk, J.K., Bogdanove, A.J., 2012. TAL effector-nucleotide targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. Nucleic Acids Res. 40, W117–W122.
Doyon, Y., McCammon, J.M., Miller, J.C., Faraji, F., Ngo, C., Katibah, G.E., Amora, R., Hocking, T.D., Zhang, L., Rebar, E.J., Gregory, P.D., Urnov, F.D., Amacher, S.L., 2008. Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. Nat. Biotechnol. 26, 702–708.
Fujii, T., Abe, H., Katsuma, S., Mita, K., Shimada, T., 2008. Mapping of sex-linked genes onto the genome sequence using various aberrations of the Z chromosome in Bombyx mori. Insect Biochem. Mol. Biol. 38, 1072–1079.
Gietz, R.D., Woods, R.A., 2002. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. Methods Enzymol. 350, 87–96.
Gupta, A., Christensen, R.G., Rayla, A.L., Lakshmanan, A., Stormo, G.D., Wolfe, S.A., 2012. An optimized two-finger archive for ZFN-mediated gene targeting. Nat. Methods 9, 588–590.
Isalan, M., Choo, Y., Klug, A., 1997. Synergy between adjacent zinc fingers in sequence-specific DNA recognition. Proc. Natl. Acad. Sci. U S A 94, 5617–5621.
Kim, Y.G., Cha, J., Chandrasegaran, S., 1996. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc. Natl. Acad. Sci. U S A 93, 1156–1160.
Maeder, M.L., Thibodeau-Beganny, S., Osiak, A., Wright, D.A., Anthony, R.M., Eichtinger, M., Jiang, T., Foley, J.E., Winfrey, R.J., Townsend, J.A., Unger-Wallace, E., Sander, J.D., Muller-Lerch, F., Fu, F., Pearlberg, J., Gobel, C., Dassie, J.P., Pruett-Miller, S.M., Porteus, M.H., Sgroi, D.C., Iafrate, A.J., Dobbs, D., McCray Jr., P.B., Cathomen, T., Voytas, D.F., Joung, J.K., 2008. Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. Mol. Cell 31, 294–301.
Marois, E., Van den Ackerveken, G., Bonas, U., 2002. The Xanthomonas type III effector protein AvrBs3 modulates plant gene expression and induces cell hypertrophy in the susceptible host. Mol. Plant Microbe. Interact. 15, 637–646.
Miller, J.C., Tan, S., Qiao, G., Barlow, K.A., Wang, J., Xia, D.F., Meng, X., Paschon, D.E., Leung, E., Hinkley, S.J., Dulay, G.P., Hua, K.L., Ankoudinova, I., Cost, G.J., Urnov, F.D., Zhang, H.S., Holmes, M.C., Zhang, L., Gregory, P.D., Rebar, E.J., 2011. A TALE nuclease architecture for efficient genome editing. Nat. Biotechnol. 29, 143–148.
Moore, F.E., Reyon, D., Sander, J.D., Martinez, S.A., Blackburn, J.S., Khayter, C., Ramirez, C.L., Joung, J.K., Langenau, D.M., 2012. Improved somatic mutagenesis in zebrafish using transcription activator-like effector nucleases (TALENs). PLoS One 7, e37877.
Moscou, M.J., Bogdanove, A.J., 2009. A simple cipher governs DNA recognition by TAL effectors. Science 326, 1501.
Mussolino, C., Morbitzer, R., Lutge, F., Dannemann, N., Lahaye, T., Cathomen, T., 2011. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. Nucleic Acids Res. 39, 9283–9293.
Pearson, H., 2008. Protein engineering: the fate of fingers. Nature 455, 160–164.
Sander, J.D., Dahlborg, E.J., Goodwin, M.J., Cade, L., Zhang, F., Cifuentes, D., Curtin, S.J., Blackburn, J.S., Thibodeau-Beganny, S., Qi, Y., Pierick, C.J., Hoffman, E., Maeder, M.L., Khayter, C., Reyon, D., Dobbs, D., Langenau, D.M., Stupar, R.M., Giraldez, A.J., Voytas, D.F., Peterson, R.T., Yeh, J.R., Joung, J.K., 2011. Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). Nat. Methods 8, 67–69.
Sun, N., Liang, J., Abil, Z., Zhao, H., 2012. Optimized TAL effector nucleases (TALENs) for use in treatment of sickle cell disease. Mol. Biosyst. 8, 1255–1263.
Szurek, B., Rossier, O., Hause, G., Bonas, U., 2002. Type III-dependent translocation of the Xanthomonas AvrBs3 protein into the plant cell. Mol. Microbiol. 46, 13–23.
Takasu, Y., Kobayashi, I., Beumer, K., Uchino, K., Sezutsu, H., Sajwan, S., Carroll, D., Tamura, T., Zurovec, M., 2010. Targeted mutagenesis in the silkworm Bombyx mori using zinc finger nuclease mRNA injection. Insect Biochem. Mol. Biol. 40, 759–765.
Tamura, T., Thibert, C., Royer, C., Kanda, T., Abraham, E., Kamba, M., Komoto, N., Thomas, J.L., Mauchamp, B., Chavancy, G., Shirk, P., Fraser, M., Prudhomme, J.C., Couble, P., 2000. Germline transformation of the silkworm Bombyx mori L. using a piggyBac transposon-derived vector. Nat. Biotechnol. 18, 81–84.
Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S., Gregory, P.D., 2010. Genome editing with engineered zinc finger nucleases. Nat. Rev. Genet. 11, 636–646.
Voth, W.P., Richards, J.D., Shaw, J.M., Stillman, D.J., 2001. Yeast vectors for integration at the HO locus. Nucleic Acids Res. 29, E59.
Yu, A.M., McVey, M., 2010. Synthesis-dependent microhomology-mediated end joining accounts for multiple types of repair junctions. Nucleic Acids Res. 38, 5706–5717.

Supplementary informations

Figure S1. Schematic illustration of a procedure for TALEN RNA preparation. The procedure involved the construction of *pENT-TAL* plasmid (upper left corner), Golden Gate assembly of TALEN open reading frames (shown on green background) and transfer of TALEN constructs into Gateway destination expression vectors for mRNA synthesis and yeast assays (upper right corner).

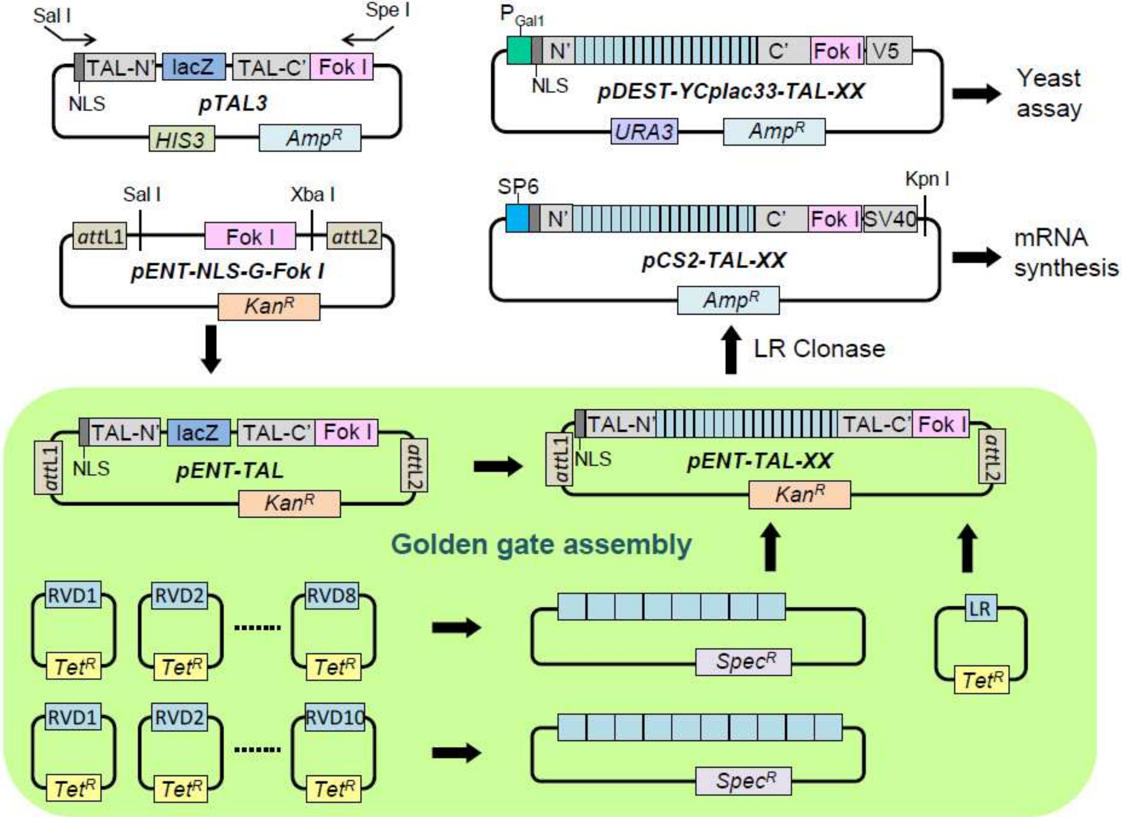


Figure S2. Mutagenesis of *BmBLOS2* gene located on Z chromosome (females are hemizygous for Z). TALEN RNA was directly microinjected to silkworm embryos, the resulting G_0 larvae were inspected for patches of mosaic epidermis. G_0 adult males were crossed with mutagenized G_0 or *wt* females as shown in the scheme below. *BmBLOS2* germline mutations are detectable among the female G_1 larvae by *oily* phenotype.
* denotes mutagenized allele.

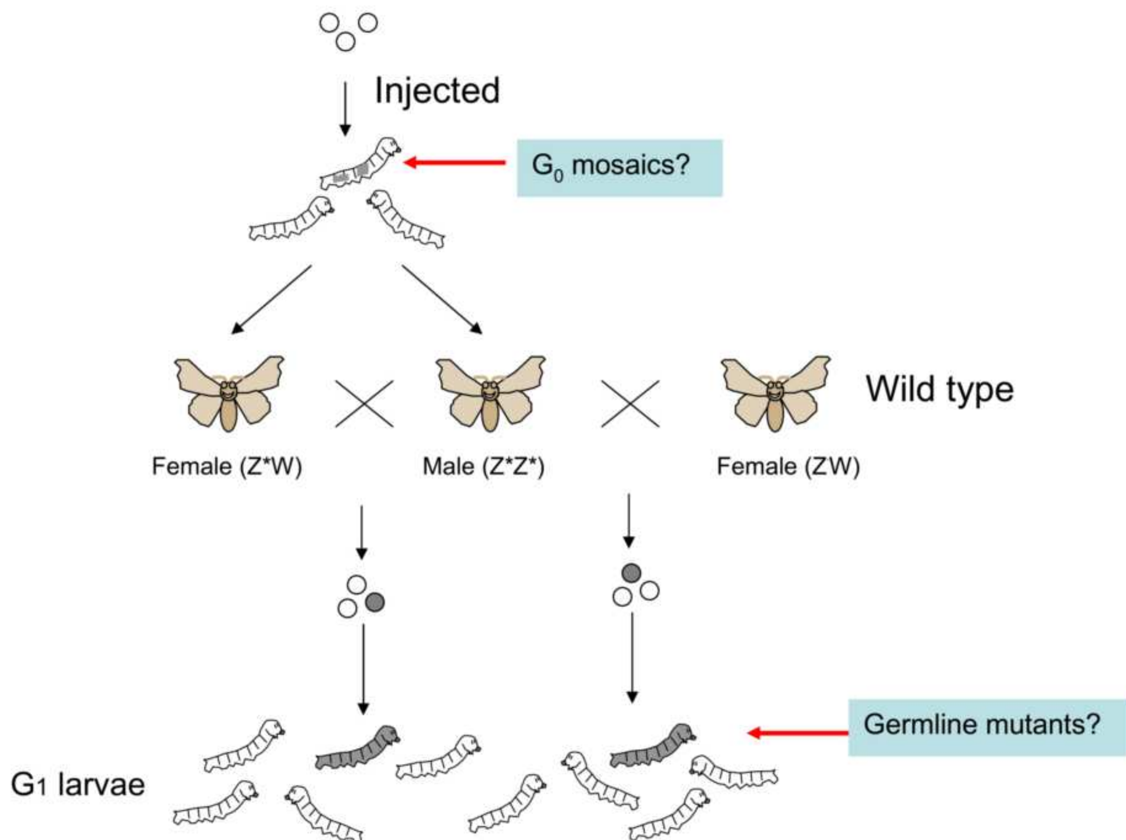


Figure S3. Construction of yeast Gateway destination vectors using *in vivo* yeast gap repair cloning. The „*Dest*“ cassette (*M-PGal1 –T7-attR1-CmR-ccdB-attR2-V5-6XHis-CYC1pA-CS*) from *pYES-DEST52* (Invitrogen) was amplified by *Pfu* polymerase using anchor primers (contained 3' sequence complementary to the Dest cassette, shown as a green line, and 5' overhangs introducing 40 bp terminal homology with the termini of linearised *YCplac* vectors shown as a black line). The PCR product was cotransformed with *SmaI*-digested shuttle vectors *YCplac33* (*Ura3, ARS1/CEN4, AmpR*) and *YCplac111* (*Leu2, ARS1/CEN4, AmpR*) into yeast strain *BY4742* (*his3Δ1; leu2Δ0; lys2Δ0; ura3Δ0*). The two completed destination vectors were isolated from yeasts grown on synthetic dropout media plates.

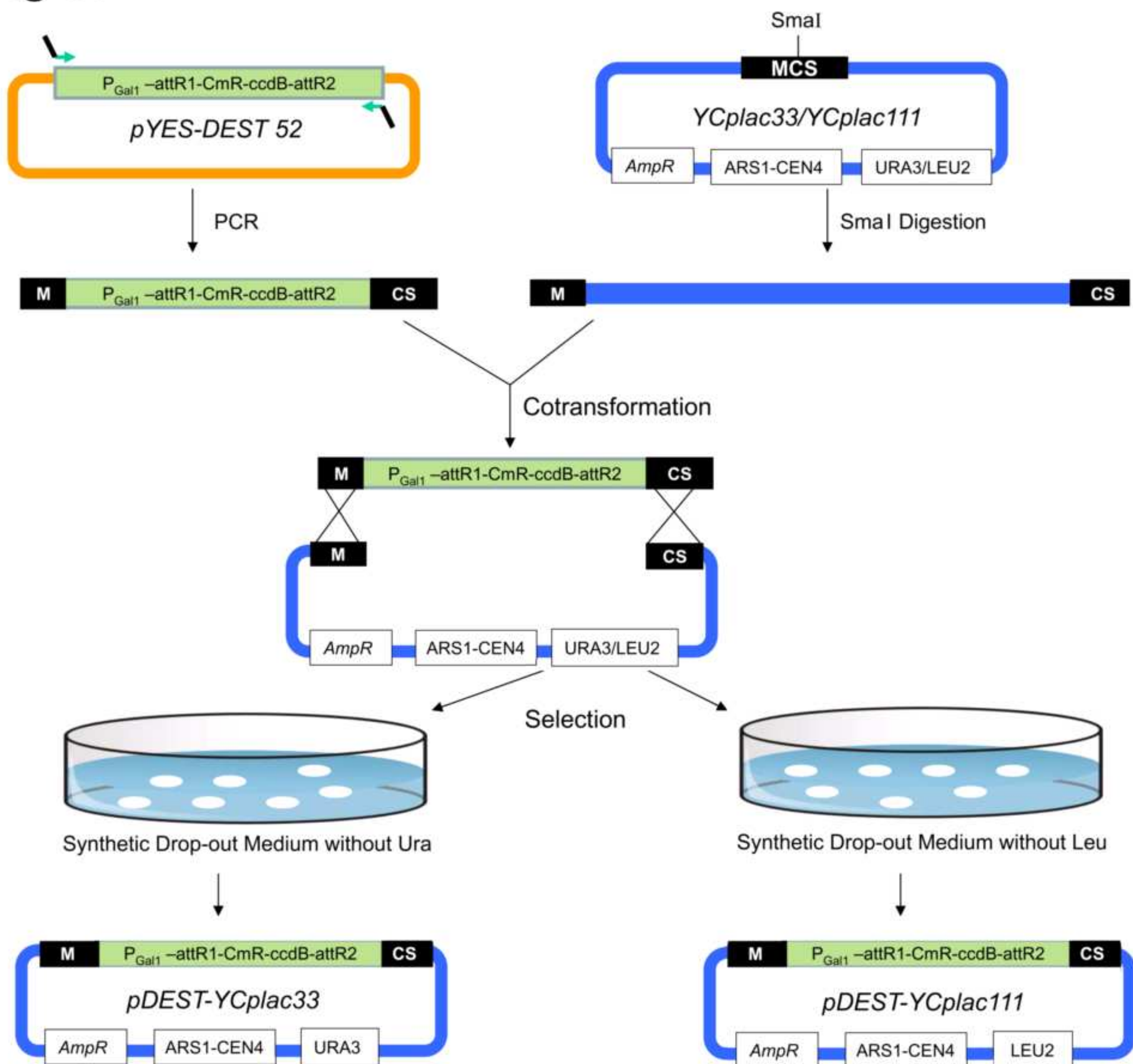
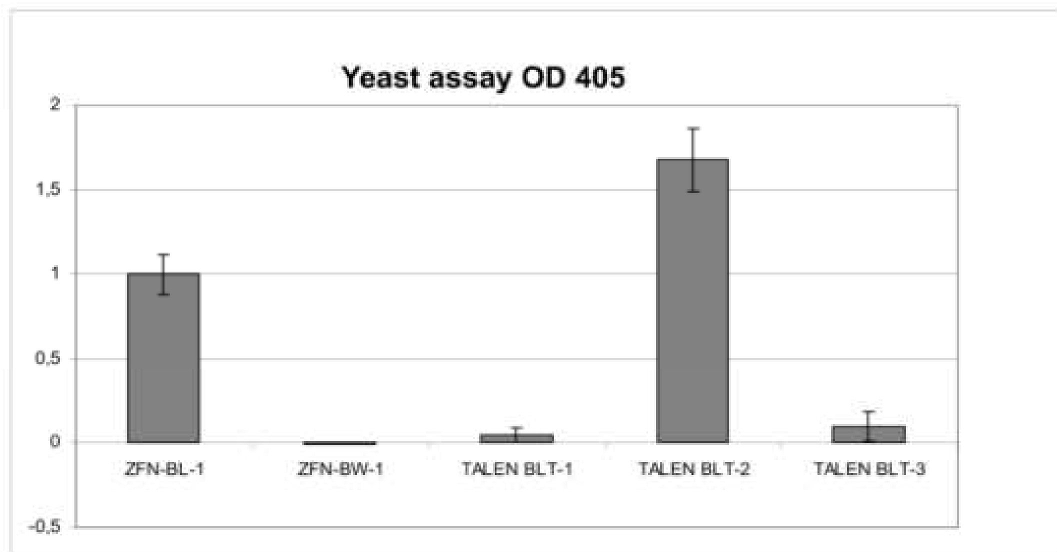


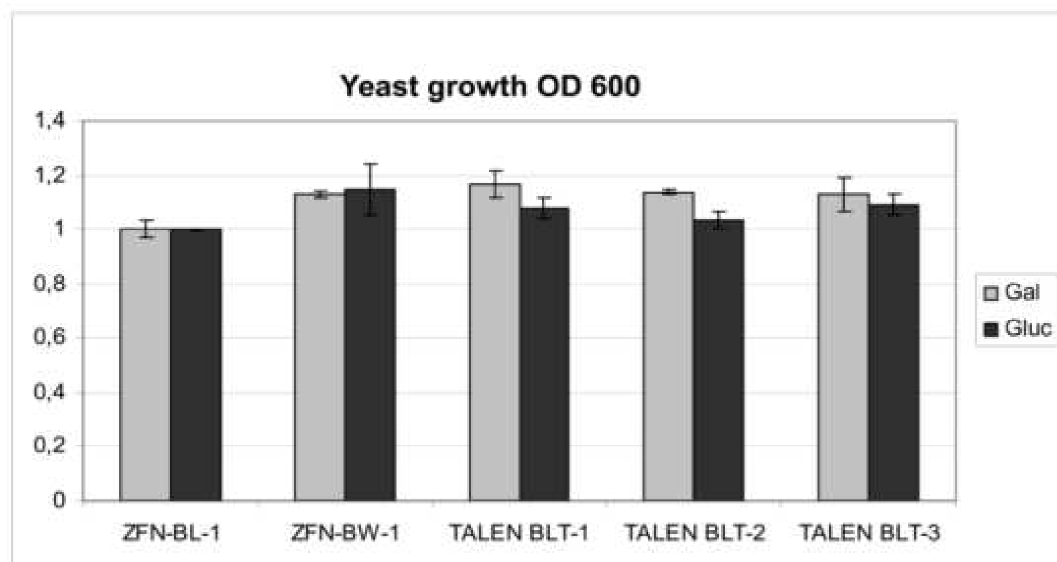
Figure S4. Yeast assays of DNA cleavage activity and toxicity. **(A)** Comparison of TALENS BLT-1-3 with two ZFN pairs used previously (Takasu et al. 2010). The BL-1 ZFN was used as a positive control while BW-1 ZFN served as a negative control. The reporter α-galactosidase activity was assayed using the chromogenic substrate X-alpha-Gal at OD_{405}. All three TALENs possessed detectable nuclease activity. **(B)** The growth of yeast cells transfected with chimeric enzyme constructs was measured at OD_{600} nm. Results suggest that the constructs did not cause significant toxicity in yeasts.

