



Universiteit
Leiden
The Netherlands

Role of non-homologous end-joining in T-DNA integration in *Arabidopsis thaliana*

Shen, H.

Citation

Shen, H. (2017, January 19). *Role of non-homologous end-joining in T-DNA integration in Arabidopsis thaliana*. Retrieved from <https://hdl.handle.net/1887/45272>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/45272>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/45272> holds various files of this Leiden University dissertation.

Author: Shen, H.

Title: Role of non-homologous end-joining in T-DNA integration in *Arabidopsis thaliana*

Issue Date: 2017-01-19

**Role of non-homologous end-joining in T-DNA
integration in *Arabidopsis thaliana***

Hexi Shen

Hexi Shen

Role of non-homologous end-joining in T-DNA integration in *Arabidopsis thaliana*

PhD thesis, Leiden University, 2017

© Hexi Shen (2017). All rights reserved. No part of this thesis may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, without the prior written permission of the copyright holder.

Cover designed by Hexi Shen

Printed by Ipskamp in the Netherlands

Role of non-homologous end-joining in T-DNA integration in *Arabidopsis thaliana*

Proefschrift

Ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van de Rector Magnificus Prof. mr. C. J. J. M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 19 Januari 2017
klokke 13.45 uur

door

Hexi Shen

Geboren te Fujian, China

3 Augustus 1984

Promotiecommissie

Promotor: Prof. Dr. P. J. J. Hooykaas

Co-promotor: Dr. B. S. de Pater

Overige leden: Dr. M. van Kregten

Dr. M. Knip (University of Amsterdam)

Prof. Dr. H. P. Spaink

Prof. Dr. C. A. M. J. J. van den Hondel

Prof. Dr. J. Memelink

Prof. Dr. M. Tijsterman

To my family and Ruobing

Table of contents

Chapter 1	General introduction	9
Chapter 2	<i>Agrobacterium</i> T-DNA integration in <i>Arabidopsis</i> non-homologous end-joining mutants	35
Chapter 3	Characterization of an <i>Arabidopsis</i> gene encoding a putative DNA ligase	51
Chapter 4	Mre11 and Ku80 control different pathways of DNA repair and T-DNA integration in <i>Arabidopsis</i>	65
Chapter 5	Sequence-specific nuclease-induced double strand break repair in <i>Arabidopsis</i> non-homologous end-joining mutants	83
Summary		109
Samenvatting		115
Curriculum vitae		121

Chapter 1

General introduction

Hexi Shen, Paul J. J. Hooykaas, Sylvia de Pater

Department of Molecular and Developmental Genetics, Institute of Biology, Leiden University,
Leiden, 2333 BE, The Netherlands

Double-strand breaks (DSBs) are considered to be one of the most lethal forms of DNA damage. They can arise as a consequence of normal cellular metabolism, but occur more frequently due to external factors such as ultraviolet (UV) radiation, ionizing radiation and genotoxic chemicals (Khanna and Jackson 2001). For instance, exposure of the chest to X-rays causes 0.0008 DSBs per cell; a body CT, 0.28 DSBs per cell; a tumor PET scan, 0.4 DSBs per cell; 10 hours Airline travel, 0.002 DSBs per cell; 10 days space mission, 0.33 DSBs per cell (Ciccia and Elledge 2010). DSBs are highly toxic if unrepaired, can cause genome rearrangements and even cell death.

To prevent this, cells have evolved complex and highly conserved systems to detect these lesions, signal their presence, trigger various downstream events and finally bring about repair. There are two main pathways for DNA DSBs repair: Homologous Recombination (HR) and Non-Homologous End-Joining (NHEJ). The HR pathway precisely restores the genomic sequence of the broken DNA ends by using the sister chromatid as a template for accurate repair. In contrast, NHEJ promotes direct ligation of the DSB ends without the requirement for sequence homology and may result in small insertions and deletions at the break site. Both pathways are highly conserved throughout eukaryotic evolution but their relative importance may be different depending on the stage of the cell cycle or the cell type. Unicellular eukaryotes such as the yeast *Saccharomyces.cerevisiae* with small genomes mostly rely on HR to repair DSBs, whereas in higher eukaryotes, like mammals and plants with large genomes containing many repeat sequences, the NHEJ pathway is the predominant repair pathway.

In this chapter, we shall review first the DNA DSBs response and its repair. Two main DSBs repair pathways, HR and NHEJ, their regulation and how these pathways affect DNA integration and *Agrobacterium*-mediated T-DNA integration in particular are discussed. This will be followed by a discussion on strategies to induce DSBs artificially at specific sites in the genome and how this can be exploited to affect gene targeting. Finally we will give an outline of the thesis.

DNA damage response

The DNA damage response (DDR) is a signal transduction pathway that affects many aspects of cellular physiology (DNA repair, cell-cycle arrest, apoptosis and senescence) (Harper and Elledge 2007) (**Figure 1**). Master regulators of this pathway are members of the phosphatidylinositol 3-kinase-like protein kinases (PIKKs) family (Block *et al.* 2004). They control downstream amplification of DNA damage signals by recruitment and phosphorylation of their substrates. The current understanding of the DDR mechanism in mammalian cells is mostly based on the study of the two most important members of the PIKKs family: ATM (Ataxia Telangiectasia Mutated) and ATR (ATM and Rad-3 related) (Falck *et al.* 2005; Shiloh 2006; Matsuoka *et al.* 2007). DSBs are recognized by the MRE11-RAD50-NBS1 (MRN) complex, which then triggers the activation of ATM to activate and/or induce the levels of DNA repair proteins to bring about repair. The appearance of areas in the genome with ssDNA as a consequence of DNA damage or repair leads to the recruitment and activation of the other master regulator ATR. After their recruitment to sites of damage, ATM

and ATR then phosphorylate a number of substrates, including the downstream protein kinases CHK1 and CHK2, which generate a downstream amplification by protein activation and repression, leading to cell cycle arrest and DNA repair. Local phosphorylation of histone H2AX at damage sites leads to local accumulation of repair proteins, which is enhanced by ubiquitination and poly (ADP) ribosylation of specific damage response pathway components.

In contrast to mammalian cells, little is known about the DNA damage response in plants. So far, only few homologous genes in plants have been characterized related to the DNA damage response. The Arabidopsis orthologs of the mammalian ATM and ATR protein play an essential role in the response to DNA damage (Garcia *et al.* 2003; Culligan *et al.* 2006). WEE1 is a critical downstream target of the ATR-ATM signaling cascades (De Schutter *et al.* 2007). Despite the conserved nature of the ATM and ATR kinases, plants and animals seem to respond distinctively different to DNA stress. Many genes needed for cell cycle and DNA damage checkpoint in animals have no ortholog in plants, such as CHK1, CHK2 and p53 (Cools and De Veylder 2009). Instead plants use a downstream factor called SOG1 (suppressor of gamma response 1) as a central hub in the DNA damage response (Yoshiyama *et al.* 2013).

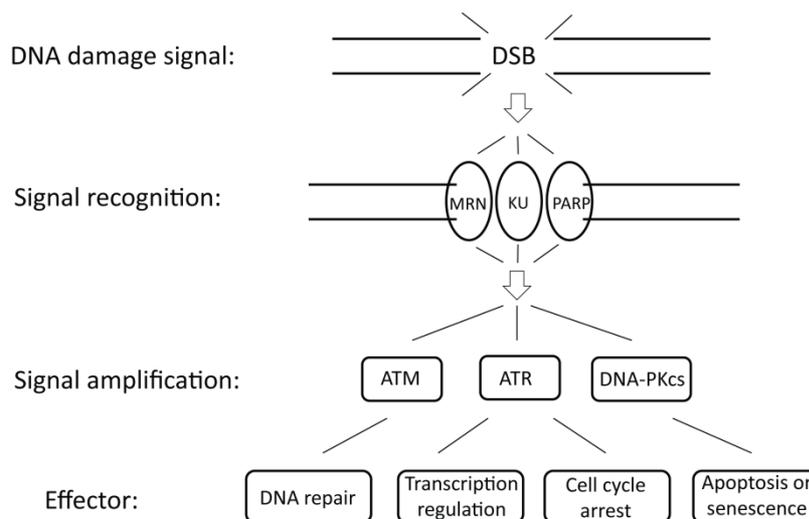


Figure 1. The DNA damage response. After signal (DSB) recognition, signal amplification occurs and effectors are produced. The presence of DSBs in the DNA is recognized by various sensor proteins, such as KU, MRN and PARP. These sensors initiate signaling pathways and activate the PIKKs: DNA-PKcs, ATM and ATR. The PIKKs amplify signals by phosphorylation of a number of downstream substrates, which ultimately leads to cell cycle arrest, DNA repair, activation of transcriptional programmes or induction of programmed cell death.

Homologous recombination

Homologous recombination (HR) promotes genome stability by facilitating error-free repair of DSBs, interstrand crosslinks (ICLs), and DNA gaps during and after DNA replication (Heyer *et al.* 2010). HR is a key repair pathway in the S and G2 phases of the cell cycle. Repair of DSBs via HR requires homologous sequences that act as a template for the reaction. In this

process, the sister chromatid is the preferred template (Kadyk and Hartwell 1992). The information from the homologous sister chromatid is copied into and replaces the damaged region resulting in precise repair. Homology search and DNA strand invasion are key steps in this process. Both are catalyzed by DNA-dependent ATPase Rad51, which is the eukaryotic RecA homolog. It can bind cooperatively to ssDNA, forming helical nucleoprotein filaments (Conway *et al.* 2004). HR can be conceptually divided into three stages and Rad51 functions in all these phases: pre-synapsis, synapsis, and post-synapsis (**Figure 2**).

In the pre-synapsis phase, the DSB is processed to generate ssDNA overhangs by 5' to 3' DNA end resection. End resection in yeast occurs by a two-step mechanism: initial resection of 50–200 nt involves the MRX (MRE11-RAD50-XRS2) complex and Sae2 (Mimitou and Symington 2008), and more extensive resection of up to several kilobases involves either Exo1 or Sgs1 in combination with Dna2 (Mimitou and Symington 2011). Functional homologs of these proteins are found in higher eukaryotes including plants as well (Mimitou and Symington 2009; Osman *et al.* 2011). After end resection, ssDNA is bound by replication protein A (RPA) to protect the ends from degradation and formation of secondary structures, which is needed for competent Rad51 filaments to assemble (Owalczykowski 1998). However, RPA bound to ssDNA also acts as a barrier against Rad51 filament assembly (Sugiyama *et al.* 1997). This inhibitory effect of RPA on Rad51 filament formation can be overcome by at least three different kinds of mediator proteins (Heyer *et al.* 2010). Rad51 paralogs and Rad52 have been identified as the key mediators of Rad51 filament formation in yeast (Sugawara *et al.* 2003). Rad52 interacts with Rad51 as well as with RPA. These interactions are required to recruit Rad51 and also accelerate displacement of RPA from ssDNA by Rad51 (Sugiyama and Kowalczykowski 2002). Besides its function as mediator, Rad52 also contributes to the later stage of recombination (Nimonkar *et al.* 2009). The third class of mediator protein is BRCA2, the human breast and ovarian cancer tumor suppressor protein, which is present in mammalian and plant cells, but not found in budding yeast (Holloman 2011; Seeliger *et al.* 2012).

The Rad51 filament facilitates fast and efficient homology search and DNA strand invasion, leading to the generation of heteroduplex DNA (D-loop). Rad54 is required for searching homology, stimulates DNA strand invasion by the Rad51 filament and also functions after synapsis (Mazin *et al.* 2003; Wolner and Peterson 2005). DNA synthesis is primed by the 3' end of the invading strand, using the donor strand as template.