Fig S4

A



B



Part II

Targeted mutagenesis in the silkworm *Bombyx mori* using zinc finger nuclease mRNA injection

Insect Biochemistry and Molecular Biology 40 (2010) 759–765

Contents lists available at ScienceDirect

Insect Biochemistry and Molecular Biology

journal homepage: www.elsevier.com/locate/ibmb

Targeted mutagenesis in the silkworm *Bombyx mori* using zinc finger nuclease mRNA injection

Yoko Takasu [a,b], Isao Kobayashi [b], Kelly Beumer [c], Keiro Uchino [b], Hideki Sezutsu [b], Suresh Sajwan [a], Dana Carroll [c], Toshiki Tamura [b], Michal Zurovec [a,*]

[a] *Biology Centre, Academy of Sciences and the Faculty of Natural Sciences, University of South Bohemia, Branisovska 31, 370 05 Ceske Budejovice, Czech Republic*
[b] *National Institute of Agrobiological Sciences, 1-2 Owashi, Tsukuba, Ibaraki 305-8634, Japan*
[c] *Department of Biochemistry, University of Utah School of Medicine, Salt Lake City, UT 84112, USA*

A R T I C L E I N F O

Article history:
Received 7 May 2010
Received in revised form
27 July 2010
Accepted 28 July 2010

Keywords:
Gene targeting
Nonhomologous end joining
Lepidoptera
BmBLOS2
Bm1 element

A B S T R A C T

Targeted mutagenesis is one of the key methods for functional gene analysis. A simplified variant of gene targeting uses direct microinjection of custom-designed Zinc Finger Nuclease (ZFN) mRNAs into *Drosophila* embryos. To evaluate the applicability of this method to gene targeting in another insect, we mutagenized the *Bombyx mori* epidermal color marker gene *BmBLOS2*, which controls the formation of uric acid granules in the larval epidermis. Our results revealed that ZFN mRNA injection is effective to induce somatic, as well as germline, mutations in a targeted gene by non-homologous end joining (NHEJ). The ZFN-induced NHEJ mutations lack end-filling and blunt ligation products, and include mainly 7 bp or longer deletions, as well as single nucleotide insertions. These observations suggest that the *B. mori* double-strand break repair system relies on microhomologies rather than on a canonical ligase IV-dependent mechanism. The frequency of germline mutants in G_1 was sufficient to be used for gene targeting relying on a screen based solely on molecular methods.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

The recent sequencing of several insect genomes for species of agricultural or medical interest, such as *Bombyx mori*, *Apis mellifera*, *Anopheles gambiae*, *Aedes aegypti*, and *Tribolium castaneum*, offers a wealth of genes potentially involved in key biological processes. Further progress in both basic and applied research is dependent on molecular and genetic methods, including transgenics and gene knockouts. Such technologies, initially developed for *Drosophila melanogaster*, are being gradually introduced and modified for other insects. Yet many methods remain to be adapted to non-drosophilid species.

The silkworm, *B. mori*, has been reared to produce silk for thousands of years. *Bombyx* genetics is well established with a great number of mutant strains and marker genes. Key molecular genetic methods have been successfully established for silkworm, including stable transgenesis of the germline (Tamura et al., 2000) targeted gene expression using the *GAL4/UAS* system (Imamura et al., 2003), and enhancer trap screening (Uchino et al., 2008). Several reports have involved RNAi gene silencing, including dsRNA injection into the silkworm embryo (Quan et al., 2002; Tomita and Kikuchi, 2009) and RNA hairpins expressed by recombinant Sindbis virus (Uhlirova et al., 2003) or in a *piggyBac* transgene (Isobe et al., 2004). Nevertheless, the knock-down of gene expression by RNAi has a serious limitation in the silkworm since gene silencing is incomplete. The absence of a general gene-targeting system to allow systematic reverse genetic studies has been a significant limitation in silkworm research. The first attempt to establish targeted mutagenesis in the silkworm was reported by Yamao et al. (1999), who induced a mutation into the *fib-L* gene by homologous recombination, using *Autographa californica nucleopolyhedrovirus* (AcNPV). However, the efficiency of this method was too low to be used as a standard technique.

Two methods of gene targeting were established for *Drosophila* by Rong and Golic (2000) and Bibikova et al. (2002). The first is based on a pair of site-specific DNA modifying enzymes from yeast, a recombinase and endonuclease that release a linear DNA fragment containing a modified sequence of the target gene in primordial germ cells. This method allows gene alteration by homologous recombination, but requires generation of transgenic fly strains expressing the yeast enzymes and a number of crossing steps to put a multitude of transgenes together. Thus the use of this technique in other insects has not been established yet. The other method is based on custom designed zinc finger nucleases (ZFNs), which are

* Corresponding author. Tel.: +420 387775283; fax: +420 385310354.
E-mail address: zurovec@entu.cas.cz (M. Zurovec).

0965-1748/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ibmb.2010.07.012

chimeric enzymes consisting of a zinc finger DNA recognition domain and the nonspecific nuclease domain of the Fok I restriction enzyme (Kim et al., 1996). This method allows a simple change in the target sequence by non-homologous end joining (NHEJ). Alternatively, specific alterations of the target region can be achieved by providing a donor plasmid with a mutated target sequence, which may recruit the homologous repair machinery. Whereas the original protocol describing ZFN mutagenesis also required transgenic flies and extensive crossing (Bibikova et al., 2002), a simplified variant of this method using direct embryo injection with mRNAs encoding ZFN was described recently in *Drosophila* (Beumer et al., 2008), zebrafish (Doyon et al., 2008) and rats (Geurts et al., 2009). This adapted method does not require ectopic expression of enzymes and avoids laborious genetic manipulations. The microinjected RNA is translated into a functional ZFN which induces double stranded breaks in a specific region of the genome. The free ends of the digested DNA then initiate a repair process, which can lead to mutations. As well, a plasmid with a mutated donor sequence can be coinjected to allow homologous recombination.

We attempted direct ZFN mRNA injection to *B. mori* embryos to evaluate the applicability of this procedure to gene targeting in the silkworm. The reported success rate of simple assembly of characterized triplet binding ZFN modules is relatively low (Ramirez et al., 2008), and targeted mutagenesis usually works only for one out of four targets (Kim et al., 2010). We chose three target sites residing in two *B. mori* epidermal color marker genes in order to offset this relatively low success rate. Both genes control the formation of uric acid granules in the larval skin, and mutations lead to the visible phenotype of translucent epidermis.

2. Materials and methods

2.1. Silkworm strains

A nondiapausing strain (*pnd*), which is *wt* for *BmBLOS2* and *Bmwh3,* was used in all of the experiments. The $w\text{-}3^{ol}$ strain is mutant for *Bmwh3* and was used as a tester. The *pnd* and $w\text{-}3^{ol}$ strains were from silkworm collections maintained at the National Institute of Agrobiological Sciences (Tsukuba, Japan) and the Graduate School of Agriculture, Kyushu University (Fukuoka Japan), respectively. The larvae were reared on an artificial diet (Nihon Nosanko, Yokohama, Japan) at 25 °C.

2.2. Target selection, DNA constructs

Two *B. mori* epidermal color marker genes, *BmBLOS2* and *Bmwh3*, were chosen for experiments and their coding regions were surveyed for the best available ZFN targets, preferably sequences close to $(NNC)_3N_6(GNN)_3$, using the "Zinc Finger Tools" program (Mandell and Barbas, 2006) available on the Carlos Barbas laboratory website (http://www.scripps.edu/mb/barbas/zfdesign/zfdesignhome.php). The candidate sites were assessed according to a table of DNA triplets (Table S1A,B and C) using the specificity evaluation reported by Carroll et al. (2006). Two sequences in the first and third exons of *BmBLOS2* and one sequence from the second exon of *Bmwh3* gene were selected and designated as targets BL-2, BL-1 and BW-1, respectively (Fig. 1 and S1, S4 and S7).

The coding regions of zinc-finger proteins that specifically bind the target gene sequence were designed and synthesized from corresponding oligonucleotides (Fig. S2-S9) combined by PCR using the method of Carroll et al. (2006). The procedure therefore included the assembly of Sp1C frameworks (consensus-based protein derived from the middle zinc finger of human transcription factor Sp1) with the ZF coding sequences using the 7-primer procedure and subcloning the resulting 294-bp DNA fragments into the pENTR-NLS-ZFN gateway vector (Desjarlais and Berg, 1992; Carroll et al., 2006). The vector also contained a nuclear localization signal (NLS) and a Fok I nuclease domain-encoding sequence (Carroll et al., 2008). For *in vitro* transcription, the resulting DNA fragments encoding the complete ZFN with NLS were transferred to the destination vector pCS2-DEST (Fig. S10) via the Clonase reaction (Invitrogen, Carlsbad, USA).

2.3. mRNA synthesis

Template plasmids for *in vitro* transcription were purified with a Qiagen HiSpeed plasmid midi kit (Qiagen, Germany), digested completely with the *Not I* restriction enzyme and subjected to *in vitro* transcription using an mMESSAGE mMACHINE kit with SP6 polymerase (Ambion, USA) according to the manufacturer's protocol. RNA was precipitated with LiCl, washed with 70% ethanol three times and air-dried.

2.4. Microinjection

Pairs of mRNA samples were dissolved at a concentration of 0.1 μg/μl for each RNA in 0.5 mM phosphate buffer (pH 7.0) containing 5 mM KCl. RNA concentration was measured by absorbance at 260 nm. RNA solutions of 3–5 nl were injected into silkworm eggs 4–8 h after oviposition, which corresponds to the syncytial preblastoderm stage. Injections were performed through the chorion as previously described by Tamura et al. (2000). The injection opening was sealed with instant glue (Aron Alpha, Konishi Co, Osaka) and the embryos were allowed to develop at 25 °C. Mutations of the target gene in epidermal cells were detected during the final larval instar as a mosaic of translucent and normal (opaque) skin of injected individuals.

2.5. The crossing scheme and screening strategy

Since the *BmBLOS2* gene is located on the *Z* chromosome, we took advantage of female hemizygosity (Fig. S11). In our screen, mutagenized G_0 males were crossed to wild-type females and G_1 females were checked for the *oily* phenotype (Fig. 2). Since the nonmosaic oily phenotype is well visible even in the first instar larvae, the cost of the screening was quite low.

The second gene used in this study, *Bmwh3,* is an autosomal mutation, which also causes the *oily* phenotype, for which a homozygous mutant tester strain $w\text{-}3^{ol}/w\text{-}3^{ol}$ was available. In the pilot experiment to investigate the appearance of somatic mutations, we scored the fifth instar larvae of the microinjected heterozygous $w\text{-}3^{ol}/+$ individuals (Fig. S12A). For the detection of germline mutants, we scored the first instars of the F_1 progeny from a cross between the microinjected G_0 *pnd* individuals and $w\text{-}3^{ol}/w\text{-}3^{ol}$ tester silkworms (Fig. S12B).

2.6. Sequence analysis

Genomic DNA of all mutant G_1 individuals was extracted using the Blood & Tissue Genomic DNA Extraction Miniprep System (Viogene, Sijhih, Taiwan) according to the manufacturer's instructions. About 30 mg of larval, pupal, or moth tissue were used per DNA extraction to obtain 3–15 μg of DNA. Fragments flanking the target region were amplified by PCR using primers 1F and 1R (Fig. 1). The PCR product was gel purified and sequenced using the same primers. Some DNA samples contained a larger deletion, and thus three sets of more distal primer pairs were used for PCR. All PCR reactions were carried out using Ex Taq DNA polymerase (Takara-bio, Kyoto, Japan). Sequencing reactions were performed with a BigDye terminator cycle sequence ready reaction kit version 3.1 and resolved on an ABI Prism 377 capillary sequencer (Applied Biosystems).

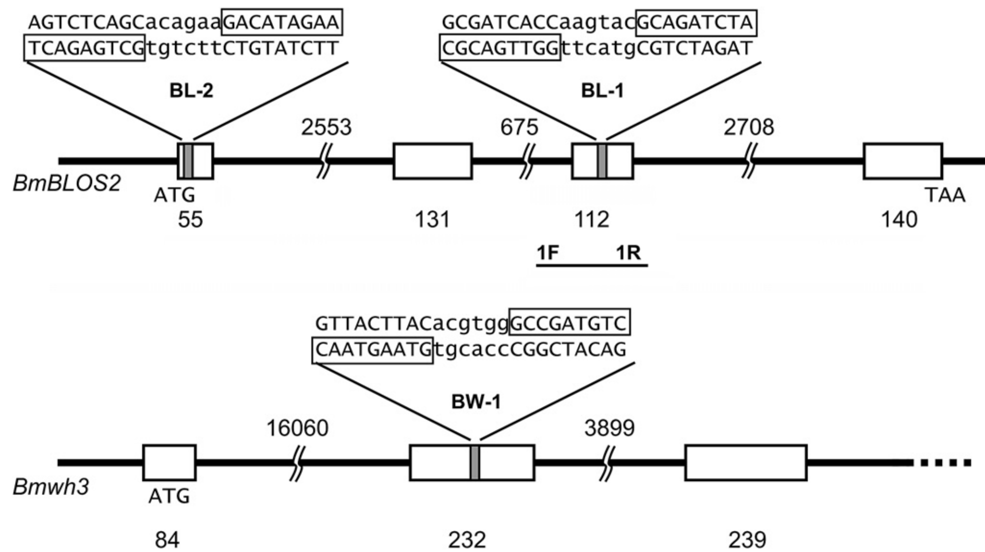


Fig. 1. Structure of the *BmBLOS2* and *Bmwh3* genes. Open boxes on the lines represent exons. The ZFN target sites are indicated as gray boxes with their sequences depicted above (boxed letters identify nine nucleotide motifs recognized by zinc fingers), two of them in the first and third exons of *BmBLOS2* gene (BL-2 and BL-1, respectively) and one in the second exon of the *Bmwh3* gene. The sizes of exons and introns (in bp) are indicated below and above the map. The pair of primers 1F-1R amplified a 683 bp fragment, including the BL-1 target site. More information on ZFN target sites is shown in Fig. S1, S4 and S7.

3. Results

3.1. The test system

We selected *B. mori* as a new insect model species for establishing ZFN genetic targeting because of its well-developed genetics and large number of characterized markers (http://www.shigen.nig.ac.jp/silkwormbase/index.jsp), as well as the availability of a microinjection system for silkworm embryos (Tamura et al., 2000).

Previous experiments with ZFN mRNA injections into *Drosophila* embryos revealed that the usage of the *yellow* (*y*) marker, influencing epidermal color, allowed the detection of somatic mosaics in the G_0 generation (Beumer et al., 2008). The somatic *y* mosaics were observed in adult flies and had no practical use in experiments. The occurrence of somatic mosaics might, however, be potentially informative about the efficiency of microinjection and function of the introduced ZFNs.

Among the best characterized genes involved in the determination of epidermal color in silkworms are the *BmBLOS2* and *Bmwh3* genes, for which null mutations are recessive. Mutant larvae display translucent skin, also known as an "*oily*" phenotype (Fujii et al., 2008), since the gene products are involved in the accumulation of urate granules in the epidermis (Tamura and Sakate, 1983). There are approximately 25 mutant loci in this biosynthetic pathway (http://www.shigen.nig.ac.jp/silkwormbase/ViewStrainGroup.do), although few genes have been identified thus far.

The *BmBLOS2* gene is located on the *Z* chromosome (Fujii et al., 2008), which allows the detection of germline mutants in hemizygous G_1 females. *BmBLOS2* is a relatively small gene consisting of four exons and encoding a protein of 145 amino acids. There were no high affinity $(NNC)_3N_6(GNN)_3$ ZF binding sites within the coding sequence; therefore, we had to use targets containing two non-GNN triplets (Fig. 1). The best available target (designated BL-1) was located in exon 3 (Fig. 1) and contained two out of six triplets of the CNN type. As shown in Table 1, only two triplets in the BL-1 target were rated as "+++" by the specificity evaluation algorithm developed by Carroll et al. (2006). The second target from the *BmBLOS2* (designated BL-2) is located in exon 1, which also exhibits two out of six triplets of the non-GNN type.

The second *Bombyx* marker used for target site selection, *Bmwh3*, is an ABC transporter gene homologous to the *Drosophila white* locus (Abraham et al., 2000) and involved in urate granule deposition. *Bmwh3* mutants also display translucent larval skin and white eggs. Since we did not find a bona fide $(NNC)_3N_6(GNN)_3$ ZF binding site within this gene either, we also selected a target site with two non-GNN triplets (Fig. 1). One of the triplets, AGT, had a low rating of a single "+" (Table 1 and S1C).

We designed open reading frame (ORF) sequences of all three ZFN pairs appropriate for targeting the previously described sequences. We then assembled the appropriate synthetic oligonucleotides corresponding to the sequences of DNA binding domains within the Sp1C ZFN framework (see Figs. S2, S3, S5, S6, S8, S9 and Materials and Methods). Finally, we synthesized the ZFN mRNAs *in vitro* and injected them into silkworm eggs at the syncitial preblastoderm stage as described in Materials and Methods (Fig. S10, S11, S12).

3.2. Mutagenesis of the BL-1 target site

The hatchability of eggs injected with the ZFN specific for the BL-1 target was 51%. Approximately 72% of the fifth instar G_0 larvae of this group showed a mosaic pattern of translucent and normal skin (Fig. 2A, B), thus representing somatic mutants. The patches of translucent epidermis appeared in the ventral side of the body. As expected, the proportion of mosaic epidermis in the larvae was higher in females than in males (70 *vs*. 56 percent, respectively), with a majority of females (80%) expressing *oily* skin extending over more than three segments. In addition, the *oily* skin areas were smaller in mosaic males. Approximately 58% of males and 20% of females displayed "small" *oily* areas, distributed over only one to three larval segments. The relatively high percentage of male mosaics showed that this ZFN was efficient enough to create somatic mutations in both chromosomal alleles.

We crossed the mutagenized BL-1G_0 males with *wt* females in order to test for germline mutants, which entailed a total of 71 crosses (Fig. S11). We obtained progeny from 55G_0 males yielding a total of 16350G_1 offspring. Among these, we detected 46 larvae, presumably females, which displayed translucent skin (Fig. 2E,F). These G_1 mutants came from at least five different G_0 males,

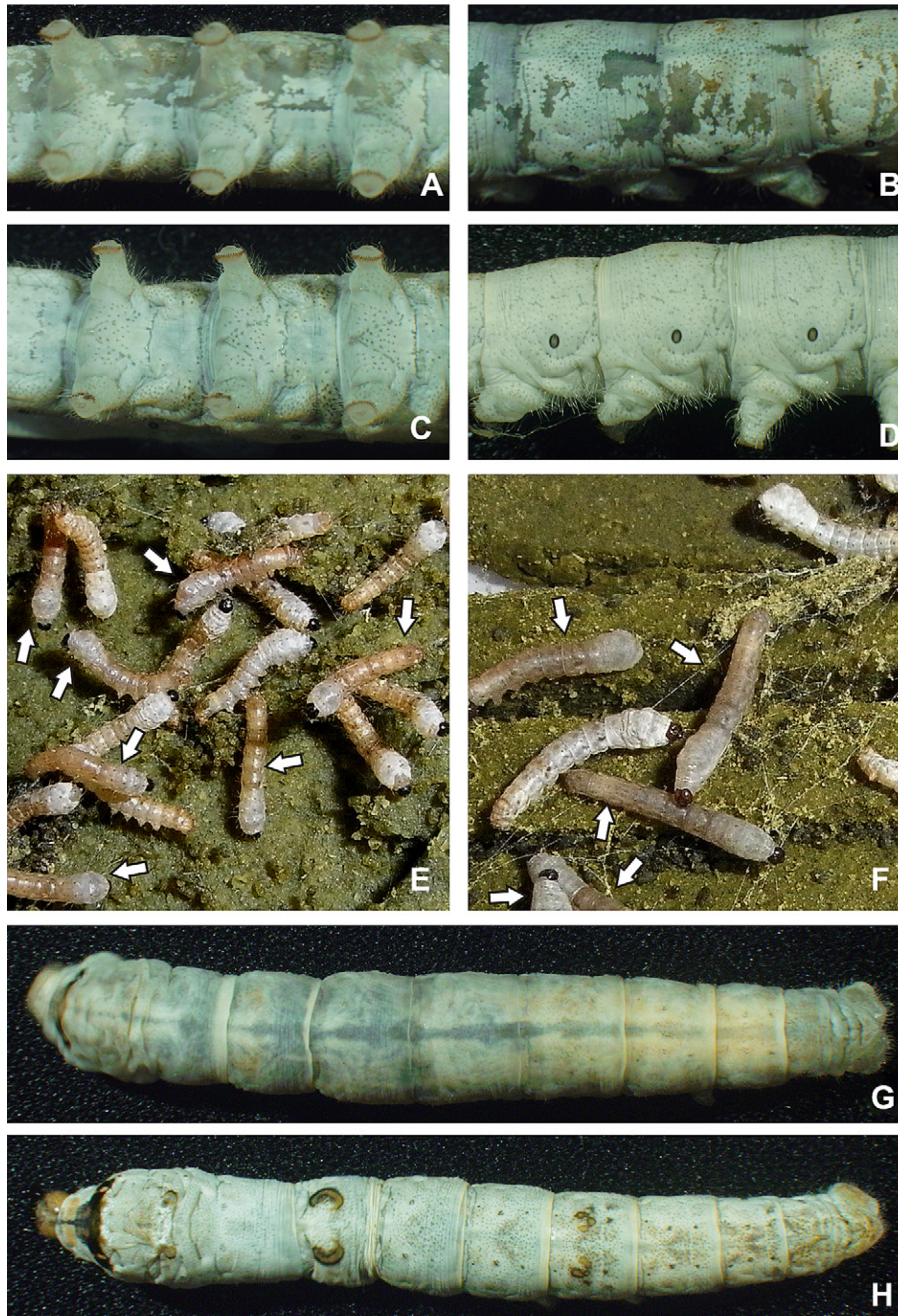


Fig. 2. Silkworms with somatic and germline mutations. Mutagenesis in epidermal cells of G_0 silkworms was detected by screening for *oily* mosaics in 5th instar larvae, which appeared predominantly on the ventral side (A), and rarely on the dorsal side (B). Ventral and dorsal sides of silkworms with normal epidermis are shown for comparison (C and D, respectively). Germline mutations were detected by the presence of a complete *oily* phenotype. G_1 mutants among 1st and 2nd instar larvae are indicated by arrows (E and F, respectively), and mutant and normal larvae in the final instar are shown in G and H, respectively.

meaning that more than 9% of fertile males yielded mutants. Eight of the G_1 mutant larvae died before metamorphosis, while the rest grew to the pupal stage, when they were confirmed to be females.

3.3. Mutagenesis of BL-2 and BW-1 target sites

The hatchability of the eggs injected with target BL-2 ZFNs (*BmBLOS2* gene) was very low. We injected 480 eggs, but only 18 of them hatched and we observed no somatic mosaics or germline G_1 individuals (Table 3). Similarly, the results for the BW-1 target site from the *Bmwh3* gene showed low hatchability. Among 144 injected eggs, 41 hatched and 27 reached the last larval instar. We observed six somatic mosaics among the G_0 individuals (22%), but no germline mutants in more than 4000G_1 larvae (see Fig. S13 for crossing schemes).

3.4. Verification of Mendelian segregation

Thirty-four of the G_1 female moths emerged from the BL-1 *oily* pupae and were crossed with male moths of the wild-type nondiapausing *pnd* strain; 25 of them laid fertile eggs. All of the G_2 males were heterozygotes showing the normal *wt* phenotype. Eleven G_2 males were crossed with females of the *pnd* strain and the G_3 larvae confirmed the expected 3:1 ratio of phenotype

Table 1
Target triplets and corresponding ZFN sequences. One, two or three "+"'s rate relative specificity of triplets for ZFN targeting according to the system of Carroll et al. (18). See Tables S1A,B and C for more information.

ZFN	DNA triplet			Recognition sequence/quality rating		
	F1	F2	F3	F1	F2	F3
BL-1R	CTA	GAT	GCA	QNSTLTE++	TSGNLVR+++	QSGDLTR++
BL-1L	CGC	GAT	GGT	HTGHLLE++	TSGNLVR+++	TSGHLVR++
BL-2R	GAA	ATA	GAC	QSGNLAR++	QKSSLTA+	DRSNLTR+
BL-2L	ACT	GAG	GCT	THLDLTR+++	RSDNLAR+++	QSSDLTR+++
BW-1R	GTA	AGT	AAC	QSSSLVR+++	HRTTLTN+	DSGNLRV++
BW-1L	GCC	GAT	GTC	DCRDLAR+++	TSGNLVR+++	DPGALVR++

segregation (50% of female progeny) supported by a χ^2 statistical test and only one G_2 male offspring line (line number 65) showed marginally significant difference from the expected ratio (Table S2).

3.5. Sequence analysis of representative BmBLOS2 alleles

We extracted genomic DNA from 46 *BmBLOS2* BL-1 mutants representing at least five independent sibling groups. We amplified the areas surrounding the mutated target points by PCR and determined the target region sequences. They showed a pattern of small deletions and insertions (Fig. 3) which are characteristic for NHEJ junctions (Beumer et al., 2006; Bibikova et al., 2002). We were able to distinguish several classes of mutations. The first class contained small 7 and 8 bp deletions. The second class contained an 18 bp deletion, which caused a loss of six amino acids. The third class had two types of insertions of a single adenosine at the target sites, causing frame shifts. A fourth class encompassed uncharacterized mutations that we were unable to amplify and most probably representing large deletions. The last class was represented by one large substitution, which occurred in brood 19, in which a 4.5 kb region of the *BmBLOS2* locus was replaced with a 596 bp fragment from chromosome 8, bearing significant homology with 131 bp *Bm1* element within the 5' upstream region of *BmBLOS2* gene (Fig. S13). Sibling groups 9, 19 and 69 included at least two types of mutations. This meant that NHEJ occurred in more than two germline cells in the G_0 embryo. The most frequent changes were small 7 and 8 bp deletions (63%) and small insertions (11%). The individuals carrying the 7 and 8 bp deletions originated in the same brood and may each represent progeny of a single germ line cell.

4. Discussion

Our work demonstrates that ZFN targeting can be successfully used to generate germline mutations in an insect species other than

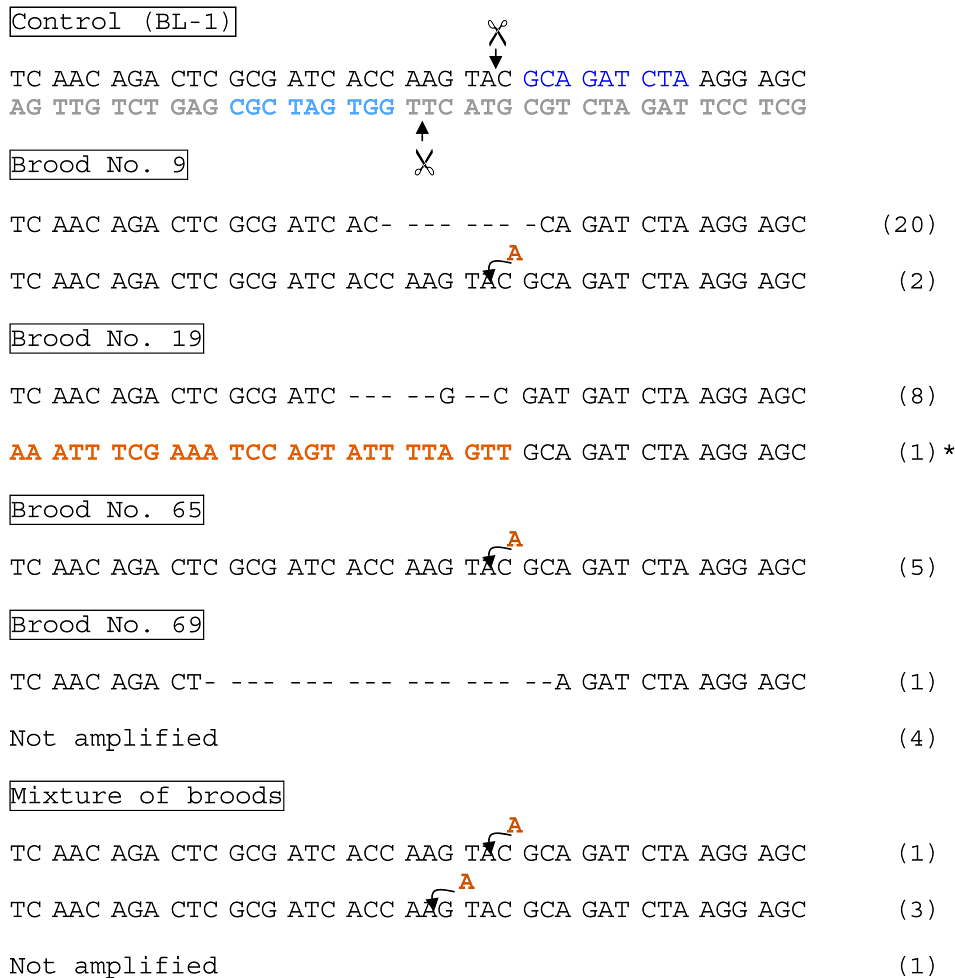


Fig. 3. Sequence analysis of novel *BmBLOS2* alleles. The sequence of the BL-1 site (both strands) is shown at the top for comparison. G_1 mutants were obtained from at least five broods, four of which (No. 9, 19, 65, and 69) were reared separately; and the remainder ("mixture of broods") came from a separate mix of several moths reared together, thus preventing the determination of the lineage of transformants. Each genotype was determined by the sequencing of PCR fragments as described in Materials and Methods. Dashes indicate deletions of residues and orange letters are inserted or substituted residues. The brood numbers with sequences of mutants are shown together with the number of individuals bearing a certain genotype (indicated in parenthesis on the right). "*" indicates that this insertion matches the sequence from chromosome 8 (see also Fig. S13).

Drosophila. The direct microinjection of mRNA into embryos allowed us to avoid the laborious construction of transgenic strains and genetic manipulations needed for the proper expression of DNA-modifying enzymes. The use of an epidermal marker gene allowed easy detection of somatic mutations and monitoring of the efficiency of microinjection. The amount of RNA microinjected into preblastoderm embryos, as well as the conditions used for microinjections, were similar to those used earlier in *Bombyx* by Uchino et al. (2007) with *Minos* mRNA.

We performed ZFN targeting of three target sites from two loci and were able to obtain germline mutants for only one target site. Our success rate was higher than the reported average for ZFNs designed by the modular assembly approach (Ramirez et al., 2008). Our target sites did not contain optimal target sequences composed solely of GNN triplets or triplets with the high "+++" rating (Beumer et al., 2006) and each of them contained two non-GNN triplets (Table 1). The results for particular targets differed significantly. The hatchability of eggs microinjected with the BL-1 target ZFNs mRNAs was 51%, which lies well within the standard range of 30–70% for injected silkworm eggs (Table 2). Also the number of fertile adults that emerged from the microinjected individuals was quite high, suggesting that the microinjection procedure worked well. In contrast, the injections of ZFN mRNAs for the BL-2 and BW-1 targets resulted in high embryo mortality, which may have been due to the lower specificity of the BL-2- and BW-1-ZFNs, which both contained triplets rated as only "+" (Table 1, S1B and S1C).

Although the hatchability of eggs was an important indicator of the possible toxicity of injected molecules, it still did not provide any information about the function of the microinjected product. The occurrence of somatic mosaics, however, indicated whether the engineered ZFN in question was or was not functioning (Table 3). Whereas we observed a high number (72%) of somatic mosaics for BL-1 target mutagenesis (*BmBLOS2*), we observed only 22% with the BW-1 target (*Bmwh3*) ZFNs and none in the BL-2 (*BmBLOS2*) targeting experiment. These results indicated that the choice of the epidermal markers for the *Bombyx* targeting experiment was important, since there was a correlation between the frequency of induced NHEJ somatic mutations with the number of germline mutants detected. The frequency of BL-1 specific somatic mosaics was comparable with the 46-80% mosaics described for the experiment with the *Drosophila yellow* gene, in which the flies were mutagenized by ZFNs using a heat-shock protocol (Bibikova et al., 2002).

We screened 16350 G_1 larvae for the *oily* phenotype and detected a total number of 46 BL-1 *BmBLOS2* target female mutants, or 0.28 percent of the total G1 progeny. The true number of mutations was probably twice as much since our phenotypic screen did not allow the detection of heterozygous male mutants. This number would still be more than an order of magnitude less than the percentage of mutants reported in a *Drosophila rosy* gene mutagenesis screen (Beumer et al., 2008); (Table 3), suggesting that the specificity of BL-1 ZFN might be lower despite its high efficiency in generating somatic mosaics. Alternatively, the lower efficiency of microinjected constructs into the silkworm germline compared to similar experiments in *Drosophila* may be a species-specific feature (T. Tamura, personal communication). This disparity may be due to differences in germ cell formation between the two species as well as the formation of the pole plasm, which is at the posterior end in *Drosophila* and at the ventral midline in silkworm (Nakao, 2009).

Table 2
Comparison of the survival of microinjected *Bombyx* embryos used for ZFN-NHEJ mutagenesis and microinjections with *Minos* vector RNA.

Experiment	Number of injected embryos	% Hatched	% Fertile adults	Reference
BL-1 *BmBLOS*	480	51	36	This study
BL-2 *BmBLOS*	480	4	N.D.	This study
BW-1 *Bmwh3*	144	28	18[a]	This study
Pmia3A	1888	40.3	26	(21)
Pmia3B1	894	35.3	23.9	(21)
Pmia3B2	678	44.5	27.1	(21)
Pmia3B3	2288	40.0	37.7	(21)

[a] These individuals were kept only until the last larval instar.

Table 3
The Efficiency of ZFN-NHEJ mutagenesis with direct embryo injection in *Bombyx (B.m.)* and *Drosophila (D.m.)*.

Experiment	Number of embryos used	Yielders	% Somatic mosaics	% Germline NHEJ mutants	Non-GNN triplets in the target	Reference
B.m. BL-1	480	5–9	72	0.28[a]	2	This study
B.m. BL-2	480	0	0	0	2	This study
B.m. BW-1	144	0	22	0	2	This study
D.m. pask	14	5	N.D.	6	1	(12)
D.m. rosy	99	41	N.D.	8.2	0	(12)
D.m. coil	45	5	N.D.	5–8	1	(12)

[a] Only females were detected by the phenotypic screen.

We confirmed the BL-1 target *BmBLOS2* mutants by molecular analysis of the targeted regions. Each brood contained mostly the same type of mutation, which suggests that numerous individuals may have originated from the same mutated germ line progenitor cell.

We did not observe 4 bp deletions or insertions, which would be expected from loss and fill in, respectively, of the 5′ overhang left by ZFN cleavage (Beumer et al., 2008; Smith et al., 2000). The nucleotide sequence pattern of mutations was therefore slightly different from the NHEJ results with wild-type flies. The silkworm mutations we observed were instead reminiscent of those obtained in flies deficient in *lig4* (DNA ligase IV) or in *Caenorhabditis elegans*, in which double-stranded repair involves microhomologies (Bozas et al., 2009; Morton et al., 2006). Perhaps the generation of the deletions observed in our *BmBLOS2* mutants is also driven by a microhomology-induced mechanism. The most common 8-bp deletion (Fig. 3) is defined by a 2-bp microhomology (CA). It would be interesting to see if this phenomenon is commonplace upon the generation of more mutants. Interestingly, the 7-bp deletion is probably a deletion-insertion, and it is not flanked by a microhomology (Fig. 3). Another unique deletion-insertion mutant recovered from brood 19 contained repair product accompanied by the insertion of extra, "captured" DNA at the breakpoint, which showed homology with *Bm1* element (Fig. S13). The 600 bp inserted fragment most probably originated from chromosome 8 (Fig. S13) and might occur by synthesis-dependent strand annealing (Kurkulos et al., 1994; Nassif et al., 1994).

Our goal is to knockout genes with unknown phenotypes. Thus, we must rely on the efficiency of the method employed and the detection of mutations based solely on molecular assays. The critical step is the design of new zinc finger combinations directed to chosen DNA sequences. The design of ZFNs is improving dramatically with a number of efficient zinc-finger combinations described (Fu et al., 2009; Maeder et al., 2008). Alternatively, several candidate target sites from the same locus can be chosen together with ZFNs designed by modular assembly, and the most efficient enzyme could be selected by a simple yeast-based assay (Doyon et al., 2008) or a CEL-I Nuclease Mismatch Assay (Miller et al., 2007). The ZFNs chosen in this fashion could then be introduced via mRNA injection to the embryos with more confidence. Non-invasive genotyping could be used for G_1 mutant screens, based on DNA obtained from a single drop (6–10 μl) of hemolymph (by pricking individual fifth instar larvae), followed by PCR amplification together with fragment length analysis and sequencing. If the percentage of G_1

germline mutants is about the same as described in this paper, we should be able to obtain several positive individuals per 1000 treated silkworms. Screening this number of larvae is well within the scale of a high-throughput silkworm laboratory.

Acknowledgments

We thank Mr. Kaoru Nakamura and Mr. Koji Hashimoto for their excellent help with *Bombyx* embryo microinjection and maintaining silkworm strains. We would also like to acknowledge Dr. Qiang Zhang for technical assistance and Dr. Natuo Kômoto (NIAS, Tsukuba) for sharing the unpublished information. We also acknowledge comments on the manuscript by Dr. Marian R. Goldsmith of the University of Rhode Island, Kingston, USA. The research was supported by grant P305/10/2406 from the Grant Agency of the Czech Republic, Grant IAA500070601 from the Grant Agency Czech Acad. Sci. and by Research Center Program MSMT – LC06077.