After extension of the 3' invading strand, repair is finalized by one of at least three different sub-pathways of HR: classical double-strand break repair (DSBR) (Szostak *et al.* 1983), synthesis-dependent strand annealing (SDSA) (Nassif *et al.* 1994), or break-induced replication (BIR) (Kraus *et al.* 2001). In the classical DSBR pathway, Rad52 and Rad59 promote capture of the second end by the D-loop, whereafter two Holliday junctions (dHJ) are formed (Wu *et al.* 2006). These dHJs are either dissolved into non-crossover products by a RecQ helicase such as yeast Sgs1 or human BLM helicase or resolved by a structure-specific endonuclease into crossover/non-crossover products. In general, non-crossovers are much more predominant in mitotic HR (Baudat and de Massy 2007). Crossovers have the potential to generate genomic rearrangements and large-scale loss of heterozygosity (LOH) (Moynahan and Jasin 2010). To avoid crossovers events in mitotic cells, the Mph1 helicase is able to

suppress the DSB repair pathway by dissociating D-loops to facilitate annealing of the extended 3' tail to complementary sequences at the other side of the DSB, leading to the SDSA sub-pathway (Prakash *et al.* 2009). It seems that the Rad51 protein has some inhibitory activity that counters capture of the second end and dHJ formation, indicating an inherent mechanistic bias toward SDSA (Wu *et al.* 2008). In the absence of a second end, the D-loop may become a replication fork, leading to the BIR sub-pathway. During BIR, the replication fork restores the integrity of a broken chromosome by copying the entire distal arm of the template chromosome, resulting in LOH (Lydeard *et al.* 2007). BIR is elevated in *rad51* and *mre11* mutants (Krishna *et al.* 2007).

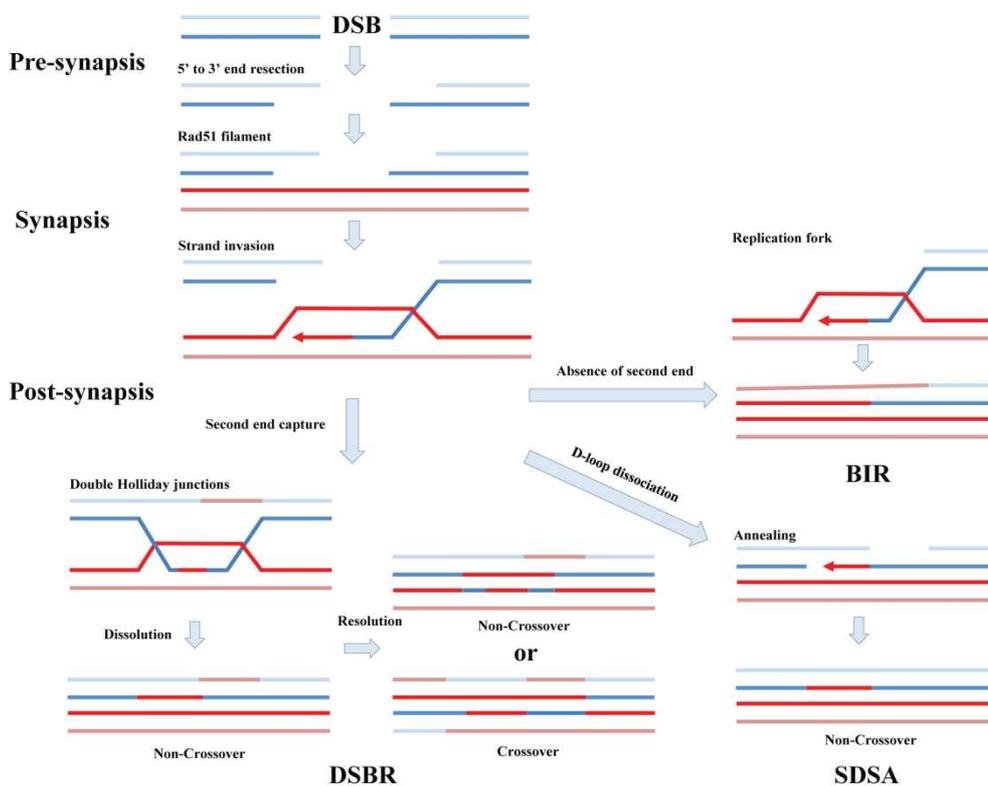


Figure 2. DSB repair by HR. DSBs can be repaired by distinctive HR pathways, including double-strand break repair (DSBR), synthesis-dependent strand annealing (SDSA) and break-induced replication (BIR). First, the DNA damage site is processed by 5' to 3' DNA end resection to generate ssDNA overhangs, which is indispensable for Rad51 filament formation. The Rad51 filament facilitates fast and efficient homology search and DNA strand invasion, leading to the generation of a D-loop structure. In the DSBR pathway, the second end is captured and a double Holliday Junction is formed. The dHJs are either dissolved into non-crossover products by BLM/Sgs1 helicase or resolved by a structure-specific endonuclease into crossover/non-crossover products. BIR happens when a second end is absent. The D-loop intermediate turns into a replication fork, leading to strand synthesis, which replicates the entire homologous template arm.

Table 1. HR proteins in human, yeast and Arabidopsis.

<i>Homo sapiens</i>	<i>Saccharomyces cerevisiae</i>	<i>Arabidopsis thaliana</i>	<i>Arabidopsis</i> gene number	Function
Rad51	Rad51	Rad51	At5g20850	RecA homologue, Strand invasion
MRN complex: Mre11-Rad50-Nbs1	MRX complex: Mre11-Rad50-Xrs2	MRN complex: Mre11-Rad50-Nbs1	At5g54260- At2g31970- At3g02680	DNA binding, Nuclease activities, DSB ends processing, DNA-damage checkpoints
CtIP	Sae2	Com1	At3g52115	DSB ends processing, DNA strand transition
Exo1	Exo1	Exo1A, Exo1B	At1g29630, At1g18090	DSB ends processing
BLM	Sgs1	RecQ4A	At1g10930	DSB ends processing, RecQ helicases
RPA1, RPA2, RPA3	RPA1, RPA2, RPA3	RPA1, RPA2, RPA3	At2g06510, At4g19130, At5g45400, At5g08020, At5g61000; At2g24490, At3g02920; At3g52630, At4g18590	ssDNA binding
Rad51B- Rad51C Rad51C- XRCC3 Rad51D- XRCC2	Rad55-Rad57	Rad51B-Rad51C Rad51C-XRCC3 Rad51D-XRCC2	At2g28560, At2g45280, At5g57450, At1g07745, At5g64520	ssDNA binding, Recombination mediator
Rad52	Rad52	Rad52	At1g71310, At5g47870	ssDNA binding and annealing, Recombination mediator, Interacts with Rad51 and RPA
BRCA1	-	BRCA1	At4g21070	Checkpoint mediator, Recombination mediator
BRCA2	-	BRCA2-1, BRCA2-2	At5g01630, At4g00020	Recombination mediator
Rad54	Rad54	Rad54	At3g19210	ATP-dependent dsDNA translocase, Stimulates the D-loop reaction
FancM	Mph1	FancM	At1g35530	Helicase activity, Dissociates D-loop formation and facilitates single strand annealing
PARI	Srs2	Srs2	At4g25120	Helicase activity, Dissociates D-loop formation and facilitates single strand annealing

Non-homologous end-joining

Non-homologous end-joining (NHEJ) is believed to be the major pathway for the repair of DSBs in most higher eukaryote cells. The basic mechanism of the NHEJ process looks relatively simple and straightforward: it rejoins broken ends directly, without requirement for long runs of end-resection and searching of a homologous repair template. Still many factors are required for NHEJ which demand precise cooperation and timely regulation. Although an increasing number of proteins involved in NHEJ have been identified and their interactions investigated in the past two decades, there is still much to learn about NHEJ.

In outline, the NHEJ pathway can be divided in three distinct steps (**Figure 3**). In the first step, the broken ends of DNA are bound by the Ku70/80 heterodimer to initiate the NHEJ pathway. In the second step, more NHEJ factors are recruited to perform end processing, including Artemis, polynucleotide kinase/phosphatase (PNKP), gap-filling DNA polymerases mu (Pol μ) and lambda (Pol λ), and the Mre11/Rad50/Nbs1 (MRN) complex. In the last step, the ligase4/XRCC4/XLF complex catalyzes the ligation of the processed DNA ends (Mahaney *et al.* 2009). In general, DSB repair by NHEJ can be precise, but may also cause small deletions and insertions of nucleotides at the junction, which alters the nucleotide sequence information surrounding the repair region (van Gent and van der Burg 2007). As a result, NHEJ is referred to as an error-prone DNA repair pathway. More recently, backup NHEJ pathways (b-NHEJ) for DNA repair, also called microhomology-mediated end-joining (MMEJ) or alternative-NHEJ (a-NHEJ), were identified from investigations of cells deficient for NHEJ components (**Figure 3**).

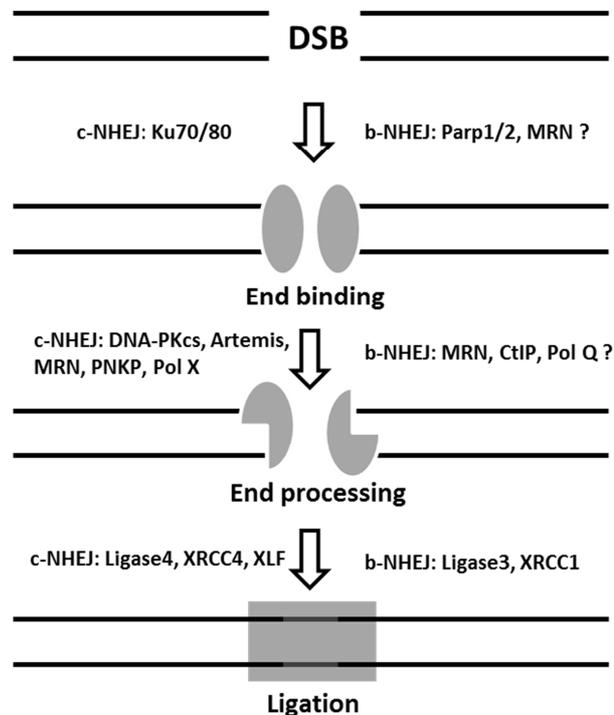


Figure 3. DSB repair by NHEJ. NHEJ comprises three distinct steps: end binding, end processing and ligation. Important factors involved in classical NHEJ and backup NHEJ, which have been identified in mammalian cells, are shown.

Classical non-homologous end-joining

Ku

The classical non-homologous end-joining (c-NHEJ) is initiated by the recognition and binding of DSBs by the Ku heterodimer (Mari *et al.* 2006; Hammel *et al.* 2010) (Figure 3). The Ku heterodimer is comprised of Ku70 and Ku80 subunits, both containing three domains: the N-terminal domain (von Willebrand domain, vWA), the Ku core domain, and the C-terminal domain. Both the N-terminus and core domains contribute to the heterodimerization of the complex. The C-terminal domain of Ku70 contains a SAF-A/B, Acinus and PIAS (SAP) domain contributing to DNA binding activity, while the Ku80 C-terminus is involved in protein-protein interactions (Zhang *et al.* 2001, 2004). The crystal structure of the human Ku heterodimer showed that the Ku heterodimer forms a ring structure on DNA broken ends in a sequence independent manner (Walker *et al.* 2001). Ku is an abundant protein that has an extraordinary affinity for dsDNA ends allowing it to quickly localize to DSBs (Mari *et al.* 2006). Once bound to a DSB, the Ku heterodimer serves as a scaffold to recruit other c-NHEJ factors to the broken end and promotes end-joining. The Ku heterodimer has been shown to interact with DNA-PKcs (Uematsu *et al.* 2007), XRCC4 (Nick McElhinny *et al.* 2000; Costantini *et al.* 2007), DNA ligase 4 (Nick McElhinny *et al.* 2000; Costantini *et al.* 2007), and XLF (Yano *et al.* 2008), directly (see below). Ku70 and Ku80 have also been identified in Arabidopsis (Bundock *et al.* 2002; Tamura *et al.* 2002; West *et al.* 2002). *Atku* mutants have been shown to display hypersensitivity to DSB-inducing treatments. The absence of the Ku protein reduced the efficiency of NHEJ in an *in vivo* plasmid-based end-joining assay in Arabidopsis (Gallego *et al.* 2003; Jia *et al.* 2012), indicating that the Ku protein plays an important role in NHEJ in Arabidopsis as well.