Appendix. Supplementary data

Supplementary data associated with this article can be found in the online version, at doi:10.1016/j.ibmb.2010.07.012.

References

Abraham, E.G., Sezutsu, H., Kanda, T., Sugasaki, T., Shimada, T., Tamura, T., 2000. Identification and characterisation of a silkworm ABC transporter gene homologous to Drosophila white. Mol. Gen. Genet. 264, 11–19.

Beumer, K., Bhattacharyya, G., Bibikova, M., Trautman, J.K., Carroll, D., 2006. Efficient gene targeting in Drosophila with zinc-finger nucleases. Genetics 172, 2391–2403.

Beumer, K.J., Trautman, J.K., Bozas, A., Liu, J.L., Rutter, J., Gall, J.G., Carroll, D., 2008. Efficient gene targeting in Drosophila by direct embryo injection with zinc-finger nucleases. Proc. Natl. Acad. Sci. U.S.A. 105, 19821–19826.

Bibikova, M., Golic, M., Golic, K.G., Carroll, D., 2002. Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. Genetics 161, 1169–1175.

Bozas, A., Beumer, K.J., Trautman, J.K., Carroll, D., 2009. Genetic analysis of zinc-finger nuclease-induced gene targeting in Drosophila. Genetics 182, 641–651.

Carroll, D., Beumer, K.J., Morton, J.J., Bozas, A., Trautman, J.K., 2008. Gene targeting in Drosophila and *Caenorhabditis elegans* with zinc-finger nucleases. Methods Mol. Biol. 435, 63–77.

Carroll, D., Morton, J.J., Beumer, K.J., Segal, D.J., 2006. Design, construction and in vitro testing of zinc finger nucleases. Nat. Protoc. 1, 1329–1341.

Desjarlais, Berg, J.M., 1992. Redesigning the DNA-binding specificity of a zinc finger protein: a data base-guided approach. Proteins 13, 272.

Doyon, Y., McCammon, J.M., Miller, J.C., Faraji, F., Ngo, C., Katibah, G.E., Amora, R., Hocking, T.D., Zhang, L., Rebar, E.J., et al., 2008. Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. Nat. Biotechnol. 26, 702–708.

Fu, F., Sander, J.D., Maeder, M., Thibodeau-Beganny, S., Joung, J.K., Dobbs, D., Miller, L., Voytas, D.F., 2009. Zinc Finger Database (ZiFDB): a repository for information on C2H2 zinc fingers and engineered zinc-finger arrays. Nucleic Acids Res. 37, D279–D283.

Fujii, T., Abe, H., Katsuma, S., Mita, K., Shimada, T., 2008. Mapping of sex-linked genes onto the genome sequence using various aberrations of the Z chromosome in *Bombyx mori*. Insect Biochem. Mol. Biol. 38, 1072–1079.

Geurts, A.M., Cost, G.J., Freyvert, Y., Zeitler, B., Miller, J.C., Choi, V.M., Jenkins, S.S., Wood, A., Cui, X., Meng, X., et al., 2009. Knockout rats via embryo microinjection of zinc-finger nucleases. Science 325, 433.

Imamura, M., Nakai, J., Inoue, S., Quan, G.X., Kanda, T., Tamura, T., 2003. Targeted gene expression using the GAL4/UAS system in the silkworm *Bombyx mori*. Genetics 165, 1329–1340.

Isobe, R., Kojima, K., Matsuyama, T., Quan, G.X., Kanda, T., Tamura, T., Sahara, K., Asano, S.I., Bando, H., 2004. Use of RNAi technology to confer enhanced resistance to BmNPV on transgenic silkworms. Arch. Virol. 149, 1931–1940.

Kim, Y.G., Cha, J., Chandrasegaran, S., 1996. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc. Natl. Acad. Sci. U.S.A. 93, 1156–1160.

Kim, J.-S., Lee, H.J., Carroll, D., 2010. Genome editing with modularly assembled zinc-finger nucleases. Nat. Methods 7, 91.

Kurkulos, M., Weinberg, J.M., Roy, D., Mount, S.M., 1994. P element-mediated in vivo deletion analysis of white-apricot: deletions between direct repeats are strongly favored. Genetics 136, 1001–1011.

Maeder, M.L., Thibodeau-Beganny, S., Osiak, A., Wright, D.A., Anthony, R.M., Eichtinger, M., Jiang, T., Foley, J.E., Winfrey, R.J., Townsend, J.A., et al., 2008. Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. Mol. Cell 31, 294–301.

Mandell, J.G., Barbas 3rd, C.F., 2006. Zinc finger tools: custom DNA-binding domains for transcription factors and nucleases. Nucleic Acids Res. 34, W516–W523.

Miller, J.C., Holmes, M.C., Wang, J., Guschin, D.Y., Lee, Y.L., Rupniewski, I., Beausejour, C.M., Waite, A.J., Wang, N.S., Kim, K.A., et al., 2007. An improved zinc-finger nuclease architecture for highly specific genome editing. Nat. Biotechnol. 25, 778–785.

Morton, J., Davis, M.W., Jorgensen, E.M., Carroll, D., 2006. Induction and repair of zinc-finger nuclease-targeted double-strand breaks in *Caenorhabditis elegans* somatic cells. Proc. Natl. Acad. Sci. U.S.A. 103, 16370–16375.

Nakao, H., 2009. Characterization of Bombyx embryo segmentation process: expression profiles of engrailed, even-skipped, caudal, and wnt1/wingless homologues. J. Exp. Zool. B Mol. Dev. Evol.

Nassif, N., Penney, J., Pal, S., Engels, W.R., Gloor, G.B., 1994. Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. Mol. Cell Biol. 14, 1613–1625.

Quan, G.X., Kim, I., Komoto, N., Sezutsu, H., Ote, M., Shimada, T., Kanda, T., Mita, K., Kobayashi, M., Tamura, T., 2002. Characterization of the kynurenine 3-monooxygenase gene corresponding to the white egg 1 mutant in the silkworm *Bombyx mori*. Mol. Genet. Genomics 267, 1–9.

Ramirez, C.L., Foley, J.E., Wright, D.A., Muller-Lerch, F., Rahman, S.H., Cornu, T.I., Winfrey, R.J., Sander, J.D., Fu, F., Townsend, J.A., et al., 2008. Unexpected failure rates for modular assembly of engineered zinc fingers. Nat. Methods 5, 374–375.

Rong, Y.S., Golic, K.G., 2000. Gene targeting by homologous recombination in Drosophila. Science 288, 2013–2018.

Smith, J., Bibikova, M., Whitby, F.G., Reddy, A.R., Chandrasegaran, S., Carroll, D., 2000. Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. Nucleic Acids Res. 28, 3361–3369.

Tamura, T., Sakate, S., 1983. Relationship between expression of oily character and uric acid incorporation in the larval integument of various oily mutants of the silkworm. Bull. Sericultural Exp. Station 28, 719–740.

Tamura, T., Thibert, C., Royer, C., Kanda, T., Abraham, E., Kamba, M., Komoto, N., Thomas, J.L., Mauchamp, B., Chavancy, G., et al., 2000. Germline transformation of the silkworm *Bombyx mori* L. using a piggyBac transposon-derived vector. Nat. Biotechnol. 18, 81–84.

Tomita, S., Kikuchi, A., 2009. Abd-B suppresses lepidopteran proleg development in posterior abdomen. Dev. Biol. 328, 403–409.

Uhlirova, M., Foy, B.D., Beaty, B.J., Olson, K.E., Riddiford, L.M., Jindra, M., 2003. Use of Sindbis virus-mediated RNA interference to demonstrate a conserved role of Broad-Complex in insect metamorphosis. Proc. Natl. Acad. Sci. U.S.A. 100, 15607–15612.

Uchino, K., Imamura, M., Shimizu, K., Kanda, T., Tamura, T., 2007. Germ line transformation of the silkworm, *Bombyx mori*, using the transposable element Minos. Mol. Genet. Genomics 277, 213–220.

Uchino, K., Sezutsu, H., Imamura, M., Kobayashi, I., Tatematsu, K., Iizuka, T., Yonemura, N., Mita, K., Tamura, T., 2008. Construction of a piggyBac-based enhancer trap system for the analysis of gene function in silkworm *Bombyx mori*. Insect Biochem. Mol. Biol. 38, 1165–1173.

Yamao, M., Katayama, N., Nakazawa, H., Yamakawa, M., Hayashi, Y., Hara, S., Kamei, K., Mori, H., 1999. Gene targeting in the silkworm by use of a baculovirus. Genes Dev. 13, 511–516.

Supplementary informations

Figure S1. Diagram showing a pair of ZFNs bound to BL-1 DNA target. Targetable triplets within the coding sequence of *BmBLOS2* gene were found by using the "Zinc Finger Tools" program (Mandell and Barbas, 2006) available on the Carlos Barbas laboratory website (http://zincfingertools). The best ***recognition*** amino acid motifs were chosen from the available candidate sequences (based on Barbas laboratory and Sangamo Biosciences data) by using a quality parameter "Q" (Carroll et al. 2006) as shown in the table S1A.

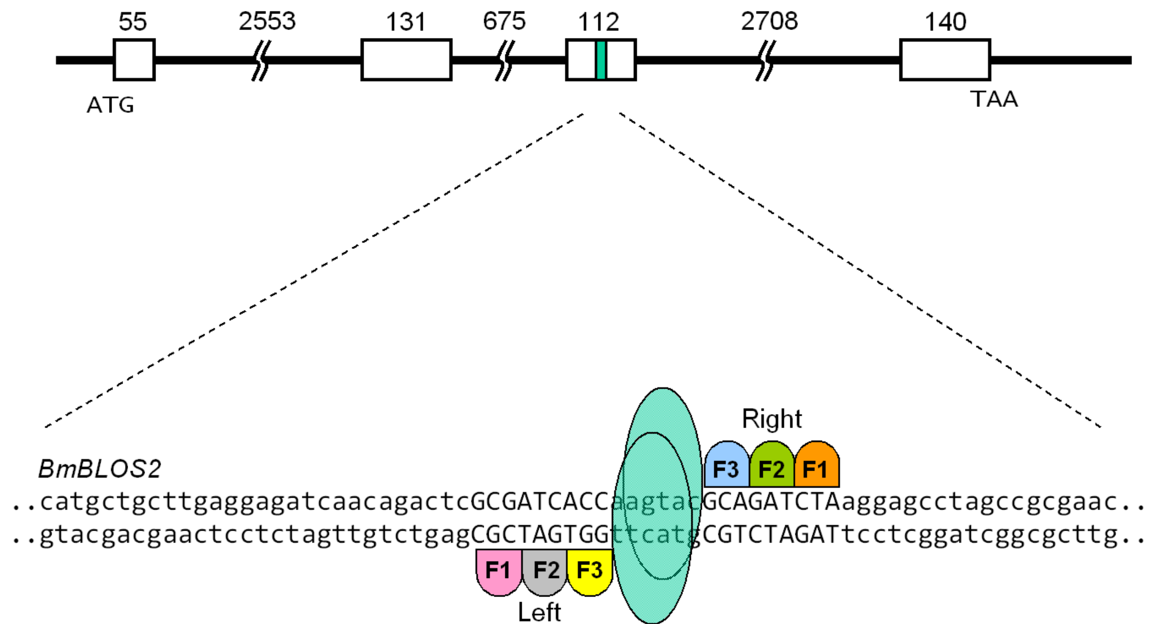


Figure S2. Design of BL-1-specific ZF. Resulting nucleotide and amino acid sequences of the left (**A**) and right (**B**) ZF coding regions in the Sp1C framework. The ***recognition*** motifs (selected from table S1A) are ***highlighted*** in different ***colors*** based on their position (see Fig. S1). Restriction enzyme sites *NdeI* and *SpeI* used for cloning are indicated.

A

```
      NdeI
gaggagcatatgcccggtgagaagccctatgcttgtccagaatgtggtaagtccttcagc
          M  P  G  E  K  P  Y  A  C  P  E  C  G  K  S  F  S
CACACAGGTCACCTGCTGGAAcaccagcgcacccatacgggtgaaaaaccatataaatgc
 H  T  G  H  L  L  E  H  Q  R  T  H  T  G  E  K  P  Y  K  C
ccagagtgcggcaagtctttcagtACTTCAGGAAATTTGGTTCGCcatcaacgcacccac
 P  E  C  G  K  S  F  S  T  S  G  N  L  V  R  H  Q  R  T  H
actggcgagaagccatacaaatgtccagaatgtggcaagtccttctctACCTCCGGCCAC
 T  G  E  K  P  Y  K  C  P  E  C  G  K  S  F  S  T  S  G  H
CTCGTCCGTcaccaacgcactcacaccggtcagaagcaactagtctgagaggag
 L  V  R  H  Q  R  T  H  T  G  Q  K  Q  L  V  -
                                          SpeI
```

B

```
      NdeI
gaggagcatatgcccggtgagaagccctatgcttgtccagaatgtggtaagtccttcagc
          M  P  G  E  K  P  Y  A  C  P  E  C  G  K  S  F  S
CAAAACTCTACACTGACAGAAcaccagcgcacccatacgggtgaaaaaccatataaatgc
 Q  N  S  T  L  T  E  H  Q  R  T  H  T  G  E  K  P  Y  K  C
ccagagtgcggcaagtctttcagtACTTCAGGAAATTTGGTTCGCcatcaacgcacccac
 P  E  C  G  K  S  F  S  T  S  G  N  L  V  R  H  Q  R  T  H
actggcgagaagccatacaaatgtccagaatgtggcaagtccttctctCAATCCGGCGAC
 T  G  E  K  P  Y  K  C  P  E  C  G  K  S  F  S  Q  S  G  D
CTCACCCGTcaccaacgcactcacaccggtcagaagcaactagtctgagaggag
 L  T  R  H  Q  R  T  H  T  G  Q  K  Q  L  V  -
                                          SpeI
```

Figure S3. Design of BL-1-specific ZF. Seven primers used for the synthesis of ZF coding regions (see Fig S10A for annealing scheme). The ***sequences encoding*** the triplet ***recognition*** motifs are highlighted in different colors. Oligonucleotides are approx. 60 nt long and their overlapping regions are underligned. The ***reverse*** primers ***are*** depicted in gray letters and are ***shown*** in reverse complement orientation. (**A**) - primers encoding the left ZFN; (**B**) - primers for the right ZFN.

A

gaggagcatatgcccggtgagaagccctatgcttgtccagaatgtggtaagtccttcagc

aatgtggtaagtccttcagcCACACAGGTCACCTGCTGGAAcaccagcgcacccatacgg

ccagcgcacccatacgggtgaaaaaccatataaatgcccagagtgcggcaagtctttcag

gtgcggcaagtctttcagtACTTCAGGAAATTTGGTTCGCcatcaacgcacccacact

catcaacgcacccacactggcgagaagccatacaaatgtccagaatgtggcaagtccttctc

atgtggcaagtccttctctACCTCCGGCCACCTCGTCCGTcaccaacgcactcacaccg

caccaacgcactcacaccggtcagaagcaactagtctgagaggag

B

gaggagcatatgcccggtgagaagccctatgcttgtccagaatgtggtaagtccttcagc

aatgtggtaagtccttcagcCAAAACTCTACACTGACAGAAcaccagcgcacccatacgg

ccagcgcacccatacgggtgaaaaaccatataaatgcccagagtgcggcaagtctttcag

gtgcggcaagtctttcagtACTTCAGGAAATTTGGTTCGCcatcaacgcacccacact

catcaacgcacccacactggcgagaagccatacaaatgtccagaatgtggcaagtccttctc

atgtggcaagtccttctctCAATCCGGCGACCTCACCCGTcaccaacgcactcacaccg

caccaacgcactcacaccggtcagaagcaactagtctgagaggag

Figure S4. Diagram showing a pair of ZFNs bound to BL-2 DNA target. Targetable triplets within the coding sequence of *BmBLOS2* gene were found by using the "Zinc Finger Tools" program (Mandell and Barbas, 2006) available on the Carlos Barbas laboratory website (http://zincfingertools). The best ***recognition*** amino acid motifs were chosen from the available candidate sequences (based on Barbas laboratory and Sangamo Biosciences data) by using a quality parameter "Q" (Carroll et al. 2006) as shown in the table S1B.

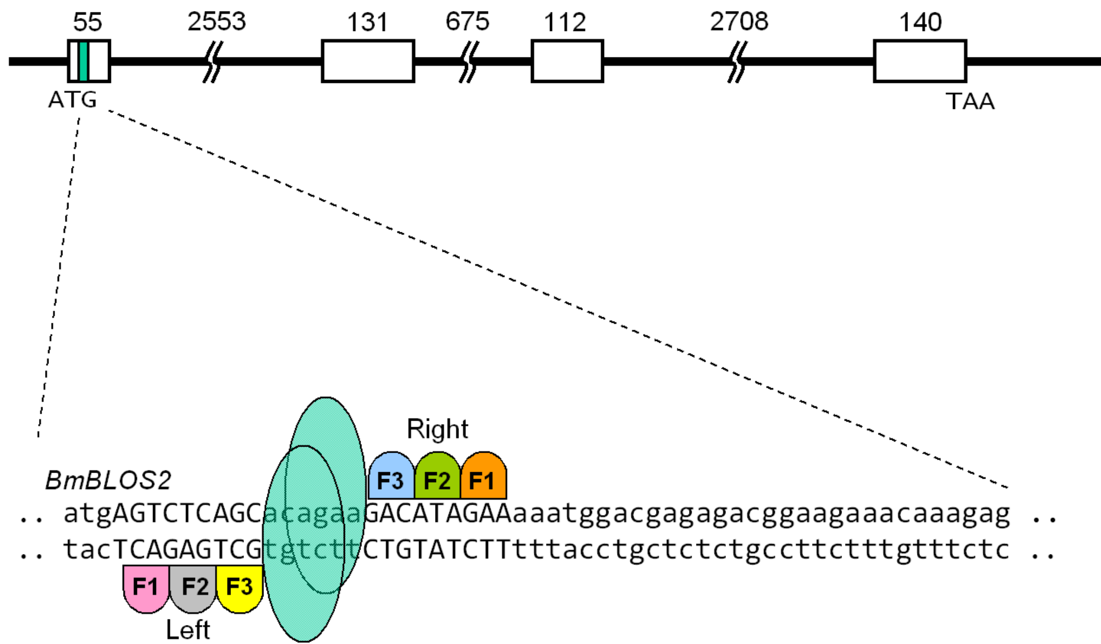


Figure S5. Design of BL-2-specific ZF. See legend of Figure S2 for description.

```
      NdeI
gaggagcatatgcccggtgagaagccctatgcttgtccagaatgtggtaagtccttcagc
      M  P  G  E  K  P  Y  A  C  P  E  C  G  K  S  F  S
ACACACCTGGACCTGATCAGAcaccagcgcacccatacgggtgaaaaaccatataaatgc
 T  H  L  D  L  I  R  H  Q  R  T  H  T  G  E  K  P  Y  K  C
ccagagtgcggcaagtctttcagtCGCTCAGATAATTTGGCCCGCcatcaacgcacccac
 P  E  C  G  K  S  F  S  R  S  D  N  L  A  R  H  Q  R  T  H
actggcgagaagccatacaaatgtccagaatgtggcaagtccttctctCAATCCTCCGAC
 T  G  E  K  P  Y  K  C  P  E  C  G  K  S  F  S  Q  S  S  D
CTCACCCGTcaccaacgcactcacaccggtcagaagcaactagtctgagaggag
 L  T  R  H  Q  R  T  H  T  G  Q  K  Q  L  V  -
                                        SpeI
```

```
      NdeI
gaggagcatatgcccggtgagaagccctatgcttgtccagaatgtggtaagtccttcagc
      M  P  G  E  K  P  Y  A  C  P  E  C  G  K  S  F  S
CAATCTGGTAACCTGGCTAGAcaccagcgcacccatacgggtgaaaaaccatataaatgc
 Q  S  G  N  L  A  R  H  Q  R  T  H  T  G  E  K  P  Y  K  C
ccagagtgcggcaagtctttcagtCAGAAGTCATCATTGATTGCCcatcaacgcacccac
 P  E  C  G  K  S  F  S  Q  K  S  S  L  I  A  H  Q  R  T  H
actggcgagaagccatacaaatgtccagaatgtggcaagtccttctctGACCGTTCCAAC
 T  G  E  K  P  Y  K  C  P  E  C  G  K  S  F  S  D  R  S  N
CTCACCCGTcaccaacgcactcacaccggtcagaagcaactagtctgagaggag
 L  T  R  H  Q  R  T  H  T  G  Q  K  Q  L  V  -
                                        SpeI
```

Figure S6. Design of BL-2-specific ZF using seven overlapping primers. See legend of Figure S3 for description.

A

gaggagcatatgcccggtgagaagccctatgcttgtccagaatgtggtaagtccttcagc

aatgtggtaagtccttcagcACACACCTGGACCTGATCAGAcaccagcgcacccatacgg

ccagcgcacccatacgggtgaaaaaccatataaatgcccagagtgcggcaagtctttcag

gtgcggcaagtctttcagtCGCTCAGATAATTTGGCCCGCcatcaacgcacccacact

catcaacgcacccacactggcgagaagccatacaaatgtccagaatgtggcaagtccttctc

atgtggcaagtccttctctCAATCCTCCGACCTCACCCGTcaccaacgcactcacaccg

caccaacgcactcacaccggtcagaagcaactagtctgagaggag

B

gaggagcatatgcccggtgagaagccctatgcttgtccagaatgtggtaagtccttcagc

aatgtggtaagtccttcagcCAATCTGGTAACCTGGCTAGAcaccagcgcacccatacgg

ccagcgcacccatacgggtgaaaaaccatataaatgcccagagtgcggcaagtctttcag

gtgcggcaagtctttcagtCAGAAGTCATCATTGATTGCCcatcaacgcacccacact

catcaacgcacccacactggcgagaagccatacaaatgtccagaatgtggcaagtccttctc

atgtggcaagtccttctctGACCGTTCCAACCTCACCCGTcaccaacgcactcacaccg

caccaacgcactcacaccggtcagaagcaactagtctgagaggag

Figure S7. Diagram showing a pair of ZFNs bound to BW-1 DNA target. Targetable triplets within the coding sequence of *Bmwh3* gene were found by using the "Zinc Finger Tools" program (Mandell and Barbas, 2006) available on the Carlos Barbas laboratory website (http://zincfingertools). The best ***recognition*** amino acid motifs were chosen from the available candidate sequences (based on Barbas laboratory and Sangamo Biosciences data) by using a quality parameter "Q" (Carroll et al. 2006) as shown in the table S1C.

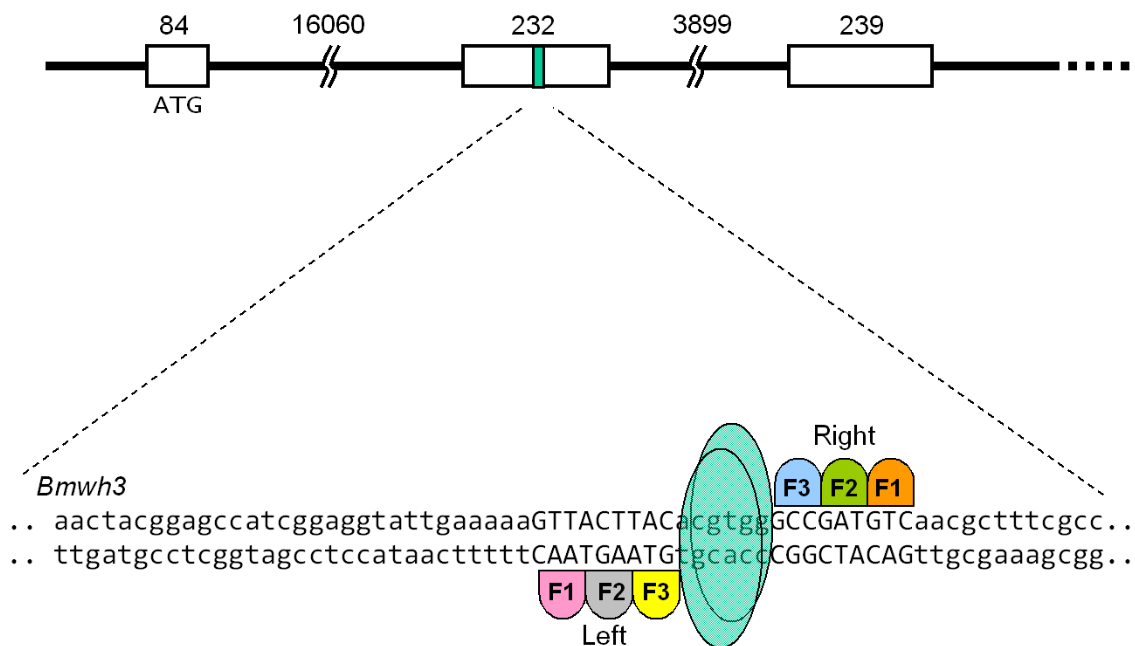


Figure S8. Design of BW-1-specific ZF. See legend of Figure S2 for description.

```
NdeI
gaggagcatatgcccggtgagaagccctatgcttgtccagaatgtggtaagtccttcagc
            M  P  G  E  K  P  Y  A  C  P  E  C  G  K  S  F  S
GACCCTGGTGCTCTGGTGAGAcaccagcgcacccatacgggtgaaaaaccatataaatgc
 D  P  G  A  L  V  R  H  Q  R  T  H  T  G  E  K  P  Y  K  C
ccagagtgcggcaagtctttcagtACTTCAGGAAATTTGGTTCGCcatcaacgcacccac
 P  E  C  G  K  S  F  S  T  S  G  N  L  V  R  H  Q  R  T  H
actggcgagaagccatacaaatgtccagaatgtggcaagtccttctctGACTGCCGTGAC
 T  G  E  K  P  Y  K  C  P  E  C  G  K  S  F  S  D  C  R  D
CTCGCACGTcaccaacgcactcacaccggtcagaagcaactagtctgagaggag
 L  A  R  H  Q  R  T  H  T  G  Q  K  Q  L  V  -
                                        SpeI
```

```
NdeI
gaggagcatatgcccggtgagaagccctatgcttgtccagaatgtggtaagtccttcagc
            M  P  G  E  K  P  Y  A  C  P  E  C  G  K  S  F  S
GACTCTGGTAACCTGAGAGTGcaccagcgcacccatacgggtgaaaaaccatataaatgc
 D  S  G  N  L  V  R  H  Q  R  T  H  T  G  E  K  P  Y  K  C
ccagagtgcggcaagtctttcagtCATCGCACTACTTTGACTAATcatcaacgcacccac
 P  E  C  G  K  S  F  S  H  R  T  T  L  T  N  H  Q  R  T  H
actggcgagaagccatacaaatgtccagaatgtggcaagtccttctctCAATCCTCCTCC
 T  G  E  K  P  Y  K  C  P  E  C  G  K  S  F  S  Q  S  S  S
CTCGTCCGTcaccaacgcactcacaccggtcagaagcaactagtctgagaggag
 L  V  R  H  Q  R  T  H  T  G  Q  K  Q  L  V  -
                                        SpeI
```

Figure S9. Design of BW-1-specific ZF using seven overlapping primers. See legend of Figure S3 for description.

A

gaggagcatatgcccggtgagaagccctatgcttgtccagaatgtggtaagtccttcagc

aatgtggtaagtccttcagcGACTCTGGTAACCTGAGAGTGcaccagcgcacccatacgg

ccagcgcacccatacgggtgaaaaaccatataaatgcccagagtgcggcaagtctttcag

gtgcggcaagtctttcagtCATCGCACTACTTTGACTAATcatcaacgcacccacact

catcaacgcacccacactggcgagaagccatacaaatgtccagaatgtggcaagtccttctc

atgtggcaagtccttctctCAATCCTCCTCCCTCGTCCGTcaccaacgcactcacaccg

caccaacgcactcacaccggtcagaagcaactagtctgagaggag

B

gaggagcatatgcccggtgagaagccctatgcttgtccagaatgtggtaagtccttcagc

aatgtggtaagtccttcagcGACCCTGGTGCTCTGGTGAGAcaccagcgcacccatacgg

ccagcgcacccatacgggtgaaaaaccatataaatgcccagagtgcggcaagtctttcag

gtgcggcaagtctttcagtACTTCAGGAAATTTGGTTCGCcatcaacgcacccacact

catcaacgcacccacactggcgagaagccatacaaatgtccagaatgtggcaagtccttctc

atgtggcaagtccttctctGACTGCCGTGACCTCGCACGTcaccaacgcactcacaccg

caccaacgcactcacaccggtcagaagcaactagtctgagaggag

Figure S10. ***Schematic overview of the procedure*** for ZFN RNA preparation. The procedure involves (**A**) the annealing of oligonucleotides and overlap extension PCR used for synthesis of the ZF coding region. The products are fractionated by electrophoresis on a 2% agarose gel. The bands of desired size (294 bp) are cut out and the DNAs are recovered with a Qiagen MinElute column. (**B**) The PCR products are cloned into a suitable vector (pENTR-NLS-ZFN), in frame with the FokI cleavage domain. (**C**) The inserts are finely transfered to the pCS2-DEST vector (which confers ampicillin resistance) using Clonase site-specific recombination reactions. (**D**) Vectors are linearized and RNAs are synthesized *in vitro*. L1, L2, B1 and B2 are *att* sites used by Clonase.

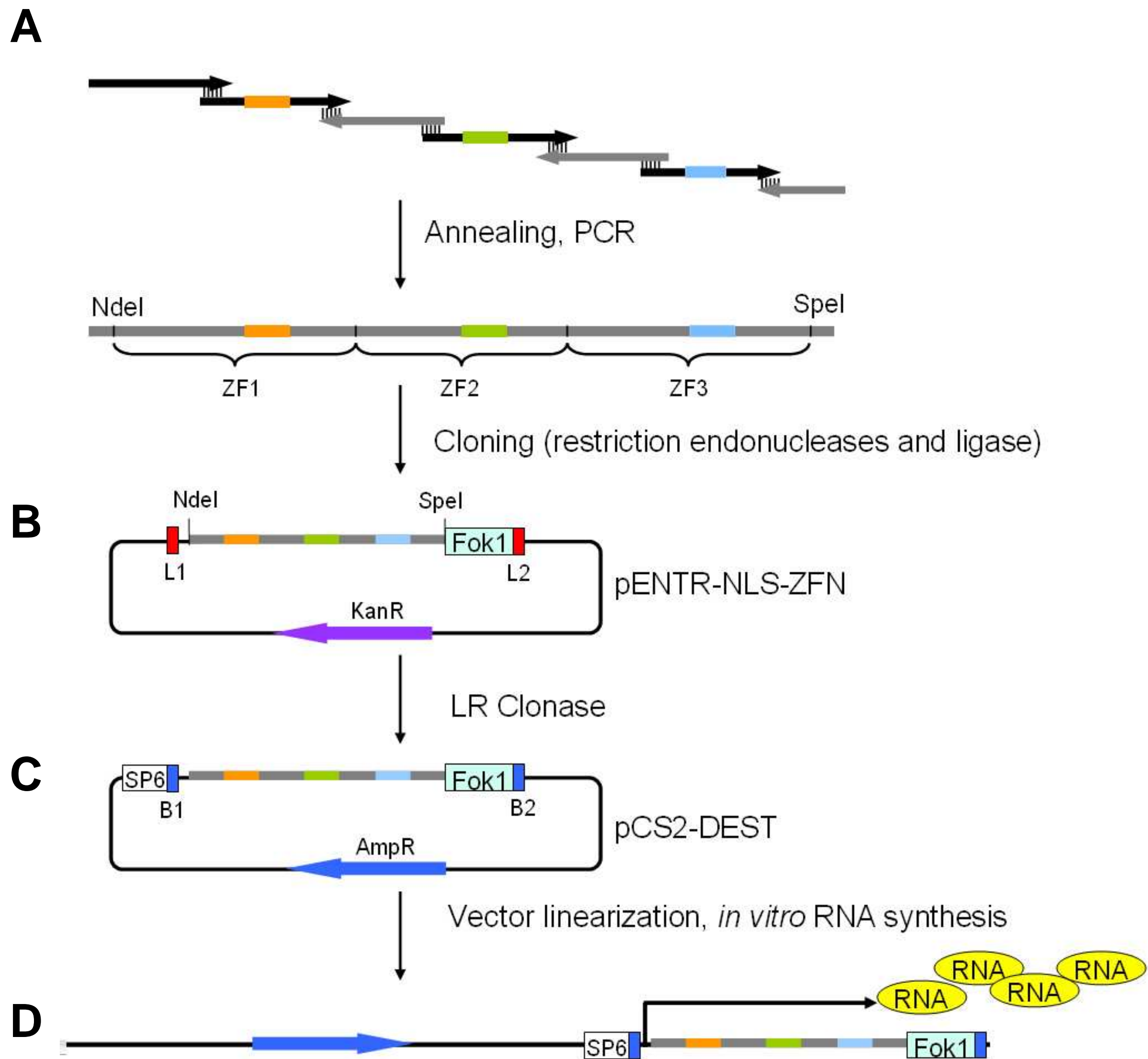


Figure S11. Direct ZFN RNA microinjection of silkworm embryos and the crossing scheme used for the detection of *BmBLOS2* mutants (The *BmBLOS2* gene is located on the Z chromosome, females are hemizygous for Z). Some of the G0 larvae display mosaicism in the epidermis. Some larve with *oily* phenotype, carrying the *BmBLOS2* germline mutations, appear among the female G1 larvae. * denotes mutagenized allele.

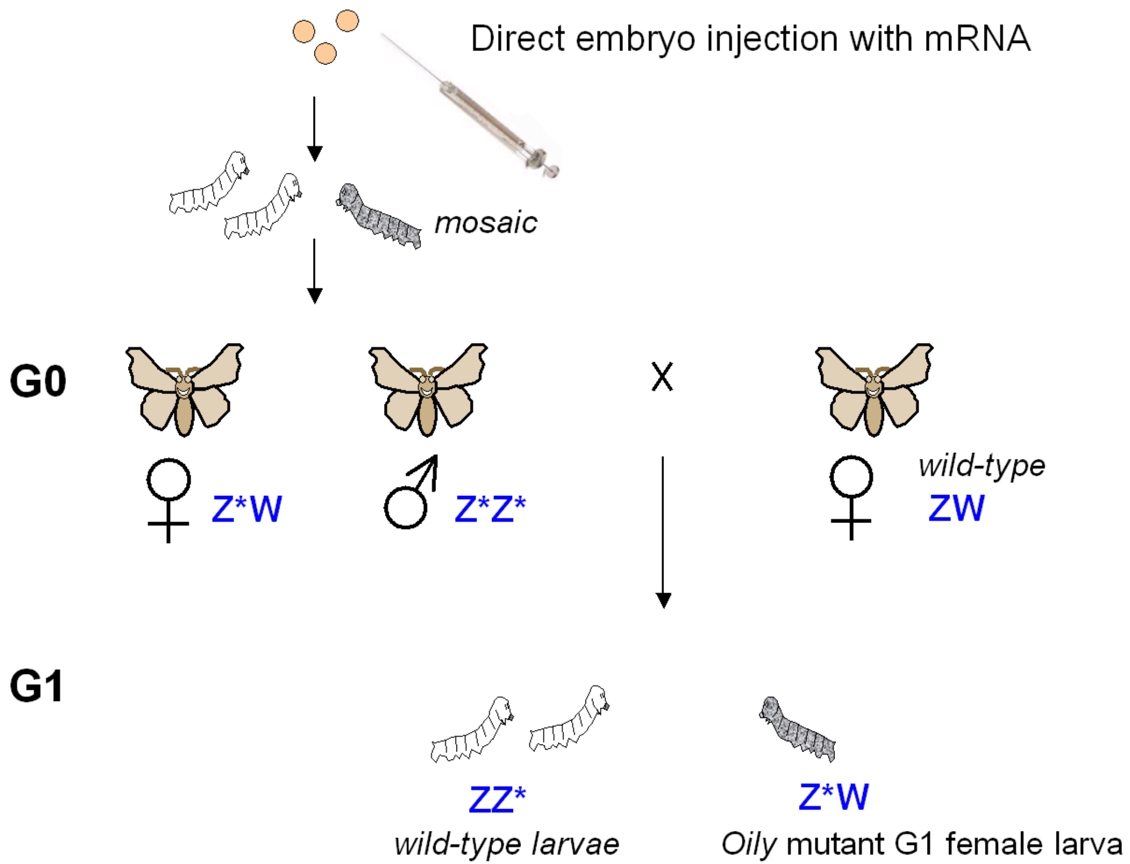


Figure S12. Direct ZFN RNA microinjection of silkworm embryos and the crossing scheme used for the detection of *Bmwh3* somatic mosaics (**A**) and G1 mutants (**B**). * denotes novel allele.