DNA-PKcs

Once Ku binds to a DSB, in mammalian cells DNA-PKcs is recruited to the DNA break very quickly. DNA-PKcs belongs to the phosphatidylinositol 3-kinase-like (PIKK) family of serine/threonine protein kinases and their kinase activity is essential for c-NHEJ in mammalian system. However, fungi and plants seem to lack DNA-PKcs. DNA-PKcs phosphorylates multiple c-NHEJ proteins *in vitro*, including Ku, Artemis, XRCC4, ligase 4, and XLF. It also shows *in vitro* autophosphorylation on multiple sites (Meek *et al.* 2008; Mahaney *et al.* 2009). DNA-PKcs can also be phosphorylated by other members of the PIKK family. The phosphorylation property is important for regulating appropriate and timely access of NHEJ proteins to the DNA ends (Neal and Meek 2011).

Artemis

After the DNA break ends have been detected and captured as described above, the next step is end processing to remove non-ligatable ends. Therefore, c-NHEJ requires additional factors to produce compatible DNA termini for ligation. These factors are well characterized in mammalian systems, but poorly understood in plants. Artemis possesses 5' to 3' exonuclease activity and acquires endonuclease activity when interacting with DNA-PKcs to form phosphorylated complexes, which are responsible for DNA hairpin opening (Ma *et al.* 2002, 2005). Lack of Artemis results in accumulation of unopened DNA hairpins at DNA ends. In

plants, its homolog called sensitive to nitrogen mustard 1 (SNM1) has been reported to function in the repair of oxidative stress-induced DNA damage and positively regulates the level of NHEJ by its exonuclease activity (Molinier *et al.* 2004; Johnson *et al.* 2011).

MRN

The Mre11-Rad50-NBS1 (MRN) complex plays not only a crucial role as a sensor in the DNA damage response, the MRN complex can also bridge the DNA ends and perform end processing in collaboration with CtIP/Sae2, whereby nicking is followed by 3' to 5' exonuclease activity in NHEJ (Paull and Gellert 1998; Chen *et al.* 2001; Williams *et al.* 2008). In plants, MRN mutants display hypersensitivity to DNA damaging agents (Gallego *et al.* 2001; Bundock and Hooykaas 2002; Waterworth *et al.* 2007), but its roles in end processing remain to be determined.

PNKP

Mammalian polynucleotide kinase 3' phosphatase (PNKP) is a bifunctional end processing enzyme. PNKP possesses both 3' DNA phosphatase and 5' DNA kinase activities which are ideally suited for processing of non-ligatable DNA termini, especially 3' overhanging and blunt ended termini, and replaces them with ligatable 5'-phosphates and 3'-OH (Karimi-Busheri *et al.* 2007; Weinfeld *et al.* 2011). PNKP is able to interact with scaffolding XRCC4 as well as poly (ADP-ribose) (PAR) and those interactions facilitate the recruitment of PNKP to the DSB sites (Koch *et al.* 2004; Li *et al.* 2013). Its Arabidopsis homolog called zinc finger DNA 3' phosphoesterase (ZDP), is a modular protein with a C-terminal 3'-phosphatase domain and an N-terminal DNA-binding domain and has been shown to bind SSBs and DSBs and to dephosphorylate 3'-phosphate ends to generate 3'-OH termini (Petrucco *et al.* 2002; Martínez-Macías *et al.* 2012).

Pol X family

The X-family of polymerases including Pol μ , Pol λ and lymphoid-specific terminal transferase (TdT), are involved in adding nucleotides or filling of gaps at DNA ends to accommodate NHEJ. These three DNA polymerases can be recruited to the DNA termini by Ku, but require the XRCC4-Lig4 complex to perform gap filling (Mahajan *et al.* 2002). As the only member of the family X DNA polymerase in plants, Pol λ was suggested to function in various repair pathways in plants and play a key role in DNA repair in the plant genome (Uchiyama *et al.* 2004; Amoroso *et al.* 2011; Roy *et al.* 2011, 2013).

Ligase4, XRCC4, and XLF

The final step in the repair of a DSB via c-NHEJ is ligation of the broken ends, which is carried out by the ligase 4-XRCC4 complex (Grawunder *et al.* 1997). DNA ligase 4 from Arabidopsis has ATP-dependent ligase activity and interacts with XRCC4 to form the Ligase 4-XRCC4 complex (West *et al.* 2000). This interaction is required for stabilization and stimulation of Ligase 4 activity. The *Atlig4* mutant is hypersensitive to the DNA-damaging agents MMS and X-rays, demonstrating that AtLIG4 is required for the repair of DNA damage (van Attikum *et al.* 2003). In animals, XLF (also called Cernunnos) has a similar structure as XRCC4, interacts with the XRCC4-Ligase4 complex and promotes NHEJ by regulating the activity of this complex (Ahnesorg *et al.* 2006).

Backup non-homologous end-joining

In the absence of the Ku complex higher eukaryotic cells still can accomplish end-joining by back up pathways. Several factors that are involved in b-NHEJ have already been identified: PARP1/2, MRN, CtIP, Ligase3, XRCC1, Pol θ (Figure 3). The molecular mechanism of b-NHEJ initiation is poorly understood, although both PARP1 and the MRN complex appear to play important roles (Deriano and Roth 2013). PARP1 has been well described as an important player of the BER/SSBR pathway responsible for the recruitment of the XRCC1-Ligase3 complex to promote repair (Caldecott 2008). Similarly, the XRCC1-Ligase3 complex also seems to contribute to the DSB ligation in the b-NHEJ pathway in mammalian cells (Audebert *et al.* 2004; Simsek *et al.* 2011). PARP1 has been reported to compete for free DNA ends with Ku and to interact with ATM which is one of the key players in the DNA damage response (Aguilar-quesada *et al.* 2007; Haince *et al.* 2008). Therefore, PARP1 may contribute to the early damage response and serve as a platform to recruit other factors to the broken ends in the b-NHEJ pathway. Recently, work in our group confirmed that the *Arabidopsis* homologs AtPARP1 and AtPARP2 are also involved in MMEJ (Jia *et al.* 2013).

Ku inhibits end resection, and in absence of Ku, the MRN complex and the MRN-interacting C-terminal-binding protein interacting protein (CtIP) probably work together to mediate DSB resection in b-NHEJ. The knockdown of Mre11 by siRNA decreased the frequency of b-NHEJ significantly, but it had no effect on the efficiency of c-NHEJ, suggesting that Mre11 is required specifically for b-NHEJ repair of DNA DSBs (Zhuang *et al.* 2009). The knockout of CtIP also results in a significant reduction in b-NHEJ (Lee-Theilen *et al.* 2011; Zhang and Jasin 2011). Human Pol λ and Pol β assist MMEJ using terminal microhomology regions (Crespan *et al.* 2012) and the interaction between PNKP and XRCC1 (Weinfeld *et al.* 2011), suggests that they may also function in end processing during b-NHEJ. Pol θ , which belongs to the DNA polymerase A-family, was reported to be involved in b-NHEJ in *Drosophila* and mammalian cells (Chan *et al.* 2010; Mateos-Gomez *et al.* 2015). It has been shown that the polymerase domain of Pol θ is able to independently carry out all the major stages of MMEJ in vitro (Kent *et al.* 2015).

The ligation step in the b-NHEJ pathway is catalyzed by Lig3 in mammalian cells. Since plants are lacking a Lig3 homolog, there must be other factors to take over ligation during b-NHEJ in plants. One of the candidates is SSB repair factor Ligase 1, which has been implicated in the b-NHEJ repair pathway in *Arabidopsis* (Waterworth *et al.* 2009). In mammals, Ligase 1 displays functional redundancy with Ligase 3 and might cooperate in b-NHEJ (Arakawa *et al.* 2012; Paul *et al.* 2013).

Table 2. NHEJ proteins in human, yeast and Arabidopsis.

<i>Homo sapiens</i>	<i>Saccharomyces cerevisiae</i>	<i>Arabidopsis thaliana</i>	<i>Arabidopsis</i> gene number	Function
Ku70	Ku70	Ku70	At1g16970	DSB end binding and protection
Ku80	Ku80	Ku80	At1g48050	DSB end binding and protection
DNA-PKcs	-	-		protein kinase
Artemis	Snm1/PSO2	Snm1	At3g26680	DNA end processing
MRN complex: Mre11-Rad50-Nbs1	MRX complex: Mre11-Rad50-Xrs2	MRN complex: Mre11-Rad50-Nbs1	At5g54260- At2g31970- At3g02680	DNA binding, Nuclease activities, DSB ends processing, DNA-damage checkpoints
PNKP	Tpp1	ZDP	At3g14890	DNA end processing
Pol λ	-	Pol λ	At1g10520	DNA polymerase, DNA end processing
53BP1	Rad9	Rad9	At3G05480	DNA end processing
DNA ligase IV	Dnl4	lig4	At5g57160	ATP-dependent DNA ligase
XRCC4	Lif1	XRCC4	At3g23100	complex with lig4
XLJ/Cernunnos	Nej1	-		complex with lig4/XRCC4
Parp1	-	Parp1	At2g31320	DNA end binding, NAD ⁺ ADP-ribosyltransferase
Parp2	-	Parp2	At4g02390	DNA end binding, NAD ⁺ ADP-ribosyltransferase
Parp3	-	Parp3	At5g22470	DNA end binding, NAD ⁺ ADP-ribosyltransferase
CtIP	Sae2	Com1	At3g52115	DNA end processing
DNA ligase III	-	-		ATP-dependent DNA ligase
XRCC1	-	XRCC1	At1g80420	complex with lig3
Pol Q	-	Pol θ (Tebichi)	At4g32700	DNA polymerase, DNA end processing

Regulation of DSBs repair pathways choice

As discussed above, HR and NHEJ are two critical pathways to repair DNA DSBs and preserve genome integrity. But when DSBs arise, how is the decision made to choose between homologous recombination or NHEJ for repair? Recent studies show that the regulation of pathway choice depends on the structure of the DNA ends and the cell cycle phase.

The initiation of end resection is an important determinant for repair pathway choice (Symington and Gautier 2011). End resection is initiated by the MRN complex combined with CtIP, which counteracts NHEJ to promote HR by limited end processing. Absence of MRN-CtIP results in a complete block of homology-directed repair (Zhu *et al.* 2008; Mimitou and Symington 2008). Loss of Ku enhanced end resection suggesting that it plays an important role in regulating the balance between NHEJ and HR. In the second step of end resection, Sgs1/BLM, Dna2 and Exo1 contribute to promote HR via generating long ssDNA

ends which are poor substrates for binding by Ku (Symington and Gautier 2011). 53BP1 and BRCA1 are two factors that are enriched at DSB sites and are supposed to be regulators in the choice between NHEJ and HR, respectively (Escribano-Díaz *et al.* 2013). 53BP1 has been implicated in limiting access of nucleases to DNA ends and contributes to promote c-NHEJ. BRCA1 may mediate HR by negatively regulating 53BP1 repair activity as well as concentrating factors required for processing and the subsequent HR resection.

Another important factor affecting the choice between the HR and NHEJ pathways is the phase of the cell cycle (Chapman *et al.* 2012). HR is mainly restricted to the S and G2 phases when the sister chromatid becomes available as template, whereas NHEJ operates throughout the cell cycle but more predominantly in the G1 phase. The resection of DSBs is prevented in G1. The activity of cyclin-dependent kinases (CDKs) is thought to be the crucial factor in both yeast and human cells to promote efficient end-resection (Ferretti *et al.* 2013). It has been shown that CDKs promote the initiation of resection via phosphorylation of Sae2/CtIP and Nbs1 and also regulate long-range resection via phosphorylation of Exo1 (Buis *et al.* 2012; Falck *et al.* 2012; Wang *et al.* 2013; Tomimatsu *et al.* 2014).