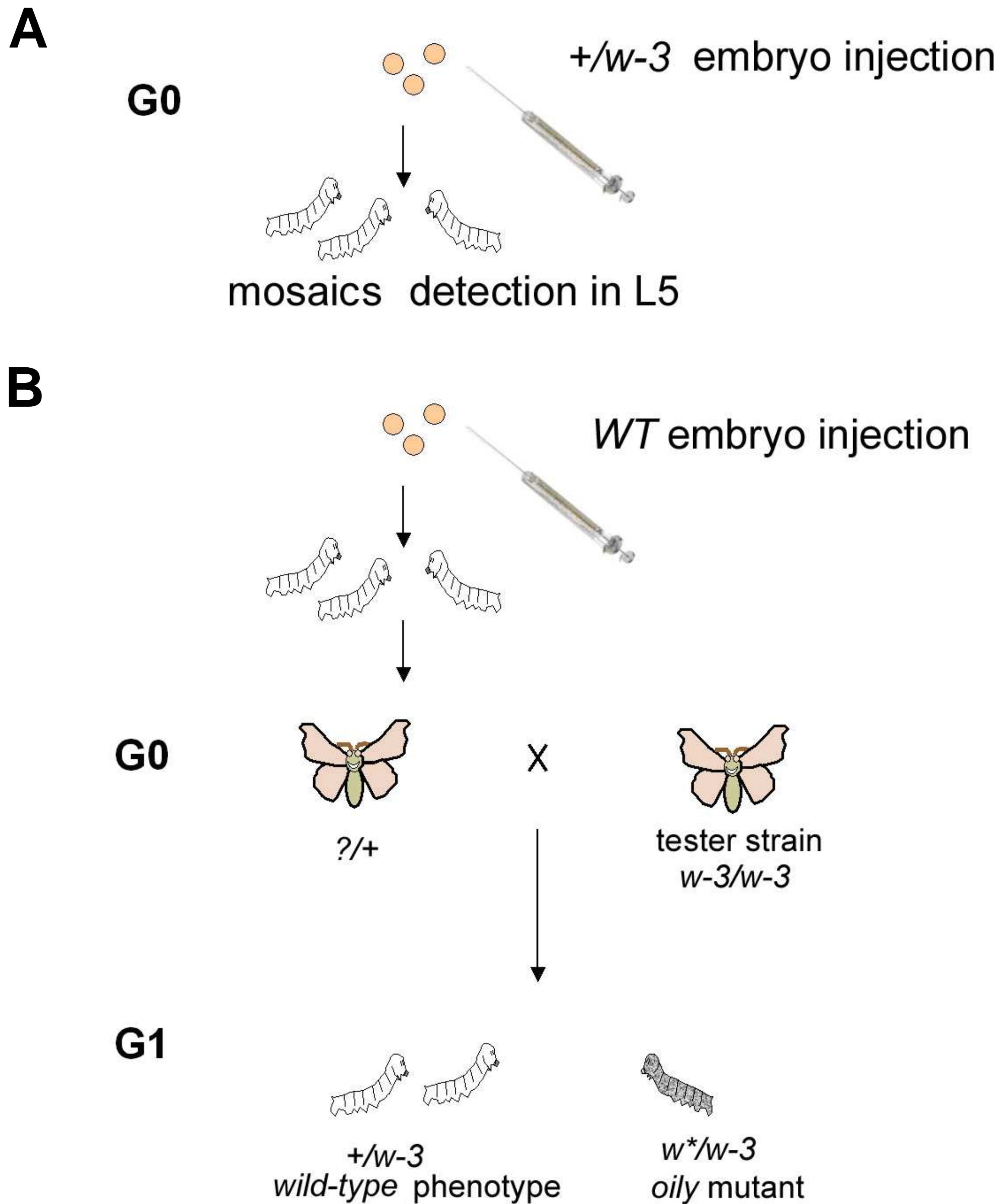
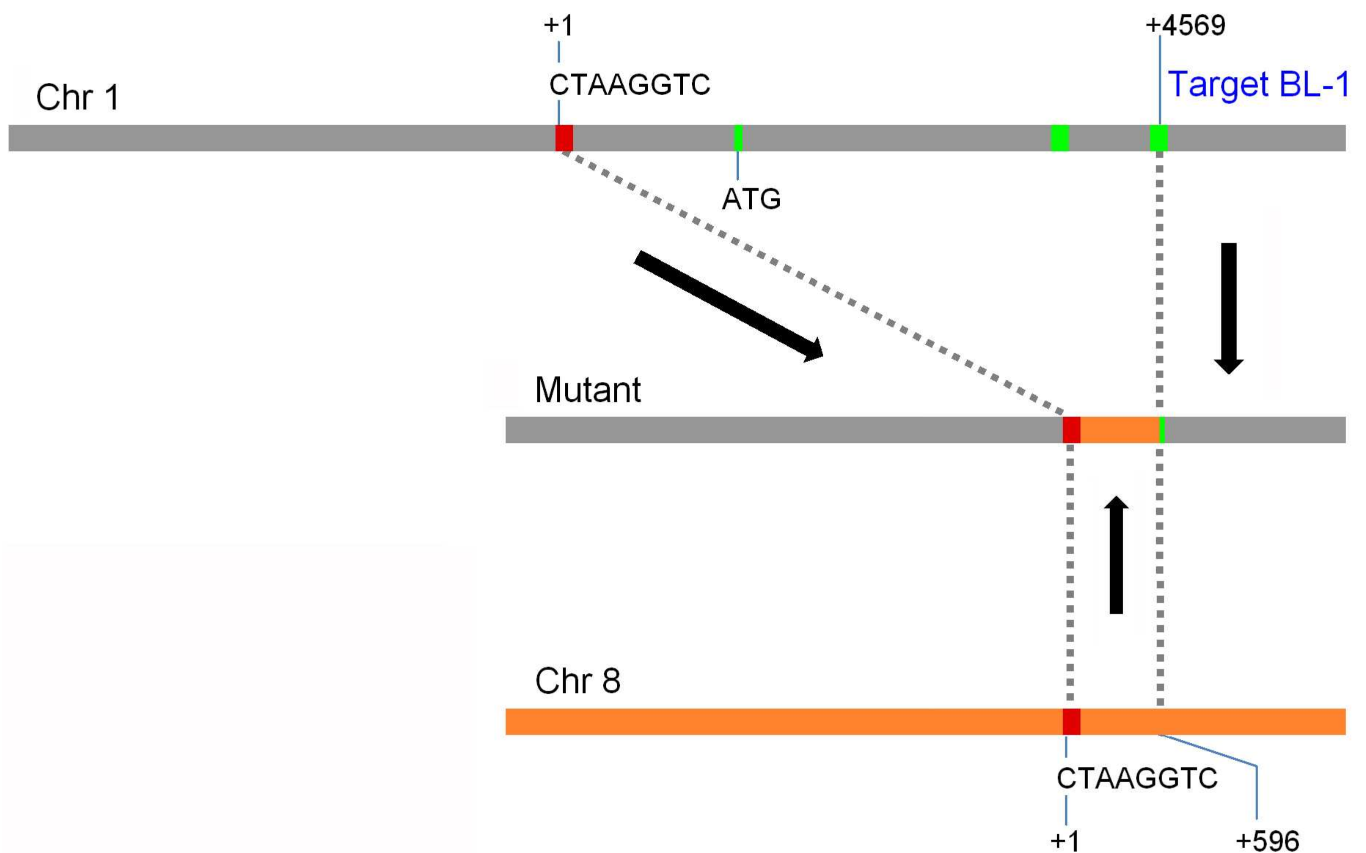


FIGURE S13. ZFN-induced deletion-insertion mutation at *BmBLOS2* locus found in one of the brood 19 silkworms. The size of the deletion was 4569 bp, and that of the filler sequence 596 bp. The inserted sequence was long enough to localize its origin in chromosome 8. The 131 bp patch of homology between the wild-type (chromosome 1) and the donor sequence (chromosome 8) was identified as a *Bm1* element. (**A**) The diagram shows the comparison of chromosomal regions of the wild-type *BmBLOS2* region (above), mutant (middle) and chromosome 8 (below). First three *BmBLOS2* exons are displayed as green rectangles. (**B**) The 5' end junction sequence. Comparison of wild-type sequence (above), mutant (middle) and putative donor from chromosome 8 (below). The first position in the sequence is numbered from the start of *Bm1* element. The homologous region of Bm1 element is framed in red; bases identical with donor sequence are boxed in orange. (**C**) The 3' end junction sequence. ZFN recognition sequence is framed in blue; numbers on the left indicate distances in base pair from the ends of sequences shown in Figure S13B; other features as in Figure S13B.



B

```
                 +1
Chr1:    TTTTAAATTATATCACTAAGGTCTCAGTATAGTTACAACGGCTGCCCCACCCTTCAAACCGAAACGCATTACTGCTTCACGGCAGAAATAGGCAGTGCGG   85
Mutant:  TTTTAAATTATATCACTAAGGTCTCAGTATAGTTACAACGGCTGCCCCACCCTTCAAACCGAAACGCATTACTGCTTCACGGTAGAAATAGGCAGGGCGG   85
Chr8:    CTTGAGATATAAGTTCTAAGGTCTCAAGTAAGTTACAACGGCTGCCCCACCCTTCAAACCGAAACGCATTACTGCTTCACGGCAGAAATAGGCAGGTGG   85

Chr1:    TGGTACCTACCCGCGCGGACTCACAAGAGGTCCTACCACCAGTAATTACGCAAATTAAAATTTTGCGGGTTTATTACACGATGTTATTCTTTGTTATTGT  185
Mutant:  TGGTACCCACCCGCGCGGACTCACAAAAGGTCCTACCACCAGTAATAGTATGGATGAACTCTACTCCGGTGCGGGCGGTGCGATGGTAAAAACGAGATGA  185
Chr8:    TGGTACCCACCCGCGCGGACTCACAAAAGGTCCTACCACCAGTAATAGTATTGCTGAACTCTACTAAGATGCGGGTCGTCCAATGGTAAAAAATAGATGA  185
```

C

Target BL-1

```
Chr1:    ......4346 bp......GCTGCTTGAGGAGATCAACAGACTCGCGATCACCAAGTACGCAGATCTAAGGAGCCTAGCCGCGAACCTGAACAAGACCCT  4613
Mutant:  ......374 bp......GGATTACTATCTTGAAATTTCGAAATCCAGTATTTTAGTTGCAGATCTAAGGAGCCTAGCCGCGAACCTGAACAAGACCCT  639
Chr8:    ......374 bp......GGATTACTATCTTGAAATTTCGAAATCCAGGATTTTAGCTGGTACTATTTATTGGTAGATTGACAACTCGACTAATAACAA  639
```

Table S1A. Evaluation of zinc finger α-helix recognition specificity according to Carroll et al. (2006). Once the best available ZFN target site BL-1 (Fig. S1) was found (individual DNA triplets are shown in red letters on yellow background) by the "Zinc Finger Tools" program (see Materials and Methods), the best ***recognition*** amino acid sequences (shown in red letters on white background) were chosen from Barbas laboratory (B) and Sangamo Biosciences (S) "finger" data by using a quality parameter "Q". This parameter is an estimate of binding based on published information (the relative specificity of one, two or three +'s).

Triplet	Finger (B)	Q	Finger (S)	Q	Triplet	Finger (B)	Q	Triplet	Finger (B)	Q	Triplet	Finger (B)	Q
GAA	QSSNLVR	+	QSGNLAR	++	AAA	QRANLRA	+	CAA	QSGNLTE	-	TAA		
GAC	DPGNLVR	+	DRSNLTR	+	AAC	DSGNLRV	++	CAC	SKKALTE	-	TAC		
GAG	RSDNLVR	++	RSDNLAR	+++	AAG	RKDNLKN	+	CAG	RADNLTE	++	TAG	REDNLHT	-
GAT	**TSGNLVR**	+++	TSANLSR	+++	AAT	TTGNLTV	-	CAT	TSGNLTE	+	TAT		
GCA	QSGDLRR	+	**QSGDLTR**	++	ACA	SPADLTR	++	CCA	TSHSLTE	+	TCA		
GCC	DCRDLAR	+++	DRSDLTR	+	ACC	DKKDLTR	+	CCC	SKKHLAE	+	TCC		
GCG	RSDDLVR	-	RSDDLQR	+++	ACG	RTDTLRD	-	CCG	RNDTLTE	+	TCG		
GCT	TSGELVR	+	QSSDLTR	+++	ACT	THLDLIR	+++	CCT	TKNSLTE	++	TCT		
GGA	QRAHLER	++	QSGHLQR	++	AGA	QLAHLRA	+	CGA	QSGHLTE	-	TGA	QAGHLAS	++
GGC	DPGHLVR	+	DRSHLAR	+	AGC			**CGC**	**HTGHLLE**	++	TGC		
GGG	RSDKLVR	+++	RSDHLSR	+++	AGG	RSDHLTN	+	CGG	RSDKLTE	-	TGG	RSDHLTT	+
GGT	TSGHLVR	+	**TSGHLVR**	++	AGT	HRTTLTN	+	CGT	SRRTCRA	++	TGT		
GTA	QSSSLVR	+++	QSGALAR	+	ATA	QKSSLIA	+	**CTA**	**QNSTLTE**	++	TTA		
GTC	DPGALVR	++	DRSALAR	+	ATC			CTC			TTC		
GTG	RSDELVR	+	RSDALTR	+++	ATG	RRDELNV	-	CTG	RNDALTE	++	TTG		
GTT	TSGSLVR	-	TSGALTR	+	ATT	HKNALQN	+	CTT	TTGALTE	-	TTT		

Table S1B. Zinc finger α-helix recognition specificity evaluation table according to Carroll et al. (2006). The ZFN target site BL-2 (Fig. S4) was found (individual DNA triplets are shown in red letters on yellow background) by the "Zinc Finger Tools" program (see Materials and Methods), the ***recognition*** amino acid sequences (shown in red letters on white background) were chosen from Barbas laboratory (B) and Sangamo Biosciences (S) "finger" data by using a quality parameter "Q". This parameter is an estimate of binding based on published information.

Triplet	Finger (B)	Q	Finger (S)	Q	Triplet	Finger (B)	Q	Triplet	Finger (B)	Q	Triplet	Finger (B)	Q
GAA	QSSNLVR	+	**QSGNLAR**	++	AAA	QRANLRA	+	CAA	QSGNLTE	-	TAA		
GAC	DPGNLVR	+	**DRSNLTR**	+	AAC	DSGNLRV	++	CAC	SKKALTE	-	TAC		
GAG	RSDNLVR	++	**RSDNLAR**	+++	AAG	RKDNLKN	+	CAG	RADNLTE	++	TAG	REDNLHT	-
GAT	TSGNLVR	+++	TSANLSR	+++	AAT	TTGNLTV	-	CAT	TSGNLTE	+	TAT		
GCA	QSGDLRR	+	QSGDLTR	++	ACA	SPADLTR	++	CCA	TSHSLTE	+	TCA		
GCC	DCRDLAR	+++	DRSDLTR	+	ACC	DKKDLTR	+	CCC	SKKHLAE	+	TCC		
GCG	RSDDLVR	-	RSDDLQR	+++	ACG	RTDTLRD	-	CCG	RNDTLTE	+	TCG		
GCT	TSGELVR	+	**QSSDLTR**	+++	**ACT**	**THLDLIR**	+++	CCT	TKNSLTE	++	TCT		
GGA	QRAHLER	++	QSGHLQR	++	AGA	QLAHLRA	+	CGA	QSGHLTE	-	TGA	QAGHLAS	++
GGC	DPGHLVR	+	DRSHLAR	+	AGC			CGC	HTGHLLE	++	TGC		
GGG	RSDKLVR	+++	RSDHLSR	+++	AGG	RSDHLTN	+	CGG	RSDKLTE	-	TGG	RSDHLTT	+
GGT	TSGHLVR	+	TSGHLVR	++	AGT	HRTTLTN	+	CGT	SRRTCRA	++	TGT		
GTA	QSSSLVR	+++	QSGALAR	+	**ATA**	**QKSSLIA**	+	CTA	QNSTLTE	++	TTA		
GTC	DPGALVR	++	DRSALAR	+	ATC			CTC			TTC		
GTG	RSDELVR	+	RSDALTR	+++	ATG	RRDELNV	-	CTG	RNDALTE	++	TTG		
GTT	TSGSLVR	-	TSGALTR	+	ATT	HKNALQN	+	CTT	TTGALTE	-	TTT		

Table S1C. Zinc finger α-helix recognition specificity evaluation table according to Carroll et al. (2006). The ZFN target site BW-1 *(*Fig. S7) was found (individual DNA triplets are shown in red letters on yellow background) by the "Zinc Finger Tools" program (see Materials and Methods), the ***recognition*** amino acid sequences (shown in red letters on white background) were chosen from Barbas laboratory (B) and Sangamo Biosciences (S)"finger" data by using a quality parameter "Q". This parameter is an estimate of binding based on published information.

.

Triplet	Finger (B)	Q	Finger (S)	Q	Triplet	Finger (B)	Q	Triplet	Finger (B)	Q	Triplet	Finger (B)	Q
GAA	QSSNLVR	+	QSGNLAR	++	AAA	QRANLRA	+	CAA	QSGNLTE	-	TAA		
GAC	DPGNLVR	+	DRSNLTR	+	AAC	DSGNLRV	++	CAC	SKKALTE	-	TAC		
GAG	RSDNLVR	++	RSDNLAR	+++	AAG	RKDNLKN	+	CAG	RADNLTE	++	TAG	REDNLHT	-
GAT	TSGNLVR	+++	TSANLSR	+++	AAT	TTGNLTV	-	CAT	TSGNLTE	+	TAT		
GCA	QSGDLRR	+	QSGDLTR	++	ACA	SPADLTR	++	CCA	TSHSLTE	+	TCA		
GCC	DCRDLAR	+++	DRSDLTR	+	ACC	DKKDLTR	+	CCC	SKKHLAE	+	TCC		
GCG	RSDDLVR	-	RSDDLQR	+++	ACG	RTDTLRD	-	CCG	RNDTLTE	+	TCG		
GCT	TSGELVR	+	QSSDLTR	+++	ACT	THLDLIR	+++	CCT	TKNSLTE	++	TCT		
GGA	QRAHLER	++	QSGHLQR	++	AGA	QLAHLRA	+	CGA	QSGHLTE	-	TGA	QAGHLAS	++
GGC	DPGHLVR	+	DRSHLAR	+	AGC			CGC	HTGHLLE	++	TGC		
GGG	RSDKLVR	+++	RSDHLSR	+++	AGG	RSDHLTN	+	CGG	RSDKLTE	-	TGG	RSDHLTT	+
GGT	TSGHLVR	+	TSGHLVR	++	AGT	HRTTLTN	+	CGT	SRRTCRA	++	TGT		
GTA	QSSSLVR	+++	QSGALAR	+	ATA	QKSSLIA	+	CTA	QNSTLTE	++	TTA		
GTC	DPGALVR	++	DRSALAR	+	ATC			CTC			TTC		
GTG	RSDELVR	+	RSDALTR	+++	ATG	RRDELNV	-	CTG	RNDALTE	++	TTG		
GTT	TSGSLVR	-	TSGALTR	+	ATT	HKNALQN	+	CTT	TTGALTE	-	TTT		

Table S2. Statistical evaluation of the cross between heterozygous *BmBLOS*/+ G_2 males with *pnd (wt)* females. The expected 3:1 *ratio* **of** *phenotype segregation* **was supported by χ^2 statistical test.**
* $0.01<P<0.05$ (marginally significant)

Line	Observed		Expected		P value
	Normal larvae	Mutant larvae	Normal larvae	Mutant larvae	
9A	237	87	243	81	0,441
9B	250	86	252	84	0,801
9C	164	56	165	55	0,876
19A	174	63	177,8	59,2	0,574
65A	217	67	213	71	0,583
65B	237	58	221,2	73,8	0,034*
69A	178	58	177	59	0,881
69B	206	81	215,2	71,8	0,207
69C	220	80	225	75	0,505
MixA	282	95	282,8	94,2	0,929
MixB	189	74	197,2	65,8	0,24

CONCLUSIONS

I. Our work demonstrated that custom ZFN mRNA injection is effective to induce somatic and germline mutations in a targeted gene by non-homologous end joining (NHEJ) mechanism in *Bombyx,* albeit at low frequency. Direct embryonic microinjection of mRNA showed the potential of avoiding the labourious construction of transgenic strains with DNA-modifying enzymes. Our results also support the observations made in other experimental studies about the low success rate and toxicity associated with modularly assembled ZFNs. The observed mutation spectrum suggested that the double-strand break repair system relies on microhomologies rather than on a canonical ligase IV dependent mechanism.

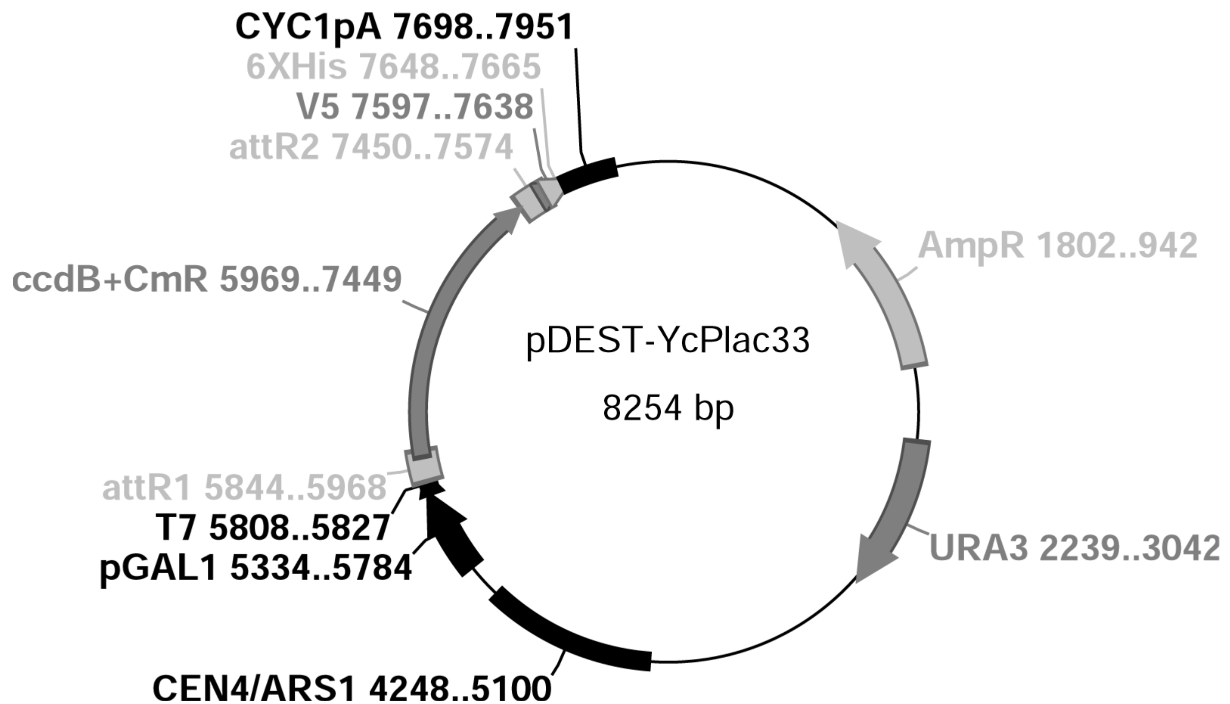
II. Our results showed that full length TALENs can be successfully used for gene mutagenesis in *Bombyx* and confirm the utility of these TALENs especially at loci lacking canonical ZFN target sites. Despite the fact that ZFNs are of a smaller size and their amino acid sequences are less repetitive in nature, which may give them an advantage for certain applications, we did not observe any disadvantages of TALENs concerning these parameters. We found that the design and construction of TALENs was very simple and fast and the success rate was better than with ZFNs since all three TALENs examined yielded germline mutants.We also showed that TALEN activity can be preexamined in yeast, further increasing the potential efficacy of the more difficult and time-consuming steps required for germline introduction and screening for mutations in the host.

Our research demonstrated that ZFNs and full length TALENs offer the potential of robust tools and represent an important step towards the routine use of these reagents in gene targeting applications in the silkworm.

APPENDIX

Sequence feature of the vectors generated during the course of this study

Sequence features of pDEST-YcPlac33



Plasmid components :

pGAL1, Full length yeast galactose inducible promoter for high-level protein expression.

T7, promoter for in vitro transcrition.

***att*R1** and ***att*R2**, recombination sites, for LR cloning of the gene of interest from an entry clone to destination vector.

***ccd*B** and **CmR** (Chloramphenicol resistance) genes for negative selection and counterselection, respectively.

V5 epitope and **6xHis** (Polyhistidine) tag for detection and purification (if required).

CYC1pA, Cytochrome c polyadenylation sequence for proper termination and processing of the transcript.

CEN4/ARS1, region for episomal maintenance and low copy replication.

URA3, auxotrophic marker for selection of yeast transformants.

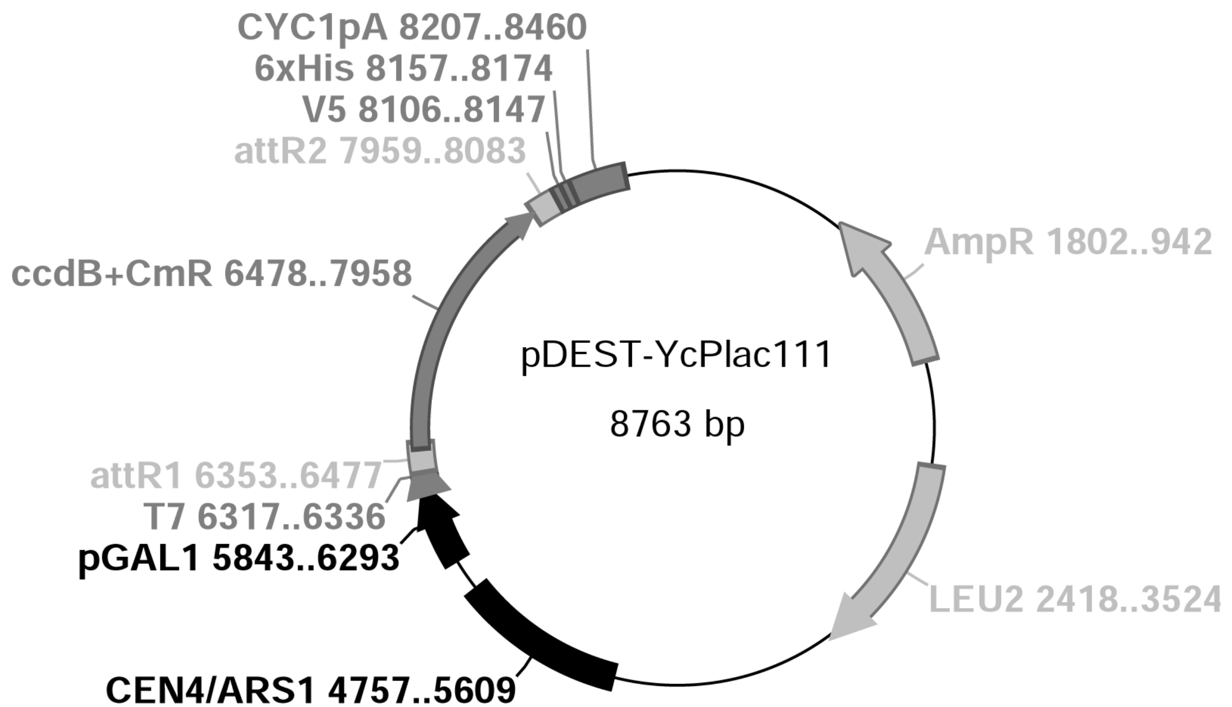
AmpR, ampicillin resistance gene for selection in *E. coli.*

```
   1 tcttccgctt cctcgctcac tgactcgctg cgctcggtcg ttcggctgcg gcgagcggta 60
  61 tcagctcact caaaggcggt aatacggtta tccacagaat caggggataa cgcaggaaag 120
 121 aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg 180
 181 tttttccata ggctccgccc ccctgacgag catcacaaaa atcgacgctc aagtcagagg 240
 241 tggcgaaacc cgacaggact ataaagatac caggcgtttc cccctggaag ctccctcgtg 300
 301 cgctctcctg ttccgaccct gccgcttacc ggatacctgt ccgcctttct cccttcggga 360
 361 agcgtggcgc tttctcatag ctcacgctgt aggtatctca gttcggtgta ggtcgttcgc 420
 421 tccaagctgg gctgtgtgca cgaacccccc gttcagcccg accgctgcgc cttatccggt 480
 481 aactatcgtc ttgagtccaa cccggtaaga cacgacttat cgccactggc agcagccact 540
 541 ggtaacagga ttagcagagc gaggtatgta ggcggtgcta cagagttctt gaagtggtgg 600
 601 cctaactacg gctacactag aaggacagta tttggtatct gcgctctgct gaagccagtt 660
 661 accttcggaa aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcggt 720
 721 ggtttttttg tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct 780
 781 ttgatctttt ctacggggtc tgacgctcag tggaacgaaa actcacgtta agggattttg 840
 841 gtcatgagat tatcaaaaag gatcttcacc tagatccttt taaattaaaa atgaagtttt 900
 901 aaatcaatct aaagtatata tgagtaaact tggtctgaca gttaccaatg cttaatcagt 960
 961 gaggcaccta tctcagcgat ctgtctattt cgttcatcca tagttgcctg actccccgtc 1020
1021 gtgtagataa ctacgatacg ggagggctta ccatctggcc ccagtgctgc aatgataccg 1080
1081 cgagacccac gctcaccggc tccagattta tcagcaataa accagccagc cggaagggcc 1140
1141 gagcgcagaa gtggtcctgc aactttatcc gcctccatcc agtctattaa ttgttgccgg 1200
1201 gaagctagag taagtagttc gccagttaat agtttgcgca acgttgttgc cattgctaca 1260
1261 ggcatcgtgg tgtcacgctc gtcgtttggt atggcttcat tcagctccgg ttcccaacga 1320
1321 tcaaggcgag ttacatgatc ccccatgttg tgcaaaaaag cggttagctc cttcggtcct 1380
1381 ccgatcgttg tcagaagtaa gttggccgca gtgttatcac tcatggttat ggcagcactg 1440
1441 cataattctc ttactgtcat gccatccgta agatgctttt ctgtgactgg tgagtactca 1500
1501 accaagtcat tctgagaata gtgtatgcgg cgaccgagtt gctcttgccc ggcgtcaata 1560
1561 cgggataata ccgcgccaca tagcagaact ttaaaagtgc tcatcattgg aaaacgttct 1620
1621 tcggggcgaa aactctcaag gatcttaccg ctgttgagat ccagttcgat gtaacccact 1680
1681 cgtgcaccca actgatcttc agcatctttt actttcacca gcgtttctgg gtgagcaaaa 1740
1741 acaggaaggc aaaatgccgc aaaaaaggga ataagggcga cacggaaatg ttgaatactc 1800
1801 atactcttcc tttttcaata ttattgaagc atttatcagg gttattgtct catgagcgga 1860
1861 tacatatttg aatgtattta gaaaaataaa caaatagggg ttccgcgcac atttccccga 1920
1921 aaagtgccac ctgacgtcta agaaaccatt attatcatga cattaaccta taaaaatagg 1980
1981 cgtatcacga ggccctttcg tcttcaagaa ttagcttttc aattcaattc atcatttttt 2040
2041 ttttattctt ttttttgatt tcggtttctt tgaaattttt ttgattcggt aatctccgaa 2100
2101 cagaaggaag aacgaaggaa ggagcacaga cttagattgg tatatatacg catatgtagt 2160
2161 gttgaagaaa catgaaattg cccagtattc ttaacccaac tgcacagaac aaaaacatgc 2220
2221 aggaaacgaa gataaatcat gtcgaaagct acatataagg aacgtgctgc tactcatcct 2280
2281 agtcctgttg ctgccaagct atttaatatc atgcacgaaa agcaaacaaa cttgtgtgct 2340
2341 tcattggatg ttcgtaccac caaggaatta ctggagttag ttgaagcatt aggtcccaaa 2400
2401 atttgtttac taaaaacaca tgtggatatc ttgactgatt tttccatgga gggcacagtt 2460
2461 aagccgctaa aggcattatc cgccaagtac aattttttac tcttcgaaga cagaaaattt 2520
2521 gctgacattg gtaatacagt caaattgcag tactctgcgg gtgtatacag aatagcagaa 2580
2581 tgggcagaca ttacgaatgc acacggtgtg gtgggcccag gtattgttag cggtttgaag 2640
2641 caggcggcag aagaagtaac aaaggaacct agaggccttt tgatgttagc agaattgtca 2700
2701 tgcaagggct ccctatctac tggagaatat actaagggta ctgttgacat tgcgaagagc 2760
2761 gacaaagatt ttgttatcgg ctttattgct caaagagaca tgggtggaag agatgaaggt 2820
2821 tacgattggt tgattatgac acccggtgtg ggtttagatg acaagggaga cgcattgggt 2880
2881 caacagtata gaaccgtgga tgatgtggtc tctacaggat ctgacattat tattgttgga 2940
2941 agaggactat ttgcaaaggg aagggatgct aaggtagagg gtgaacgtta cagaaaagca 3000
3001 ggctgggaag catatttgag aagatgcggc cagcaaaact aaaaaactgt attataagta 3060
3061 aatgcatgta tactaaactc acaaattaga gcttcaattt aattatatca gttattaccc 3120
3121 aattctcatg tttgacagct tatcatcgat cgtccaactg catggagatg agtcgtggca 3180
3181 agaataccaa gagttcctcg gtttgccagt tattaaaaga ctcgtatttc caaaagactg 3240
3241 caacatacta ctcagtgcag cttcacagaa acctcattcg tttattccct tgtttgattc 3300
3301 agaagcaggt gggacaggtg aacttttgga ttggaactcg atttctgact gggttggaag 3360
3361 gcaagagagc cccgagagct tacattttat gttagctggt ggactgacgc cagaaaatgt 3420
3421 tggtgatgcg cttagattaa atggcgttat tggtgttgat gtaagcggag gtgtggagac 3480
3481 aaatggtgta aaagactcta acaaaatagc aaatttcgtc aaaaatgcta agaaataggt 3540
3541 tattactgag tagtatttat ttaagtattg tttgtgcact tgcctgcaag cctttttgaaa 3600
```

3601 agcaagcata aaagatctaa acataaaatc tgtaaaataa caagatgtaa agataatgct 3660
3661 aaatcatttg gctttttgat tgattgtaca ggaaaatata catcgcaggg ggttgacttt 3720
3721 taccatttca ccgcaatgga atcaaacttg ttgaagagaa tgttcacagg cgcatacgct 3780
3781 acaatgaccc gattcttgct agccttttct cggtcttgca aacaaccgcc ggcagcttag 3840
3841 tatataaata cacatgtaca tacctctctc cgtatcctcg taatcatttt cttgtattta 3900
3901 tcgtcttttc gctgtaaaaa ctttatcaca cttatctcaa atacacttat taaccgcttt 3960
3961 tactattatc ttctacgctg acagtaatat caaacagtga cacatattaa acacagtggt 4020
4021 ttctttgcat aaacaccatc agcctcaagt cgtcaagtaa agatttcgtg ttcatgcaga 4080
4081 tagataacaa tctatatgtt gataattagc gttgcctcat caatgcgaga tccgtttaac 4140
4141 cggaccctag tgcacttacc ccacgttcgg tccactgtgt gccgaacatg ctccttcact 4200
4201 attttaacat gtggaattaa ttctcatgtt tgacagctta tcatcgaact ctaagaggtg 4260
4261 atacttattt actgtaaaac tgtgacgata aaaccggaag gaagaataag aaaactcgaa 4320
4321 ctgatctata atgcctattt tctgtaaaga gtttaagcta tgaaagcctc ggcattttgg 4380
4381 ccgctcctag gtagtgcttt ttttccaagg acaaaacagt ttctttttct tgagcaggtt 4440
4441 ttatgtttcg gtaatcataa acaataaata aattatttca tttatgttta aaaataaaaa 4500
4501 ataaaaaagt attttaaatt tttaaaaaag ttgattataa gcatgtgacc ttttgcaagc 4560
4561 aattaaattt tgcaatttgt gattttaggc aaaagttaca atttctggct cgtgtaatat 4620
4621 atgtatgcta aagtgaactt ttacaaagtc gatatggact tagtcaaaag aaatttttctt 4680
4681 aaaaatatat agcactagcc aatttagcac ttctttatga gatatattat agactttatt 4740
4741 aagccagatt tgtgtattat atgtatttac ccggcgaatc atggacatac attctgaaat 4800
4801 aggtaatatt ctctatggtg agacagcata gataacctag gatacaagtt aaaagctagt 4860
4861 actgttttgc agtaattttt ttcttttttta taagaatgtt accacctaaa taagttataa 4920
4921 agtcaatagt taagtttgat atttgattgt aaaataccgt aatatatttg catgatcaaa 4980
4981 aggctcaatg ttgactagcc agcatgtcaa ccactatatt gatcaccgat atatggactt 5040
5041 ccacaccaac tagtaatatg acaataaatt caagatattc ttcatgagaa tggcccagcg 5100
5101 atatatgcgg tgtgaaatac cgcacagatg cgtaaggaga aaataccgca tcaggcgcca 5160
5161 ttcgccattc aggctgcgca actgttggga agggcgatcg gtgcgggcct cttcgctatt 5220
5221 acgccagctg gcgaaagggg gatgtgctgc aaggcgatta agttgggtaa cgccagggtt 5280
5281 ttcccagtca cgacgttgta aaacgacggc cagtgaattc gagctcggta cccacggatt 5340
5341 agaagccgcc gagcgggtga cagccctccg aaggaagact ctcctccgtg cgtcctcgtc 5400
5401 ttcaccggtc gcgttcctga aacgcagatg tgcctcgcgc cgcactgctc cgaacaataa 5460
5461 agattctaca atactagctt ttatggttat gaagaggaaa aattggcagt aacctggccc 5520
5521 cacaaacctt caaatgaacg aatcaaatta acaaccatag gatgataatg cgattagttt 5580
5581 tttagcctta tttctggggt aattaatcag cgaagcgatg atttttgatc tattaacaga 5640
5641 tatataaatg caaaaactgc ataaccactt taactaatac tttcaacatt ttcggtttgt 5700
5701 attacttctt attcaaatgt aataaaagta tcaacaaaaa attgttaata tacctctata 5760
5761 ctttaacgtc aaggagaaaa aaccccggat cggactacta gcagctgtaa tacgactcac 5820
5821 tatagggaat attaagctat caaacaagtt tgtacaaaaa agctgaacga gaaacgtaaa 5880
5881 atgatataaa tatcaatata ttaaattaga ttttgcataa aaaacagact acataatact 5940
5941 gtaaaacaca acatatccag tcactatggc ggccgcgggt gatgctgcca acttagcggc 6000
6001 cgcattaggc accccaggct ttacacttta tgcttccggc tcgtataatg tgtggatttt 6060
6061 gagttaggat ccgtcgagat ttcaggagc taaggaagct aaaatggaga aaaaaatcac 6120
6121 tggatatacc accgttgata tatcccaatg gcatcgtaaa gaacattttg aggcatttca 6180
6181 gtcagttgct caatgtacct ataaccagac cgttcagctg gatattacgg cctttttaaa 6240
6241 gaccgtaaag aaaaataagc acaagtttta tccggccttt attcacattc ttgcccgcct 6300
6301 gatgaatgct catccggaat tccgtatggc aatgaaagac ggtgagctgg tgatatggga 6360
6361 tagtgttcac ccttgttaca ccgttttcca tgagcaaact gaaacgtttt catcgctctg 6420
6421 gagtgaatac cacgacgatt tccggcagtt tctacacata tattcgcaag atgtggcgtg 6480
6481 ttacggtgaa aacctggcct atttccctaa agggtttatt gagaatatgt ttttcgtctc 6540
6541 agccaatccc tgggtgagtt tcaccagttt tgatttaaac gtggccaata tggacaactt 6600
6601 cttcgccccc gttttcacca tgggcaaata ttatacgcaa ggcgacaagg tgctgatgcc 6660
6661 gctggcgatt caggttcatc atgccgtttg tgatggcttc catgtcggca gaatgcttaa 6720
6721 tgaattacaa cagtactgcg atgagtggca gggcggggcgt aatctagagg atccggctta 6780
6781 ctaaaagcca gataacagta tgcgtatttg cgcgctgatt tttgcggtat aagaatatat 6840
6841 actgatatgt atacccgaag tatgtcaaaa agaggtatgc tatgaagcag cgtattacag 6900
6901 tgacagttga cagcgacagc tatcagttgc tcaaggcata tatgatgtca atatctccgg 6960
6961 tctggtaagc acaaccatgc agaatgaagc ccgtcgtctg cgtgccgaac gctggaaagc 7020
7021 ggaaaatcag gaagggatgg ctgaggtcgc ccggtttatt gaaatgaacg gctcttttgc 7080
7081 tgacgagaac aggggctggt gaaatgcagt ttaaggttta cacctataaa agagagagcc 7140
7141 gttatcgtct gtttgtggat gtacagagtg atattattga cacgcccggg cgacggatgg 7200
7201 tgatccccct ggccagtgca cgtctgctgt cagataaagt cccccgtgaa ctttacccgg 7260

7261 tggtgcatat cggggatgaa agctggcgca tgatgaccac cgatatggcc agtgtgccgg 7320
7321 tctccgttat cggggaagaa gtggctgatc tcagccaccg cgaaaatgac atcaaaaacg 7380
7381 ccattaacct gatgttctgg ggaatataaa tgtcaggctc cgttatacac agccagtctg 7440
7441 caggtcgacc atagtgactg gatatgttgt gttttacagt attatgtagt ctgtttttta 7500
7501 tgcaaaatct aatttaatat attgatattt atatcatttt acgtttctcg ttcagctttc 7560
7561 ttgtacaaag tggttcgatc tagagggccc ttcgaaggta agcctatccc taaccctctc 7620
7621 ctcggtctcg attctacgcg taccggtcat catcaccatc accattgagt ttaaacccgc 7680
7681 tgatcctaga gggccgcatc atgtaattag ttatgtcacg cttacattca cgccctcccc 7740
7741 ccacatccgc tctaaccgaa aaggaaggag ttagacaacc tgaagtctag gtccctattt 7800
7801 atttttttat agttatgtta gtattaagaa cgttatttat atttcaaatt tttctttttt 7860
7861 ttctgtacag acgcgtgtac gcatgtaaca ttatactgaa aaccttgctt gagaaggttt 7920
7921 tgggacgctc gaaggcttta atttgcaagc tgcggccctg cattaatgaa tcggccaacg 7980
7981 cgcgggggat cctctagagt cgacctgcag gcatgcaagc ttggcgtaat catggtcata 8040
8041 gctgtttcct gtgtgaaatt gttatccgct cacaattcca cacaacatac gagccggaag 8100
8101 cataaagtgt aaagcctggg gtgcctaatg agtgagctaa ctcacattaa ttgcgttgcg 8160
8161 ctcactgccc gctttccagt cgggaaacct gtcgtgccag ctgcattaat gaatcggcca 8220
8221 acgcgcgggg agaggcggtt tgcgtattgg gcgc 8254

Sequence features of pDEST-YcPlac111



Plasmid components :

pGAL1, Full length yeast galactose inducible promoter for high-level protein expression.