***Agrobacterium*-mediated T-DNA integration and NHEJ pathways**

Agrobacterium tumefaciens mediated genetic transformation involves the transfer of T-DNA from its tumor-inducing plasmid to the host cell nuclear genome. *Agrobacterium* is nowadays used as a vector to produce genetically modified plants. T-DNA is at random positions in the plant genome (Gelvin 2000), which may lead to mutation and position effects altering the expression of the transgenes. Therefore, there is great interest to develop methods for the controlled and targeted integration of T-DNA. In yeast this can be accomplished by providing a segment of yeast homologous DNA in the T-DNA. The HR machinery of yeast then mediates integration at the homologous site (Bundock *et al.* 1995). In plants, homologous recombination can occur between a chromosomal locus and a homologous T-DNA introduced via *Agrobacterium* (Offringa *et al.* 1990), but only with a very low efficiency. Two possible models for T-DNA integration have been proposed (Tzfira *et al.* 2004; Gelvin 2010). In the strand-invasion model, T-DNA integration depends on the microhomology between T-DNA and plant DNA sequences. It was suggested that single stranded T-DNA is preferred for integration (Rodenburg *et al.* 1989; Gheysen *et al.* 1991; Mayerhofer *et al.* 1991). In the DNA DSB repair model, the single stranded T-DNA is first converted to a double stranded T-DNA, whereafter this double strand form integrates into the genome at double strand break sites (Salomon and Puchta 1998). This was supported by the fact that DSBs are preferential targets for T-DNA integration and that T-DNA can be cut by a restriction enzyme before integration (Chilton and Que 2003; Tzfira *et al.* 2003).

Previously, it was shown in our lab that *Agrobacterium* T-DNA integration in yeast (*Saccharomyces cerevisiae*) depends on NHEJ proteins, such as Ku70 and DNA ligase 4 (van Attikum *et al.* 2001; van Attikum and Hooykaas 2003). Arabidopsis NHEJ mutants have subsequently been studied for T-DNA integration in plants. However, the results obtained in plant T-DNA integration by different research groups are inconsistent and reveal either no or limited negative effects (Friesner and Britt 2003; van Attikum *et al.* 2003; Gallego *et al.* 2003;

Li *et al.* 2005). Disruption of multiple DNA repair pathways at the same time did not eliminate transformation (Jia *et al.* 2012; Mestiri *et al.* 2014; Park *et al.* 2015), suggesting that there must be other unknown proteins and pathways that mediated T-DNA integration in plants.

Homologous recombination pathway and gene targeting

Gene targeting (GT) is a genetic technique based on homologous recombination to change or replace endogenous genes. It is a powerful technique and has been widely used to study gene function. Several approaches have been developed to select and detect GT events, including gene specific selection (GSS) and positive-negative selection (PNS). In GSS schemes, an endogenous target gene is replaced by a copy of the same gene with a selectable mutation. In PNS schemes, selection of homologous recombination relies on positive and negative selectable markers placed within and outside the homologous sequence, respectively. The PNS-based approach has proven to be very successful and is also used in plant species such as rice and *Arabidopsis*. GT can be achieved efficiently especially in yeast and a few other organisms. However, GT usually is inefficient in cells of multicellular eukaryotes, especially in those of plants and animals due to a much lower efficiency of HR than NHEJ. In order to establish a feasible tool for GT, two options have been tested to improve the frequency of GT events based on the mechanism of HR. The first option was to facilitate the HR pathway by either increasing the synthesis of proteins involved in HR (Reiss *et al.* 2000; Shaked *et al.* 2005) or by inhibiting synthesis of proteins involved in the NHEJ pathway (Jia *et al.* 2012). This was successful in fungi, where prevention of NHEJ by deletion of Ku or Lig4 resulted in very efficient GT (van Attikum and Hooykaas 2003; Kooistra *et al.* 2004; de Boer *et al.* 2010). Unfortunately, in plants hardly any improvement of GT efficiency was seen. Another approach was to enhance GT by inducing genomic DSBs at the target site, which became possible by the development of different classes of artificial nucleases (see below). In this way, the frequency of GT was increased significantly in different organisms including plants (Durai *et al.* 2005; de Pater *et al.* 2009; Urnov *et al.* 2010). GT was also achieved in *Arabidopsis thaliana* by expression of a site-specific endonuclease that not only cuts within the target but also the chromosomal transgenic donor (*in planta* GT), leading to an excised targeting vector in each plant cell (Fauser *et al.* 2012). Due to the recent developments, especially the application of genome editing tools, we expect that a method of efficient GT will be developed for plants in the near future.

Strategies for DSB induction

As DNA recombination events including transgene integration and gene targeting, is increased at break sites in the genome, it has been a strategy to induce local DNA breaks to stimulate these events. Ionizing radiation (X-ray) and genotoxic chemicals (Bleomycin, MMS, etc.) were initially used to induce such DNA breaks, but as they affect the genome in an uncontrolled manner and cause mutation this was not very successful.

Site-specific nucleases have been developed by which DSBs can be induced at a preferred site in the genome. The advent of meganucleases such as SceI for the first time

offered the possibility to induce a DSB at a specific site(s) in a large genome. Such local break led to a significant increase in DNA integration (Salomon and Puchta 1998) and in gene targeting (Smih *et al.* 1995). Since then three classes of nucleases have been used extensively: Zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and CRISPR/Cas (for “clustered regularly interspersed short palindromic repeats/CRISPR associated”) system.

ZFNs consist of zinc finger arrays fused to the nuclease domain of the type II restriction enzyme Fok I. Each zinc finger typically recognizes three nucleotides and engineered fingers have been combined to recognize specific longer DNA sequences. ZFNs function as a dimer to create a DSB within the spacer between the binding sites of two ZFN monomers. ZFN-mediated gene modification has been reported in different eukaryotic organisms (Urnov *et al.* 2010) and also in *Arabidopsis* by others and our group (de Pater *et al.* 2009, 2013; Qi *et al.* 2013). Like ZFNs, TALENs are composed of DNA-binding domains linked to the Fok I nuclease domain. Each binding domain contains a variable number of amino acid repeats, each with a RVD (repeat variable di-residues) present at amino acid positions 12 and 13 of each repeat, which are able to specifically recognize a single base pair of DNA (Joung and Sander 2013). TALENs and ZFNs make DSBs with 5' overhangs. TALENs are considered to be more efficient, specific and reproducible, because TALENs are less affected by context of targeting sequences than ZFNs, as shown in yeast (Christian *et al.* 2010), human (Zhang *et al.* 2011) and *Arabidopsis* (Cermak *et al.* 2011).

More recently, a RNA-guided CRISPR/Cas nuclease system was described for inducing DNA DSBs at specific genomic loci (Jinek *et al.* 2012). CRISPR/Cas originates from a microbial adaptive immune system that uses RNA-guided nucleases to cleave foreign genetic material (Jinek *et al.* 2012). The CRISPR/Cas9 used for DSB induction in eukaryote organisms is based on bacterial type II CRISPR/Cas systems, consisting of CRISPR-associated protein Cas9 and a single guide RNA chimera (sgRNA) which was engineered from the tracrRNA (trans-activated CRISPR RNA) and crRNA (CRISPR RNA) (Jinek *et al.* 2012). Cas9 contains two endonuclease domains (HNH- and RuvC-like domains) cleaving both strands of the target DNA, directed by the sgRNA via Watson-Crick base-pairing to the target DNA sequence (Jinek *et al.* 2012). The cleavage locations are also determined by a protospacer adjacent motif (PAM) which is juxtaposed to the complementary region in the target DNA (Jinek *et al.* 2012). In contrast to ZFNs and TALENs, the CRISPR/Cas9 system is markedly easier to design. It requires only a change in the guide RNA sequence and is highly specific and efficient for a vast number of cell types and organisms. During the past three years, the CRISPR/Cas9 technique has been applied successfully in mammals, plants and yeast (review by Kim 2016).

Outline of the thesis

The non-homologous end joining (NHEJ) repair pathway has been well characterized in yeast and mammals, but there is still much to learn in plants. In plants, the main factors of c-NHEJ have been identified; our previous work showed that the b-NHEJ pathway also exist in plants, depending on Parp1 and Parp2 (Jia *et al.* 2013). As it was possible that the same DNA repair

pathways that mediate T-DNA integration in yeast might act also in plants, the first objective of this project was to analyze whether *Agrobacterium*-mediated transformation efficiency was affected by the absence of NHEJ factors in *Arabidopsis*. The second objective was to investigate the NHEJ pathways in *Arabidopsis*.

In **chapter 2**, AtParp3 and AtXrcc1 were functionally characterized, and double and triple mutants in NHEJ factors were obtained by crossing of single mutants. Together with the *ku70*, *ku80*, *parp1*, *parp2*, *parp1parp2*, *lig4*, *lig6* and *lig4lig6* mutants, which were characterized by our lab previously, they were all tested for T-DNA integration in a root transformation assay. In plants, *lig4lig6* double mutants are still able to perform end-joining (Jia, 2011), suggesting that there must still be another ligase to take over their function. Therefore, a putative DNA ligase gene (*LIG1a*) was studied in **chapter 3**. Mre11 has been reported to play a role both in c-NHEJ and b-NHEJ pathways in mammals. In order to analyze whether Mre11 also functions in b-NHEJ in plants, the *ku80mre11* double mutant was obtained by crossing. The function of Mre11 in plants was studied in **chapter 4**. **Chapter 5** focuses on how exactly the DNA ends join in the *Arabidopsis* NHEJ mutants *in vivo*, which were deficient in c-NHEJ, b-NHEJ or both. Sequence-specific nucleases were used to induced double strand breaks in *Arabidopsis* NHEJ mutants.

References

- Aguilar-quesada, R., J. A. Muñoz-gámez, D. Martín-Oliva, A. Peralta, M. T. Valenzuela *et al.*, 2007 Interaction between ATM and PARP-1 in response to DNA damage and sensitization of ATM deficient cells through PARP inhibition. *BMC Mol. Biol.* 8: 29.
- Ahnesorg, P., P. Smith, and S. P. Jackson, 2006 XLF interacts with the XRCC4-DNA Ligase IV complex to promote DNA nonhomologous end-joining. *Cell* 124: 301–313.
- Alani, E., J. D. Griffith, D. Kolodner, N. Carolina, and C. Hill, 1992 Characterization of DNA-binding and strand-exchange stimulation properties of γ -RPA, a yeast protein DNA-binding properties. *J. Mol. Bio.* 227: 54–71.
- Amoroso, A., L. Concia, C. Maggio, C. Raynaud, C. Bergounioux *et al.*, 2011 Oxidative DNA damage bypass in *Arabidopsis thaliana* requires DNA polymerase λ and proliferating cell nuclear antigen 2. *Plant Cell* 23: 806–822.
- Arakawa, H., T. Bednar, M. Wang, K. Paul, E. Mladenov *et al.*, 2012 Functional redundancy between DNA ligases I and III in DNA replication in vertebrate cells. *Nucleic Acids Res.* 40: 2599–610.
- Audebert, M., B. Salles, and P. Calsou, 2004 Involvement of poly(ADP-ribose) polymerase-1 and XRCC1/DNA ligase III in an alternative route for DNA double-strand breaks rejoining. *J. Biol. Chem.* 279: 55117–55126.
- Baudat, F., and B. de Massy, 2007 Regulating double-stranded DNA break repair towards crossover or non-crossover during mammalian meiosis. *Chromosome Res.* 15: 565–77.
- Block, W. D., Y. Yu, and S. P. Lees-Miller, 2004 Phosphatidyl inositol 3-kinase-like serine/threonine protein kinases (PIKKs) are required for DNA damage-induced phosphorylation of the 32 kDa subunit of replication protein A at threonine 21. *Nucleic Acids Res.* 32: 997–1005.
- de Boer, P., J. Bastiaans, H. Touw, R. Kerkman, J. Bronkhof *et al.*, 2010 Highly efficient gene targeting in *Penicillium chrysogenum* using the bi-partite approach in Δ lig4 or Δ ku70 mutants. *Fungal*

- Genet. Biol. 47: 839–846.
- Buis, J., T. Stoneham, E. Spehalski, and D. O. Ferguson, 2012 Mre11 regulates CtIP-dependent double-strand break repair by interaction with CDK2. *Nat. Struct. Mol. Biol.* 19: 246–252.
- Bundock, P., H. van Attikum, and P. Hooykaas, 2002 Increased telomere length and hypersensitivity to DNA damaging agents in an Arabidopsis KU70 mutant. *Nucleic Acids Res.* 30: 3395–3400.
- Bundock, P., A. den Dulk-Ras, A. Beijersbergen, and P. J. J. Hooykaas, 1995 Trans-kingdom T-DNA transfer from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae*. *EMBO J.* 14: 3206–14.
- Bundock, P., and P. Hooykaas, 2002 Severe developmental defects, hypersensitivity to DNA-damaging agents, and lengthened telomeres in Arabidopsis MRE11 mutants. *Plant Cell* 14: 2451–2462.
- Caldecott, K. W., 2008 Single-strand break repair and genetic disease. *Nat. Rev. Genet.* 9: 619–631.
- Cermak, T., E. Doyle, and M. Christian, 2011 Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic acids research* 39: e82.
- Chan, S. H., A. M. Yu, and M. McVey, 2010 Dual roles for DNA polymerase theta in alternative end-joining repair of double-strand breaks in *Drosophila*. *PLoS Genet.* 6: 1–16.
- Chapman, J. R., M. R. G. Taylor, and S. J. Boulton, 2012 Playing the end game: DNA double-strand break repair pathway choice. *Mol. Cell* 47: 497–510.
- Chen, L., K. Trujillo, W. Ramos, P. Sung, and A. E. Tomkinson, 2001 Promotion of Dnl4-Catalyzed DNA end-joining by the Rad50/Mre11/Xrs2 and Hdf1/Hdf2 complexes. *Mol. Cell* 8: 1105–1115.
- Chilton, M.-D. M., and Q. Que, 2003 Targeted integration of T-DNA into the tobacco genome at double-stranded breaks: new insights on the mechanism of T-DNA integration. *Plant Physiol.* 133: 956–965.
- Christian, M., T. Cermak, E. L. Doyle, C. Schmidt, F. Zhang *et al.*, 2010 Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* 186: 756–761.
- Ciccia, A., and S. J. Elledge, 2010 The DNA Damage Response: making it safe to play with knives. *Mol. Cell* 40: 179–204.
- Conway, A. B., T. W. Lynch, Y. Zhang, G. S. Fortin, C. W. Fung *et al.*, 2004 Crystal structure of a Rad51 filament. *Nat. Struct. Mol. Biol.* 11: 791–6.
- Cools, T., and L. De Veylder, 2009 DNA stress checkpoint control and plant development. *Curr. Opin. Plant Biol.* 12: 23–28.
- Costantini, S., L. Woodbine, L. Andreoli, P. A. Jeggo, and A. Vindigni, 2007 Interaction of the Ku heterodimer with the DNA ligase IV/Xrcc4 complex and its regulation by DNA-PK. *DNA Repair* 6: 712–722.
- Crespan, E., T. Czabany, G. Maga, and U. Hübscher, 2012 Microhomology-mediated DNA strand annealing and elongation by human DNA polymerases λ and β on normal and repetitive DNA sequences. *Nucleic Acids Res.* 40: 5577–5590.
- Culligan, K. M., C. E. Robertson, J. Foreman, P. Doerner, and A. B. Britt, 2006 ATR and ATM play both distinct and additive roles in response to ionizing radiation. *Plant J.* 48: 947–961.
- de Pater, S., L. W. Neuteboom, J. E. Pinas, P. J. J. Hooykaas, and B. J. Van Der Zaal, 2009 ZFN-induced mutagenesis and gene-targeting in Arabidopsis through *Agrobacterium*-mediated floral dip transformation. *Plant Biotechnol. J.* 7: 821–835.
- de Pater, S., J. E. Pinas, P. J. J. Hooykaas, and B. J. van der Zaal, 2013 ZFN-mediated gene targeting of the Arabidopsis protoporphyrinogen oxidase gene through *Agrobacterium*-mediated floral dip

- transformation. *Plant Biotechnol. J.* 11: 510–515.
- Deriano, L., and D. B. Roth, 2013 Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. *Annu. Rev. Genet.* 47: 433–455.
- Durai, S., M. Mani, K. Kandavelou, J. Wu, M. H. Porteus *et al.*, 2005 Zinc finger nucleases: Custom-designed molecular scissors for genome engineering of plant and mammalian cells. *Nucleic Acids Res.* 33: 5978–5990.
- Elliott, B., and M. Jasin, 2002 Cellular and Molecular Life Sciences Human Genome and Diseases : Review Double-strand breaks and translocations in cancer. *CMLS, Cell. Mol. Life Sci.* 59: 373–385.
- Escribano-Díaz, C., A. Orthwein, A. Fradet-Turcotte, M. Xing, J. T. F. Young *et al.*, 2013 A Cell Cycle-Dependent Regulatory Circuit Composed of 53BP1-RIF1 and BRCA1-CtIP Controls DNA Repair Pathway Choice. *Mol. Cell* 49: 872–883.
- Falck, J., J. Coates, and S. P. Jackson, 2005 Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage. *Nature* 434: 605–611.
- Falck, J., J. V. Forment, J. Coates, M. Mistrik, J. Lukas *et al.*, 2012 CDK targeting of NBS1 promotes DNA-end resection, replication restart and homologous recombination. *EMBO Rep.* 13: 561–568.
- Fausser, F., N. Roth, M. Pacher, G. Ilg, R. Sanchez-Fernandez *et al.*, 2012 In planta gene targeting. *Proc. Natl. Acad. Sci. USA.* 109: 7535–7540.
- Ferretti, L. P., L. Lafranchi, and A. Sartori, 2013 Controlling DNA-end resection: a new task for CDKs. *Front. Genet.* 4: 99.
- Friesner, J., and A. B. Britt, 2003 Ku80- and DNA ligase IV-deficient plants are sensitive to ionizing radiation and defective in T-DNA integration. *Plant J.* 34: 427–440.
- Gallego, M. E., J.-Y. Bleuyard, S. Daoudal-Cotterell, N. Jallut, and C. I. White, 2003 Ku80 plays a role in non-homologous recombination but is not required for T-DNA integration in Arabidopsis. *Plant J.* 35: 557–565.
- Gallego, M. E., M. Jeanneau, F. Granier, D. Bouchez, N. Bechtold *et al.*, 2001 Disruption of the Arabidopsis RAD50 gene leads to plant sterility and MMS sensitivity. *Plant J.* 25: 31–41.
- Garcia, V., H. Bruchet, D. Camescasse, F. Granier, D. Bouchez *et al.*, 2003 AtATM is essential for meiosis and the somatic response to DNA damage in plants. *Plant Cell* 15: 119–132.
- Gasior, S. L., A. K. Wong, Y. Kora, A. Shinohara, and D. K. Bishop, 1998 Rad52 associates with RPA and functions with Rad55 and Rad57 to assemble meiotic recombination complexes. *Genes Dev.* 12: 2208–2221.
- Gelvin, S. B., 2000 *Agrobacterium* and Plant Genes involved in T-DNA transfer and integration. *Annu. Rev. Plant Biol.* 51: 223–256.
- Gelvin, S. B., 2010 Plant proteins involved in *Agrobacterium*-mediated genetic transformation. *Annu. Rev. Phytopathol.* 48: 45–68.
- van Gent, D. C., and M. van der Burg, 2007 Non-homologous end-joining, a sticky affair. *Oncogene* 26: 7731–7740.
- Gheysen, G., R. Villarroel, and M. Van Montagu, 1991 Illegitimate recombination in plants: A model for T-DNA integration. *Genes Dev.* 5: 287–297.
- Grawunder, U., M. Wilm, X. Wu, P. Kulesza, T. E. Wilson *et al.*, 1997 Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. *Nature* 388: 492–

- 495.
- Habert, E., 1996 Double-strand break repair in the absence of RAD51 in yeast: A possible role for break-induced DNA replication. *Proc. Natl. Acad. Sci. USA.* 93: 7131–7136.
- Haince, J. F., D. McDonald, A. Rodrigue, U. Dery, J. Y. Masson *et al.*, 2008 PARP1-dependent kinetics of recruitment of MRE11 and NBS1 proteins to multiple DNA damage sites. *J. Biol. Chem.* 283: 1197–1208.
- Hammel, M., Y. Yu, B. L. Mahaney, B. Cai, R. Ye *et al.*, 2010 Ku and DNA-dependent protein kinase dynamic conformations and assembly regulate DNA binding and the initial non-homologous end joining complex. *J. Biol. Chem.* 285: 1414–1423.
- Harper, J. W., and S. J. Elledge, 2007 The DNA damage response: ten years after. *Mol. Cell* 28: 739–745.
- Heyer, W.-D., K. T. Ehmsen, and J. Liu, 2010 Regulation of homologous recombination in eukaryotes. *Annu. Rev. Genet.* 44: 113–139.
- Holloman, W. K., 2011 Unraveling the mechanism of BRCA2 in homologous recombination. *Nat. Struct. Mol. Biol.* 18: 748–754.
- Jia, Q., P. Bundock, P. J. J. Hooykaas, and S. de Pater, 2012 *Agrobacterium tumefaciens* T-DNA integration and gene targeting in *Arabidopsis thaliana* Non-Homologous End-Joining mutants. *J. Bot.* 2012: 1–13.
- Jia, Q., A. Den Dulk-Ras, H. Shen, P. J. J. Hooykaas, and S. de Pater, 2013 Poly(ADP-ribose)polymerases are involved in microhomology mediated back-up non-homologous end joining in *Arabidopsis thaliana*. *Plant Mol. Biol.* 82: 339–351.
- Jia, Q., 2011 DNA repair and gene targeting in plant end-joining mutants. PhD thesis. Leiden University, Leiden, The Netherlands.
- Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J. a Doudna *et al.*, 2012 A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337: 816–821.
- Johnson, R. A., R. P. Hellens, and D. R. Love, 2011 A transient assay for recombination demonstrates that *Arabidopsis* SNM1 and XRCC3 enhance non-homologous recombination. *Genet. Mol. Res.* 10: 2104–2132.
- Joung, J. K., and J. D. Sander, 2013 TALENs: a widely applicable technology for targeted genome editing. *Nat. Rev. Mol. Cell Biol.* 14: 49–55.
- Kadyk, L. C., and L. H. Hartwell, 1992 Sister chromatids are preferred over homologs as substrates for recombination repair in *Saccharomyces cerevisiae*. *Genetics* 2.: 387–402.
- Karimi-Busheri, F., A. Rasouli-Nia, J. Allalunis-Turner, and M. Weinfeld, 2007 Human polynucleotide kinase participates in repair of DNA double-strand breaks by nonhomologous end joining but not homologous recombination. *Cancer Res.* 67: 6619–6625.
- Kent, T., G. Chandramouly, S. M. McDevitt, A. Y. Ozdemir, and R. T. Pomerantz, 2015 Mechanism of microhomology-mediated end-joining promoted by human DNA polymerase theta. *Nat Struct Mol Biol* 22: 230–237.
- Khanna, K. K., and S. P. Jackson, 2001 DNA double-strand breaks: signaling, repair and the cancer connection. *Nat. Genet.* 27: 247–254.
- Kim, J.-S., 2016 Genome editing comes of age. *Nat. Protoc.* 11: 1573–1578.
- Koch, C. A., R. Agyei, S. Galicia, P. Metalnikov, P. O'Donnell *et al.*, 2004 Xrcc4 physically links DNA end processing by polynucleotide kinase to DNA ligation by DNA ligase IV. *EMBO J.* 23: 3874–

3885.