T7, promoter for in vitro transcrition.

***att*R1** and ***att*R2**, recombination sites, for LR cloning of the gene of interest from an entry clone to destination vector.

***ccd*B** and **CmR** (Chloramphenicol resistance) genes for negative selection and counterselection, respectively.

V5 epitope and **6xHis** (Polyhistidine) tag for detection and purification (if required).

CYC1pA, Cytochrome c polyadenylation sequence for proper termination and processing of the transcript.

CEN4/ARS1, region for episomal maintenance and low copy replication.

LEU2, auxotrophic marker for selection of yeast transformants.

AmpR, ampicillin resistance gene for selection in *E. coli.*

```
   1 tcttccgctt cctcgctcac tgactcgctg cgctcggtcg ttcggctgcg gcgagcggta 60
  61 tcagctcact caaaggcggt aatacggtta tccacagaat caggggataa cgcaggaaag 120
 121 aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg 180
 181 tttttccata ggctccgccc ccctgacgag catcacaaaa atcgacgctc aagtcagagg 240
 241 tggcgaaacc cgacaggact ataaagatac caggcgtttc cccctggaag ctccctcgtg 300
 301 cgctctcctg ttccgaccct gccgcttacc ggatacctgt ccgcctttct cccttcggga 360
 361 agcgtggcgc tttctcatag ctcacgctgt aggtatctca gttcggtgta ggtcgttcgc 420
 421 tccaagctgg gctgtgtgca cgaacccccc gttcagcccg accgctgcgc cttatccggt 480
 481 aactatcgtc ttgagtccaa cccggtaaga cacgacttat cgccactggc agcagccact 540
 541 ggtaacagga ttagcagagc gaggtatgta ggcggtgcta cagagttctt gaagtggtgg 600
 601 cctaactacg gctacactag aaggacagta tttggtatct gcgctctgct gaagccagtt 660
 661 accttcggaa aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcggt 720
 721 ggtttttttg tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct 780
 781 ttgatctttt ctacggggtc tgacgctcag tggaacgaaa actcacgtta agggattttg 840
 841 gtcatgagat tatcaaaaag gatcttcacc tagatccttt taaattaaaa atgaagtttt 900
 901 aaatcaatct aaagtatata tgagtaaact tggtctgaca gttaccaatg cttaatcagt 960
 961 gaggcaccta tctcagcgat ctgtctattt cgttcatcca tagttgcctg actccccgtc 1020
1021 gtgtagataa ctacgatacg ggagggctta ccatctggcc ccagtgctgc aatgataccg 1080
1081 cgagacccac gctcaccggc tccagattta tcagcaataa accagccagc cggaagggcc 1140
1141 gagcgcagaa gtggtcctgc aactttatcc gcctccatcc agtctattaa ttgttgccgg 1200
1201 gaagctagag taagtagttc gccagttaat agtttgcgca acgttgttgc cattgctaca 1260
1261 ggcatcgtgg tgtcacgctc gtcgtttggt atggcttcat tcagctccgg ttcccaacga 1320
1321 tcaaggcgag ttacatgatc ccccatgttg tgcaaaaaag cggttagctc cttcggtcct 1380
1381 ccgatcgttg tcagaagtaa gttggccgca gtgttatcac tcatggttat ggcagcactg 1440
1441 cataattctc ttactgtcat gccatccgta agatgctttt ctgtgactgg tgagtactca 1500
1501 accaagtcat tctgagaata gtgtatgcgg cgaccgagtt gctcttgccc ggcgtcaata 1560
1561 cgggataata ccgcgccaca tagcagaact ttaaaagtgc tcatcattgg aaaacgttct 1620
1621 tcggggcgaa aactctcaag gatcttaccg ctgttgagat ccagttcgat gtaacccact 1680
1681 cgtgcaccca actgatcttc agcatctttt actttcacca gcgtttctgg gtgagcaaaa 1740
1741 acaggaaggc aaaatgccgc aaaaaaggga ataagggcga cacggaaatg ttgaatactc 1800
1801 atactcttcc tttttcaata ttattgaagc atttatcagg gttattgtct catgagcgga 1860
1861 tacatatttg aatgtattta gaaaaataaa caaatagggg ttccgcgcac atttccccga 1920
1921 aaagtgccac ctgacgtcta agaaaccatt attatcatga cattaaccta taaaaatagg 1980
1981 cgtatcacga ggccctttcg tcttcaagaa ttaactgtgg gaatactcag gtatcgtaag 2040
2041 atgcaagagt tcgaatctct tagcaaccat tatttttttc ctcaacataa cgagaacaca 2100
2101 caggggcgct atcgcacaga atcaaattcg atgactggaa atttttttgtt aatttcagag 2160
2161 gtcgcctgac gcatatacct ttttcaactg aaaaattggg agaaaaagga aaggtgagag 2220
2221 gccggaaccg gcttttcata tagaatagag aagcgttcat gactaaatgc ttgcatcaca 2280
2281 atacttgaag ttgacaatat tatttaagga cctattgttt tttccaatag gtggttagca 2340
2341 atcgtcttac tttctaactt ttcttacctt ttacatttca gcaatatata tatatatttc 2400
2401 aaggatatac cattctaatg tctgccccta tgtctgcccc taagaagatc gtcgttttgc 2460
2461 caggtgacca cgttggtcaa gaaatcacag ccgaagccat taaggttctt aaagctattt 2520
2521 ctgatgttcg ttccaatgtc aagttcgatt tcgaaaatca tttaattggt ggtgctgcta 2580
2581 tcgatgctac aggtgtccca cttccagatg aggcgctgga agcctccaag aaggttgatg 2640
2641 ccgttttgtt aggtgctgtg ggtggtccta aatggggtac aggtagtgtt agacctgaac 2700
2701 aaggtttact aaaaatccgt aaagaacttc aattgtacgc caacttaaga ccatgtaact 2760
2761 ttgcatccga ctctctttta gacttatctc caatcaagcc acaatttgct aaaggtactg 2820
2821 acttcgttgt tgtcagagaa ttagtgggag gtatttactt tggtaagaga aaggaagacg 2880
2881 atggtgatgg tgtcgcttgg gatagtgaac aatacaccgt tccagaagtg caaagaatca 2940
2941 caagaatggc cgctttcatg gccctacaac atgagccacc attgcctatt tggtccttgg 3000
3001 ataaagctaa tgttttggcc tcttcaagat tatggagaaa aactgtggag gaaaccatca 3060
3061 agaacgaatt tcctacattg aaggttcaac atcaattgat tgattctgcc gccatgatcc 3120
3121 tagttaagaa cccaacccac ctaaatggta ttataatcac cagcaacatg tttggtgata 3180
3181 tcatctccga tgaagcctcc gttatcccag gttccttggg tttgttgcca tctgcgtcct 3240
3241 tggcctcttt gccagacaag aacaccgcat ttggtttgta cgaaccatgc cacggttctg 3300
3301 ctccagattt gccaaagaat aaggttgacc ctatcgccac tatcttgtct gctgcaatga 3360
3361 tgttgaaatt gtcattgaac ttgcctgaag aaggtaaggc cattgaagat gcagttaaaa 3420
3421 aggttttgga tgcaggtatc agaactggtg atttaggtgg ttccaacagt accaccgaag 3480
3481 tcggtgatgc tgtcgccgaa gaagttaaga aaatccttgc ttaaaaagat tctctttttt 3540
3541 tatgatattt gtacataaac tttataaatg aaattcataa tagaaacgac acgaaattac 3600
3601 aaaatggaat atgttcatag ggtagaatta attctcatgt ttgacagctt atcatcgatc 3660
```

```
3661 gtccaactgc atggagatga gtcgtggcaa gaataccaag agttcctcgg tttgccagtt 3720
3721 attaaaagac tcgtatttcc aaaagactgc aacatactac tcagtgcagc ttcacagaaa 3780
3781 cctcattcgt ttattccctt gtttgattca gaagcaggtg ggacaggtga acttttggat 3840
3841 tggaactcga tttctgactg ggttggaagg caagagagcc ccgagagctt acattttatg 3900
3901 ttagctggtg gactgacgcc agaaaatgtt ggtgatgcgc ttagattaaa tggcgttatt 3960
3961 ggtgttgatg taagcggagg tgtggagaca aatggtgtaa aagactctaa caaaatagca 4020
4021 aatttcgtca aaaatgctaa gaaataggtt attactgagt agtatttatt taagtattgt 4080
4081 ttgtgcactt gcctgcaagc cttttgaaaa gcaagcataa aagatctaaa cataaaatct 4140
4141 gtaaaataac aagatgtaaa gataatgcta aatcatttgg ctttttgatt gattgtacag 4200
4201 gaaaatatac atcgcagggg gttgactttt accatttcac cgcaatggaa tcaaacttgt 4260
4261 tgaagagaat gttcacaggc gcatacgcta caatgacccg attcttgcta gccttttctc 4320
4321 ggtcttgcaa acaaccgccg gcagcttagt atataaatac acatgtacat acctctctcc 4380
4381 gtatcctcgt aatcattttc ttgtatttat cgtcttttcg ctgtaaaaac tttatcacac 4440
4441 ttatctcaaa tacacttatt aaccgctttt actattatct tctacgctga cagtaatatc 4500
4501 aaacagtgac acatattaaa cacagtggtt tctttgcata aacaccatca gcctcaagtc 4560
4561 gtcaagtaaa gatttcgtgt tcatgcagat agataacaat ctatatgttg ataattagcg 4620
4621 ttgcctcatc aatgcgagat ccgtttaacc ggaccctagt gcacttaccc cacgttcggt 4680
4681 ccactgtgtg ccgaacatgc tccttcacta ttttaacatg tggaattaat tctcatgttt 4740
4741 gacagcttat catcgaactc taagaggtga tacttattta ctgtaaaact gtgacgataa 4800
4801 aaccggaagg aagaataaga aaactcgaac tgatctataa tgcctatttt ctgtaaagag 4860
4861 tttaagctat gaaagcctcg gcattttggc cgctcctagg tagtgctttt tttccaagga 4920
4921 caaaacagtt tctttttctt gagcaggttt tatgtttcgg taatcataaa caataaataa 4980
4981 attatttcat ttatgtttaa aaataaaaaa taaaaaagta ttttaaattt ttaaaaaagt 5040
5041 tgattataag catgtgacct tttgcaagca attaaatttt gcaatttgtg atttaggca 5100
5101 aaagttacaa tttctggctc gtgtaatata tgtatgctaa agtgaacttt tacaaagtcg 5160
5161 atatggactt agtcaaaaga aattttctta aaaatatata gcactagcca atttagcact 5220
5221 tctttatgag atatattata gactttatta agccagattt gtgtattata tgtatttacc 5280
5281 cggcgaatca tggacataca ttctgaaata ggtaatattc tctatggtga gacagcatag 5340
5341 ataacctagg atacaagtta aaagctagta ctgttttgca gtaatttttt tctttttat 5400
5401 aagaatgtta ccacctaaat aagttataaa gtcaatagtt aagtttgata tttgattgta 5460
5461 aaataccgta atatatttgc atgatcaaaa ggctcaatgt tgactagcca gcatgtcaac 5520
5521 cactatattg atcaccgata tatggacttc cacaccaact agtaatatga caataaattc 5580
5581 aagatattct tcatgagaat ggcccagcga tatatgcggt gtgaaatacc gcacagatgc 5640
5641 gtaaggagaa aataccgcat caggcgccat tcgccattca ggctgcgcaa ctgttgggaa 5700
5701 gggcgatcgg tgcgggcctc ttcgctatta cgccagctgg cgaaaggggg atgtgctgca 5760
5761 aggcgattaa gttgggtaac gccagggttt tcccagtcac gacgttgtaa aacgacggcc 5820
5821 agtgaattcg agctcggtac ccacggatta gaagccgccg agcgggtgac agccctccga 5880
5881 aggaagactc tcctccgtgc gtcctcgtct tcaccggtcg cgttcctgaa acgcagatgt 5940
5941 gcctcgcgcc gcactgctcc gaacaataaa gattctacaa tactagcttt tatggttatg 6000
6001 aagaggaaaa attggcagta acctggcccc acaaaccttc aaatgaacga atcaaattaa 6060
6061 caaccatagg atgataatgc gattagtttt ttagccttat ttctggggta attaatcagc 6120
6121 gaagcgatga tttttgatct attaacagat atataaatgc aaaaactgca taaccacttt 6180
6181 aactaatact ttcaacattt tcggtttgta ttacttctta ttcaaatgta ataaaagtat 6240
6241 caacaaaaaa ttgttaatat acctctatac tttaacgtca aggagaaaaa accccggatc 6300
6301 ggactactag cagctgtaat acgactcact atagggaata ttaagctatc aaacaagttt 6360
6361 gtacaaaaaa gctgaacgag aaacgtaaaa tgatataaat atcaatatat taaattagat 6420
6421 tttgcataaa aaacagacta cataatactg taaaacacaa catatccagt cactatggcg 6480
6481 gccgcgggtg atgctgccaa cttagcggcc gcattaggca ccccaggctt tacactttat 6540
6541 gcttccggct cgtataatgt gtggatttg agttaggatc cgtcgagatt ttcaggagct 6600
6601 aaggaagcta aaatggagaa aaaaatcact ggatatacca ccgttgatat atcccaatgg 6660
6661 catcgtaaag aacattttga ggcatttcag tcagttgctc aatgtaccta taaccagacc 6720
6721 gttcagctgg atattacggc ctttttaaag accgtaaaga aaaataagca caagttttat 6780
6781 ccggccttta ttcacattct tgcccgcctg atgaatgctc atccggaatt ccgtatggca 6840
6841 atgaaagacg gtgagctggt gatatgggat agtgttcacc cttgttacac cgttttccat 6900
6901 gagcaaactg aaacgttttc atcgctctgg agtgaatacc acgacgattt ccggcagttt 6960
6961 ctacacatat attcgcaaga tgtggcgtgt tacggtgaaa acctggccta tttccctaaa 7020
7021 gggtttattg agaatatgtt tttcgtctca gccaatccct gggtgagttt caccagtttt 7080
7081 gatttaaacg tggccaatat ggacaacttc ttcgcccccg ttttcaccat gggcaaatat 7140
7141 tatacgcaag gcgacaaggt gctgatgccg ctggcgattc aggttcatca tgccgtttgt 7200
7201 gatggcttcc atgtcggcag aatgcttaat gaattacaac agtactgcga tgagtggcag 7260
7261 gcggggcgta atctagagga tccggcttac taaaagccag ataacagtat gcgtatttgc 7320
```

7321 gcgctgattt ttgcggtata agaatatata ctgatatgta tacccgaagt atgtcaaaaa 7380
7381 gaggtatgct atgaagcagc gtattacagt gacagttgac agcgacagct atcagttgct 7440
7441 caaggcatat atgatgtcaa tatctccggt ctggtaagca caaccatgca gaatgaagcc 7500
7501 cgtcgtctgc gtgccgaacg ctggaaagcg gaaaatcagg aagggatggc tgaggtcgcc 7560
7561 cggtttattg aaatgaacgg ctcttttgct gacgagaaca ggggctggtg aaatgcagtt 7620
7621 taaggtttac acctataaaa gagagagccg ttatcgtctg tttgtggatg tacagagtga 7680
7681 tattattgac acgcccgggc gacggatggt gatccccctg gccagtgcac gtctgctgtc 7740
7741 agataaagtc cccgtgaac tttacccggt ggtgcatatc ggggatgaaa gctggcgcat 7800
7801 gatgaccacc gatatggcca gtgtgccggt ctccgttatc ggggaagaag tggctgatct 7860
7861 cagccaccgc gaaaatgaca tcaaaaacgc cattaacctg atgttctggg gaatataaat 7920
7921 gtcaggctcc gttatacaca gccagtctgc aggtcgacca tagtgactgg atatgttgtg 7980
7981 ttttacagta ttatgtagtc tgtttttat gcaaaatcta atttaatata ttgatattta 8040
8041 tatcatttta cgtttctcgt tcagctttct tgtacaaagt ggttcgatct agagggccct 8100
8101 tcgaaggtaa gcctatccct aaccctctcc tcggtctcga ttctacgcgt accggtcatc 8160
8161 atcaccatca ccattgagtt taaacccgct gatcctagag ggccgcatca tgtaattagt 8220
8221 tatgtcacgc ttacattcac gccctccccc cacatccgct ctaaccgaaa aggaaggagt 8280
8281 tagacaacct gaagtctagg tccctattta ttttttata gttatgttag tattaagaac 8340
8341 gttatttata tttcaaattt ttcttttttt tctgtacaga cgcgtgtacg catgtaacat 8400
8401 tatactgaaa accttgcttg agaaggtttt gggacgctcg aaggctttaa tttgcaagct 8460
8461 gcggccctgc attaatgaat cggccaacgc gcgggggatc ctctagagtc gacctgcagg 8520
8521 catgcaagct tggcgtaatc atggtcatag ctgtttcctg tgtgaaattg ttatccgctc 8580
8581 acaattccac acaacatacg agccggaagc ataaagtgta aagcctgggg tgcctaatga 8640
8641 gtgagctaac tcacattaat tgcgttgcgc tcactgcccg ctttccagtc gggaaacctg 8700
8701 tcgtgccagc tgcattaatg aatcggccaa cgcgcgggga gaggcggttt gcgtattggg 8760
8761 cgc 8763

© for non-published parts Suresh Sajwan

sajsuresh@gmail.com

Gene targeting in Silkworm (*Bombyx mori*) by Engineered Endonucleases
Ph.D. Thesis, 2013

All rights reserved
For non-commercial use only

University of South Bohemia in České Budějovice
Faculty of Science
Branišovská 31
CZ-37005 České Budějovice, Czech Republic

Phone: +420 387 772 244
www.prf.jcu.cz, e-mail: sekret@prf.jcu.cz