- Kooistra, R., P. J. J. Hooykaas, and H. Y. Steensma, 2004 Efficient gene targeting in *Kluyveromyces lactis*. *Yeast* 21: 781–792.
- Kraus, E., W. Y. Leung, and J. E. Haber, 2001 Break-induced replication: a review and an example in budding yeast. *Proc. Natl. Acad. Sci. USA.* 98: 8255–8262.
- Krejci, L., V. Altmannova, M. Spirek, and X. Zhao, 2012 Homologous recombination and its regulation. *Nucleic Acids Res.* 40: 5795–5818.
- Krishna, S., B. M. Wagener, H. P. Liu, Y.-C. Lo, R. Sterk *et al.*, 2007 Mre11 and Ku regulation of double-strand break repair by gene conversion and break-induced replication. *DNA Repair* 6: 797–808.
- Lee-Theilen, M., A. J. Matthews, D. Kelly, S. Zheng, and J. Chaudhuri, 2011 CtIP promotes microhomology-mediated alternative end joining during class-switch recombination. *Nat. Struct. Mol. Biol.* 18: 75–79.
- Li, M., L. Y. Lu, C. Y. Yang, S. Wang, and X. Yu, 2013 The FHA and BRCT domains recognize ADP-ribosylation during DNA damage response. *Genes Dev.* 27: 1752–1768.
- Li, J., M. Vaidya, C. White, A. Vainstein, V. Citovsky *et al.*, 2005 Involvement of KU80 in T-DNA integration in plant cells. *Proc. Natl. Acad. Sci. USA.* 102: 19231–19236.
- Lisby, M., J. H. Barlow, R. C. Burgess, R. Rothstein, W. Street *et al.*, 2004 Choreography of the DNA damage response: spatiotemporal relationships among checkpoint and repair proteins. *Cell* 118: 699–713.
- Lydeard, J. R., S. Jain, M. Yamaguchi, and J. E. Haber, 2007 Break-induced replication and telomerase-independent telomere maintenance require Pol32. *Nature* 448: 820–823.
- Ma, Y., U. Pannicke, K. Schwarz, and M. R. Lieber, 2002 Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. *Cell* 108: 781–794.
- Ma, Y., K. Schwarz, and M. R. Lieber, 2005 The Artemis:DNA-PKcs endonuclease cleaves DNA loops, flaps, and gaps. *DNA Repair* 4: 845–851.
- Mahajan, K. N., S. A. Nick McElhinny, B. S. Mitchell, and D. A. Ramsden, 2002 Association of DNA polymerase mu (pol mu) with Ku and ligase IV: role for pol mu in end-joining double-strand break repair. *Mol. Cell. Biol.* 22: 5194–5202.
- Mahaney, B. L., K. Meek, and S. P. Lees-Miller, 2009 Repair of ionizing radiation-induced DNA double-strand breaks by non-homologous end-joining. *Biochem. J.* 417: 639–650.
- Mari, P.-O., B. I. Florea, S. P. Persengiev, N. S. Verkaik, H. T. Brüggewirth *et al.*, 2006 Dynamic assembly of end-joining complexes requires interaction between Ku70/80 and XRCC4. *Proc. Natl. Acad. Sci. USA.* 103: 18597–18602.
- Martínez-Macías, M. I., W. Qian, D. Miki, O. Pontes, Y. Liu *et al.*, 2012 A DNA 3' Phosphatase Functions in Active DNA Demethylation in Arabidopsis. *Mol. Cell* 45: 357–370.
- Mateos-Gomez, P. A., F. Gong, N. Nair, K. M. Miller, E. Lazzerini-Denchi *et al.*, 2015 Mammalian polymerase θ promotes alternative NHEJ and suppresses recombination. *Nature* 518: 254–257.
- Matsuoka, S., B. A. Ballif, A. Smogorzewska, E. R. McDonald, K. E. Hurov *et al.*, 2007 ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* 316: 1160–1166.
- Mayerhofer, R., Z. Koncz-kalman, C. Nawrath, G. Bakkeren, A. Cramer *et al.*, 1991 T-DNA

- integration : in plants mode of illegitimate recombination. *EMBO J.* 10: 697–704.
- Mazin, A. V., A. Alexeev, and S. C. Kowalczykowski, 2003 A novel function of Rad54 protein. Stabilization of the Rad51 nucleoprotein filament. *J. Biol. Chem.* 278: 14029–14036.
- McIlwraith, M. J., and S. C. West, 2008 DNA repair synthesis facilitates RAD52-mediated second-end capture during DSB repair. *Mol. Cell* 29: 510–516.
- Meek, K., V. Dang, and S. P. Lees-Miller, 2008 DNA-PK: the means to justify the ends? *Adv. Immunol.* 99: 33–58.
- Mestiri, I., F. Norre, M. E. Gallego, and C. I. White, 2014 Multiple host-cell recombination pathways act in *Agrobacterium*-mediated transformation of plant cells. *Plant J.* 77: 511–520.
- Mimitou, E. P., and L. S. Symington, 2011 DNA end resection--unraveling the tail. *DNA Repair* 10: 344–348.
- Mimitou, E. P., and L. S. Symington, 2009 Nucleases and helicases take center stage in homologous recombination. *Trends Biochem. Sci.* 34: 264–272.
- Mimitou, E. P., and L. S. Symington, 2008 Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. *Nature* 455: 770–774.
- Molinier, J., M.-E. Stamm, and B. Hohn, 2004 SNM-dependent recombinational repair of oxidatively induced DNA damage in *Arabidopsis thaliana*. *EMBO Rep.* 5: 994–999.
- Moynahan, M. E., and M. Jasin, 2010 Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis. *Nat. Rev. Mol. Cell Biol.* 11: 196–207.
- Nassif, N., J. Penney, S. Pal, W. R. Engels, and G. B. Gloor, 1994 Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. *Mol. Cell. Biol.* 14: 1613–1625.
- Neal, J. A., and K. Meek, 2011 Choosing the right path: Does DNA-PK help make the decision? *Mutat. Res. - Fundam. Mol. Mech. Mutagen.* 711: 73–86.
- Nick McElhinny, S. A., C. M. Snowden, J. McCarville, and D. A. Ramsden, 2000 Ku recruits the XRCC4-ligase IV complex to DNA ends. *Mol. Cell. Biol.* 20: 2996–3003.
- Nimonkar, A. V., R. A. Sica, and S. C. Kowalczykowski, 2009 Rad52 promotes second-end DNA capture in double-stranded break repair to form complement-stabilized joint molecules. *Proc. Natl. Acad. Sci. USA.* 106: 3077–3082.
- Offringa, R., M. J. de Groot, H. J. Haagsman, M. P. Does, P. J. van den Elzen *et al.*, 1990 Extrachromosomal homologous recombination and gene targeting in plant cells after *Agrobacterium* mediated transformation. *EMBO J.* 9: 3077–3084.
- Ogawa, T., X. Yu, A. Shinohara, and E. H. Egelman, 1993 Similarity of the yeast RAD51 filament to the bacterial RecA filament. *Science* 374: 1991–1994.
- Osman, K., J. D. Higgins, E. Sanchez-Moran, S. J. Armstrong, and F. C. H. Franklin, 2011 Pathways to meiotic recombination in *Arabidopsis thaliana*. *New Phytol.* 190: 523–544.
- Owalczykowski, S. T. C. K., 1998 DNA annealing by Rad52 Protein is stimulated by specific interaction with the complex of replication protein A and single-stranded DNA. *Proc. Natl. Acad. Sci. USA.* 95: 6049–6054.
- Park, S.-Y., Z. Vaghchhipawala, B. Vasudevan, L.-Y. Lee, Y. Shen *et al.*, 2015 *Agrobacterium* T-DNA integration into the plant genome can occur without the activity of key non-homologous end-joining proteins. *Plant J.* 81: 934–946.
- Paul, K., M. Wang, E. Mladenov, A. Bencsik-Theilen, T. Bednar *et al.*, 2013 DNA ligases I and III cooperate in alternative non-homologous end-joining in vertebrates. *PLoS One* 8: e59505.

- Paull, T. T., and M. Gellert, 1998 The 3' to 5' Exonuclease Activity of Mre11 Facilitates Repair of DNA Double-Strand Breaks. *Mol. Cell* 1: 969–979.
- Petrucchio, S., G. Volpi, A. Bolchi, C. Rivetti, and S. Ottonello, 2002 A nick-sensing DNA 3'-repair enzyme from *Arabidopsis*. *J. Biol. Chem.* 277: 23675–23683.
- Prakash, R., D. Satory, E. Dray, A. Papusha, J. Scheller *et al.*, 2009 Yeast Mph1 helicase dissociates Rad51-made D-loops: implications for crossover control in mitotic recombination. *Genes Dev.* 23: 67–79.
- Qi, Y., Y. Zhang, F. Zhang, J. a. Baller, S. C. Cleland *et al.*, 2013 Increasing frequencies of site-specific mutagenesis and gene targeting in *Arabidopsis* by manipulating DNA repair pathways. *Genome Res.* 23: 547–554.
- Reiss, B., I. Schubert, K. Köpchen, E. Wendeler, J. Schell *et al.*, 2000 RecA stimulates sister chromatid exchange and the fidelity of double-strand break repair, but not gene targeting, in plants transformed by *Agrobacterium*. *Proc. Natl. Acad. Sci. USA.* 97: 3358–3363.
- Richardson, C., M. Jasin, and N. Hprt, 2000 Frequent chromosomal translocations induced by DNA double-strand breaks. *Nature* 1: 697–700.
- Rodenburg, K. W., M. J. de Groot, R. a Schilperoort, and P. J. J. Hooykaas, 1989 Single-stranded DNA used as an efficient new vehicle for transformation of plant protoplasts. *Plant Mol. Biol.* 13: 711–719.
- Roy, S., S. R. Choudhury, D. N. Sengupta, and K. P. Das, 2013 Involvement of AtPol λ in the repair of high salt- and DNA cross-linking agent-induced double strand breaks in *Arabidopsis*. *Plant Physiol.* 162: 1195–1210.
- Roy, S., S. R. Choudhury, S. K. Singh, and K. P. Das, 2011 AtPol λ , a homolog of mammalian DNA polymerase λ in *Arabidopsis thaliana*, is involved in the repair of UV-B induced DNA damage through the dark repair pathway. *Plant Cell Physiol.* 52: 448–467.
- Salomon, S., and H. Puchta, 1998 Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. *EMBO J.* 17: 6086–6095.
- De Schutter, K., J. Joubes, T. Cools, A. Verkest, F. Corellou *et al.*, 2007 *Arabidopsis* WEE1 kinase controls cell cycle arrest in response to activation of the DNA integrity checkpoint. *Plant Cell* 19: 211–225.
- Seeliger, K., S. Dukowic-Schulze, R. Wurz-Wildersinn, M. Pacher, and H. Puchta, 2012 BRCA2 is a mediator of RAD51- and DMC1-facilitated homologous recombination in *Arabidopsis thaliana*. *New Phytol.* 193: 364–375.
- Shaked, H., C. Melamed-Bessudo, and A. A. Levy, 2005 High-frequency gene targeting in *Arabidopsis* plants expressing the yeast RAD54 gene. *Proc. Natl. Acad. Sci. USA.* 102: 12265–12269.
- Shiloh, Y., 2006 The ATM-mediated DNA-damage response: taking shape. *Trends Biochem. Sci.* 31: 402–410.
- Simsek, D., E. Brunet, S. Y. W. Wong, S. Katyal, Y. Gao *et al.*, 2011 DNA ligase III promotes alternative nonhomologous end-joining during chromosomal translocation formation. *PLoS Genet.* 7(6): e1002080.
- Smih, F., P. Rouet, P. J. Romanienko, and M. Jasin, 1995 Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells. *Nucleic Acids Res.* 23: 5012–5019.
- Sugawara, N., and J. E. Haber, 1992 Characterization of double-strand break-induced recombination: homology requirements and single-stranded DNA formation. *Mol. Cell. Biol.* 12: 563–575.

- Sugawara, N., X. Wang, J. E. Haber, and G. A. L. Ho, 2003 In vivo roles of Rad52 , Rad54 , and Rad55 proteins in Rad51-mediated recombination. *Mol. Cell* 12: 209–219.
- Sugiyama, T., N. Kantake, Y. Wu, and S. C. Kowalczykowski, 2006 Rad52-mediated DNA annealing after Rad51-mediated DNA strand exchange promotes second ssDNA capture. *EMBO J.* 25: 5539–5548.
- Sugiyama, T., and S. C. Kowalczykowski, 2002 Rad52 protein associates with replication protein A (RPA)-single-stranded DNA to accelerate Rad51-mediated displacement of RPA and presynaptic complex formation. *J. Biol. Chem.* 277: 31663–31672.
- Sugiyama, T., E. M. Zaitseva, and S. C. Kowalczykowski, 1997 A single-stranded DNA-binding protein is needed for efficient presynaptic complex formation by the *Saccharomyces cerevisiae* Rad51 protein. *J. Biol. Chem.* 272: 7940–7945.
- Sung, P., 1994 Catalysis of ATP-dependent homologous DNA pairing and strand exchange by yeast RAD51 protein. *Science* 265: 1241–1243.
- Symington, L. S., and J. Gautier, 2011 Double-Strand Break End Resection and Repair Pathway Choice. *Annu. Rev. Genet.* 45: 247–271.
- Szostak, J. W., T. L. Orr-Weaver, R. J. Rothstein, and F. W. Stahl, 1983 The double-strand-break repair model for recombination. *Cell* 33: 25–35.
- Tamura, K., Y. Adachi, K. Chiba, K. Oguchi, and H. Takahashi, 2002 Identification of Ku70 and Ku80 homologues in *Arabidopsis thaliana*: Evidence for a role in the repair of DNA double-strand breaks. *Plant J.* 29: 771–781.
- Tomimatsu, N., B. Mukherjee, M. Catherine Hardebeck, M. Ilcheva, C. Vanessa Camacho *et al.*, 2014 Phosphorylation of EXO1 by CDKs 1 and 2 regulates DNA end resection and repair pathway choice. *Nat. Commun.* 5: 3561.
- Tzfira, T., L. R. Frankman, M. Vaidya, and V. Citovsky, 2003 Site-specific integration of *Agrobacterium tumefaciens* T-DNA via double-stranded intermediates. *Plant Physiol.* 133: 1011–1023.
- Tzfira, T., J. Li, B. Lacroix, and V. Citovsky, 2004 *Agrobacterium* T-DNA integration: molecules and models. *Trends Genet.* 20: 375–383.
- Uchiyama, Y., S. Kimura, T. Yamamoto, T. Ishibashi, and K. Sakaguchi, 2004 Plant DNA polymerase λ , a DNA repair enzyme that functions in plant meristematic and meiotic tissues. *Eur. J. Biochem.* 271: 2799–2807.
- Uematsu, N., E. Weterings, K. I. Yano, K. Morotomi-Yano, B. Jakob *et al.*, 2007 Autophosphorylation of DNA-PKCS regulates its dynamics at DNA double-strand breaks. *J. Cell Biol.* 177: 219–229.
- Urnov, F. D., E. J. Rebar, M. C. Holmes, H. S. Zhang, and P. D. Gregory, 2010 Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* 11: 636–646.
- van Attikum, H., P. Bundock, and P. J. J. Hooykaas, 2001 Non-homologous end-joining proteins are required for *Agrobacterium* T-DNA integration. *EMBO J.* 20: 6550–6558.
- van Attikum, H., P. Bundock, R. M. Overmeer, L. Y. Lee, S. B. Gelvin *et al.*, 2003 The *Arabidopsis* AtLIG4 gene is required for the repair of DNA damage, but not for the integration of *Agrobacterium* T-DNA. *Nucleic Acids Res.* 31: 4247–4255.
- van Attikum, H., and P. J. J. Hooykaas, 2003 Genetic requirements for the targeted integration of *Agrobacterium* T-DNA in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 31: 826–832.
- Walker, J. R., R. A. Corpina, and J. Goldberg, 2001 Structure of the Ku heterodimer bound to DNA

- and its implications for double-strand break repair. *Nature* 412: 607–614.
- Wang, H., L. Z. Shi, C. C. L. Wong, X. Han, P. Y. H. Hwang *et al.*, 2013 The interaction of CtIP and Nbs1 connects CDK and ATM to regulate HR-mediated double-strand break repair. *PLoS Genet.* 9(2): e1003277.
- Waterworth, W. M., C. Altun, S. J. Armstrong, N. Roberts, P. J. Dean *et al.*, 2007 NBS1 is involved in DNA repair and plays a synergistic role with ATM in mediating meiotic homologous recombination in plants. *Plant J.* 52: 41–52.
- Waterworth, W. M., J. Kozak, C. M. Provost, C. M. Bray, K. J. Angelis *et al.*, 2009 DNA ligase 1 deficient plants display severe growth defects and delayed repair of both DNA single and double strand breaks. *BMC Plant Biol.* 9: 79.
- Weinfeld, M., R. S. Mani, I. Abdou, R. D. Aceytuno, and J. N. M. Glover, 2011 Tidying up loose ends: The role of polynucleotide kinase/phosphatase in DNA strand break repair. *Trends Biochem. Sci.* 36: 262–271.
- West, C. E., W. M. Waterworth, Q. Jiang, and C. M. Bray, 2000 Arabidopsis DNA ligase IV is induced by γ -irradiation and interacts with an Arabidopsis homologue of the double strand break repair protein XRCC4. *Plant J.* 24: 67–78.
- West, C. E., W. M. Waterworth, G. W. Story, P. A. Sunderland, Q. Jiang *et al.*, 2002 Disruption of the Arabidopsis AtKu80 gene demonstrates an essential role for AtKu80 protein in efficient repair of DNA double-strand breaks in vivo. *Plant J.* 31: 517–528.
- Williams, R. S., G. Moncalian, J. S. Williams, Y. Yamada, O. Limbo *et al.*, 2008 Mre11 dimers coordinate DNA end bridging and nuclease processing in double-strand-break repair. *Cell* 135: 97–109.
- Wolner, B., and C. L. Peterson, 2005 ATP-dependent and ATP-independent roles for the Rad54 chromatin remodeling enzyme during recombinational repair of a DNA double strand break. *J. Biol. Chem.* 280: 10855–10860.
- Wu, Y., N. Kantake, T. Sugiyama, and S. C. Kowalczykowski, 2008 Rad51 protein controls Rad52-mediated DNA annealing. *J. Biol. Chem.* 283: 14883–14892.
- Wu, Y., T. Sugiyama, and S. C. Kowalczykowski, 2006 DNA annealing mediated by Rad52 and Rad59 proteins. *J. Biol. Chem.* 281: 15441–15449.
- Yano, K., K. Morotomi-Yano, S.-Y. Wang, N. Uematsu, K.-J. Lee *et al.*, 2008 Ku recruits XLF to DNA double-strand breaks. *EMBO Rep.* 9: 91–96.
- Yoshiyama, K. O., J. Kobayashi, N. Ogita, M. Ueda, S. Kimura *et al.*, 2013 ATM-mediated phosphorylation of SOG1 is essential for the DNA damage response in Arabidopsis. *EMBO Rep.* 14: 817–822.
- Zhang, F., L. Cong, S. Lodato, S. Kosuri, G. M. Church *et al.*, 2011 Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat. Biotechnol.* 29: 149–153.
- Zhang, Z., W. Hu, L. Cano, T. D. Lee, D. J. Chen *et al.*, 2004 Solution structure of the C-terminal domain of Ku80 suggests important sites for protein-protein interactions. *Structure* 12: 495–502.
- Zhang, Y., and M. Jasin, 2011 An essential role for CtIP in chromosomal translocation formation through an alternative end-joining pathway. *Nat. Struct. Mol. Biol.* 18: 80–84.
- Zhang, Z., L. Zhu, D. Lin, F. Chen, D. J. Chen *et al.*, 2001 The Three-dimensional Structure of the C-terminal DNA-binding Domain of Human Ku70. *J. Biol. Chem.* 276: 38231–38236.

Zhu, Z., W.-H. Chung, E. Y. Shim, S. E. Lee, and G. Ira, 2008 Sgs1 helicase and two nucleases Dna2 and Exo1 resect DNA double-strand break ends. *Cell* 134: 981–994.

Zhuang, J., G. Jiang, H. Willers, and F. Xia, 2009 Exonuclease function of human Mre11 promotes deletional nonhomologous end joining. *J. Biol. Chem.* 284: 30565–30573.

Chapter 2

***Agrobacterium* T-DNA integration in *Arabidopsis* non-homologous end-joining mutants**

Hexi Shen, Amke den Dulk-Ras, Kelly van Kooperen, Paul J. J. Hooykaas,
Sylvia de Pater

Department of Molecular and Developmental Genetics, Institute of Biology, Leiden University,
Leiden, 2333 BE, The Netherlands

Abstract

Arabidopsis thaliana *parp3* or *xrcc1* mutant was isolated, and subsequently double and triple mutants (*parp1parp3*, *ku80xrcc1*, *parp1parp2parp3*) were obtained by crossing. Treatments with DNA damaging agents showed that *PARP3* and *XRCC1* are involved in DNA repair. We further examined transient and stable root transformation frequencies of these mutants after co-cultivation with *Agrobacterium*. Knocking out components of either the c-NHEJ or b-NHEJ pathway, did not lead to a significant decrease in root transformation. However, the *ku80xrcc1* and *ku80p1p2* mutants, in which both c-NHEJ and b-NHEJ pathways are inactivated, showed a significant decrease in stable root transformation efficiency. However, no significant differences were observed in transient transformation. These results indicate that T-DNA integration requires the known NHEJ repair pathways for optimal transformation, but that there must still be other important factors and/or pathways involved in T-DNA integration.

Introduction

Genetic transformation of plants by *Agrobacterium* involves the transfer of T-DNA from its tumor-inducing plasmid to the host cell nuclear genome. In this way, *Agrobacterium* has provided us with a means to produce genetically modified plants. T-DNA is transferred as a single stranded molecule (T-strand) from the bacteria to the plant cell nucleus. During this process, several *Agrobacterium* Vir proteins accompany the T-strand. The T-DNA integrates at a random position in the nuclear genome of the plant cells, but the precise mechanism of T-DNA integration into the plant genome remains unclear. Two possible models for T-DNA integration have been proposed over the years (see for reviews Tzfira *et al.* 2004; Gelvin 2010). In the strand-invasion model, T-DNA integration depends on the microhomology between T-strand and plant DNA sequences. It was suggested that single stranded DNA is preferential for integration (Rodenburg *et al.* 1989; Gheysen *et al.* 1991; Mayerhofer *et al.* 1991). In the DNA double strand break repair model, the T-strand is first converted into double stranded DNA and then this is integrated into a double strand break site in the genome. The second model was supported by the fact that DSBs are the preferential targets for T-DNA integration (Salomon and Puchta 1998; Chilton and Que 2003; Tzfira *et al.* 2003).

Using yeast as a model it was shown in our lab that random T-DNA integration in yeast (*Saccharomyces cerevisiae*) is dependent on the non-homologous end joining (NHEJ) pathway of DSB repair, and that proteins such as Ku70, Ku80 and DNA ligase IV are essential for T-DNA integration (van Attikum *et al.* 2001). Inactivation of the NHEJ pathway still allowed integration via homologous recombination provided that the T-DNA carried an area of homology with the yeast genome (van Attikum and Hooykaas 2003). In this way gene targeting could be facilitated in yeasts and fungi (Kooistra *et al.* 2004). Many proteins involved in the NHEJ pathway are conserved in plants, including Ku70, Ku80 and Lig4 and therefore attempts were made to facilitate gene targeting in plants by inactivating NHEJ in plants. However, studies on T-DNA integration with Arabidopsis NHEJ mutants gave variable results. Initial publications reported about strongly or mildly decreased stable transformation in Arabidopsis *ku70* and *ku80* mutants (Friesner and Britt 2003; Li *et al.* 2005; Jia *et al.* 2012; Mestiri *et al.* 2014), whereas a more recent publication even reported increased T-DNA integration in such mutants (Park *et al.* 2015). A decrease in floral dip transformation was reported for a Arabidopsis *lig4* mutant (Friesner and Britt 2003), but others found no alteration in the frequency of stable transformation in both floral dip and the root tumorigenesis assay (van Attikum *et al.* 2003; Jia *et al.* 2012; Park *et al.* 2015). Therefore, in plants T-DNA integration must be possible by another pathway. In animal and plant cells in the absence of c-NHEJ, DSB repair is possible by backup pathways (b-NHEJ) that so far are not fully characterized and these may play a role in T-DNA integration.

In the b-NHEJ repair pathway in animal cells, Parp1 has been identified to play an essential role together with XRCC1, DNA ligase III, Mre11, as well as other end processing proteins (Audebert *et al.* 2004; Cheng *et al.* 2011). Poly(ADP-ribose) polymerases (PARPs) are abundant nuclear enzymes that have been implicated in many cellular processes in higher eukaryotes, including stress responses, mitosis, transcription and DNA repair (Schreiber *et al.* 2006). Three Parp proteins have been found to be activated in response to DNA damage in animals: Parp1, Parp2 and Parp3 (Gibson and Kraus 2012). Parp1 and Parp2 are involved in

the DNA damage response (DDR), DNA single strand break repair (SSBR) and base excision repair (BER) (Beck, Robert, *et al.* 2014). Parp3 was only recently discovered as a factor recruited to DNA damage sites and which accelerates non-homologous end-joining probably by its interaction with the components of the c-NHEJ pathway, including Ku70/Ku80 and APLF proteins (Boehler *et al.* 2011; Beck *et al.* 2014). Thus, Parp3 is thought to be involved in DSB repair via c-NHEJ.

In plants, XRCC1 has been identified to play an important role in NHEJ in the absence of Ku (Charbonnel *et al.* 2010). Our previous work showed that the Arabidopsis Parp1 and Parp2 are involved in a b-NHEJ pathway, called MMEJ (microhomology-mediated end-joining) which uses microhomology for repair (Jia *et al.* 2013). Triple mutants inactivating both c-NHEJ and b-NHEJ were constructed in our lab. This included the triple mutant *ku80parp1parp2* which was hypersensitive to DNA damage, but still conferred T-DNA integration (Jia, 2011). These results suggest that there may be further redundancy or still other factors are involved in T-DNA integration. Recently, the *PARP3* gene was discovered in plants (Rissel *et al.* 2014; Stolarek *et al.* 2015). Its activity might explain why there was still T-DNA integration in the *ku80parp1parp2* triple mutant.

In order to get a better understanding of NHEJ repair pathways and *Agrobacterium*-mediated T-DNA integration in plants, mutants in either c-NHEJ, b-NHEJ or both pathways were tested in root transformation assays. In order to extend our collection of NHEJ mutants, homozygous *parp3* and *xrcc1* mutants were isolated and characterized. The *parp3* mutant was crossed with *parp1* and *parp2* mutants to obtain the homozygous triple mutant of *parp1parp2parp3* (*p1p2p3*). The single, double and triple mutants were tested for the sensitivity to DNA damaging agents and in a root transformation assay. The *xrcc1* mutant was crossed with *ku80* to obtain the *ku80xrcc1* double mutant. Together with other NHEJ mutants including *ku80*, *ku70*, *xrcc1*, *lig4*, *lig6*, *lig4lig6*, and *ku80p1p2*, the *p1p2p3* and *k80xrcc1* mutant lines were also analyzed for *Agrobacterium*-mediated T-DNA integration in root transformation assays.

Materials and methods

Plant material

The *Atparp3* (At5g22470; SALK_108092) and *Atxrcc1* (At1g80420; SALK_125373) T-DNA insertion lines, ecotype Col-0, were obtained from the SALK collection (Alonso *et al.* 2003). The *Atnre11-2* (Bundock and Hooykaas 2002), *Atparp1* (GABI-Kat Line 692A05) (Jia *et al.* 2013), *Atparp2* (SALK_640400) (Jia *et al.* 2013), *Atlig4* (SALK_044027), *Atku70* (SALK_123114) and *Atku80* (SALK_016627) (Jia *et al.* 2012) and *Atlig6* (SALK_065307) (Jia, 2011) mutants have been previously described.

Characterization of the Atparp3 and Atxrcc1 mutants

DNA was extracted from individual plants using the CTAB DNA isolation protocol. The T-DNA insertion sites of the *parp3* and *xrcc1* mutants were mapped with a gene-specific primer (SP539 or SP544 for *parp3*; SP170 or SP171 for *xrcc1*) and a T-DNA specific primer (LBa1 or RB) and PCR products were cloned and sequenced. Pairs of gene-specific primers around the insertion site were used to determine whether the plants were homozygous or heterozygous

for the T-DNA insertion. The sequences of all the primers are listed in Table1. For Southern blot analysis, 10 µg DNA from the mutants were digested with *Hind*III and separated on a 0.7% agarose gel and transferred onto positively charged Hybond-N membrane (Amersham Biosciences). The hybridization and detection procedures were done according to the DIG protocol from Roche Applied Sciences.

Assays for sensitivity to bleomycin and methyl methane sulfonate (MMS)

Seeds of wild-type, *parp1*, *parp2*, *parp3*, *p1p2*, *p1p3*, *p1p2p3*, *ku80*, *xrcc1*, *ku80xrcc1* mutants were surface-sterilized as described (Weijers *et al.* 2001) and germinated on solid ½ MS medium without additions or containing 0.02 µg/ml bleomycin (Sigma), 0.05 µg/ml bleomycin, 0.005% (v/v) MMS (Sigma) or 0.007% MMS. After 2 weeks the mutants photographed.

Root transformation and GUS assays

Root transformation were performed as described (Vergunst *et al.* 2000). Root segments were infected with *A.tumefaciens* LBA1100 harboring the binary vector pCambia3301. The T-DNA from pCambia3301 contains a phosphinothricin selection cassette and a GUS gene. After co-cultivation on callus induction medium containing 100 µM acetosyringone for 48 hours, root segments were washed, dried and incubated on shoot induction medium plus phosphinothricin 30 µg/ml, 500 µg/ml carbenicillin and 100 µg/ml vancomycin. After 3-4 weeks, plates were photographed and transformation efficiencies was scored as infected root segments that produced any form of callus.

For transient GUS activity assays, root segments were washed after 72 hr cocultivation, dried and stained with X-Gluc overnight at 37°C. Root segments were destained with 70% ethanol and visualized using a microscope.

For quantification of GUS activity, root segments were washed, dried and disrupted to a powder under liquid N₂ in a TissueLyser (Retch, Haan, Germany). Protein extraction buffer (1x Na-phosphate/EDTA buffer PH 7.0, 0.1% sodium lauryl sarcosine (SLS), 0.1% Triton-X100, 10 mM β-mercaptoethanol) was added to tissue powder. Soluble protein was isolated by centrifugation at 4°C. The protein concentration was determined by using the BIO-RAD protein assay reagent. Ten µl protein extracts in duplicate were co-cultivated in 190 µl extraction buffer/MUG at 37°C. The fluorescence value was scored after 0.5, 1, 3, 5 hours by a Perkin Elmer 1420 Fluorimeter. GUS activities from triplicate transformations were normalized against total protein content to correct for differences in protein extraction efficiencies. Statistical analyses were performed using Prism version 5 (GraphPad Software Inc.).

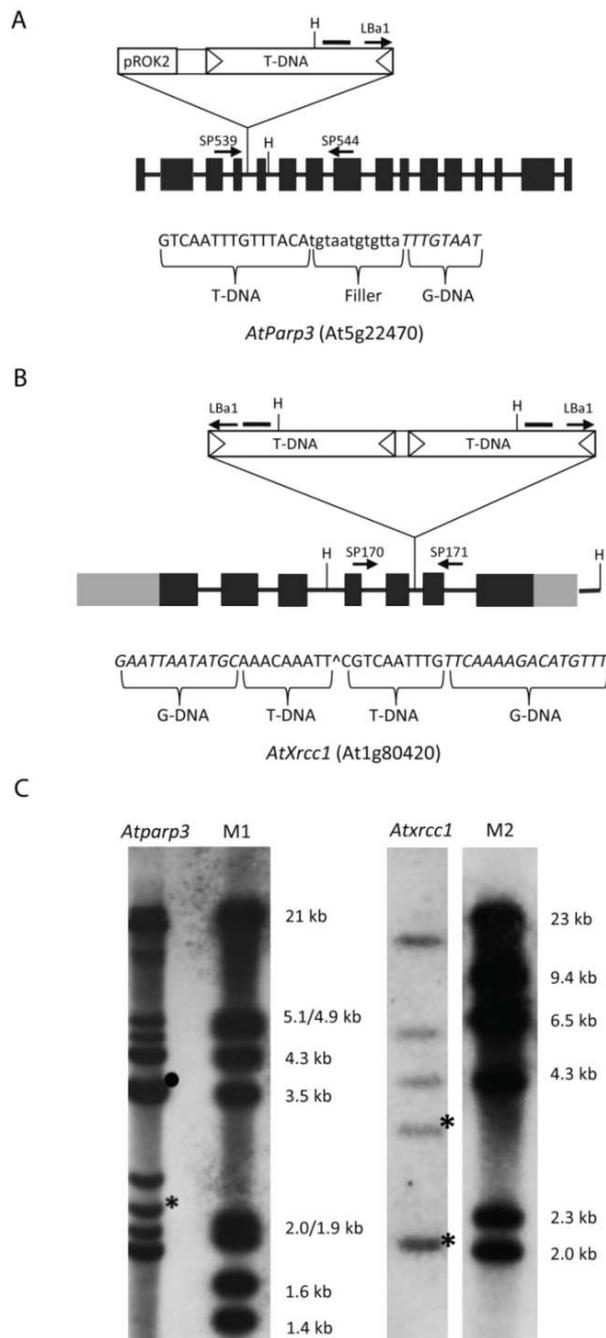


Figure 1. Molecular analysis of the T-DNA insertions in the *PARP3* and *XRCC1* loci. Genomic organization of the *PARP3* (A) and *XRCC1* (B) locus is indicated with the positions of the inserted T-DNA. Exons are shown as black boxes, 3' and 5' UTRs are shown as gray boxes and introns are shown as lines. The probe (—) and the restriction enzyme digestion site (H: *Hind*III) used for Southern blot analysis, are indicated. Genomic DNA sequences (gDNA) flanking the T-DNA insertion are shown in italics. T-DNA border indicated with triangle. (C) Southern blot analysis of the *parp3* and *xrcc1* T-DNA insertion mutant. The genomic DNA was digested by *Hind*III. M1, M2: Lambda DNA *Eco*RI+*Hind*III, Lambda DNA *Hind*III marker. The expected bands of 2174 bp (*parp3* mutant), 2173 bp and 3340 bp (*xrcc1* mutant) are indicated with an asterisk. Tandem repeat of 3.6 kb is indicated with a dot.

Results

*Identification of the *parp3* and *xrcc1* mutants*

Previously it was found that in the triple *ku80p1p2* mutant inactivating both c-NHEJ and b-NHEJ at the same time, T-DNA integration was still observed. Recently, a third *PARP* gene was discovered in plants called *PARP3* (Rissel *et al.* 2014). As *PARP3* is possibly redundant to *PARP2* and *PARP1*, we isolated and characterized the *parp3* homozygous mutant and crossed it with *p1p2* (**Figure 1**). Homozygous mutants were identified by using T-DNA specific primers for the left border or right border in combination with gene-specific primers flanking the insertion site. No PCR products were obtained for homozygous mutants using two gene-specific primers. The insertion site of the T-DNA was mapped by sequencing. The T-DNA was integrated in the *PARP3* gene in intron 4, whereby 12 bp filler DNA was inserted at the LB end (**Figure 1A**). The RB was integrated with part of the original pROK2 vector. Therefore, the RB integration site could not be mapped. For Southern blotting, genomic DNA of the *parp3* mutant was digested with *HindIII*. A diagnostic band of 2174 bp was expected connecting the T-DNA with the *PARP3* gene, which can indeed be seen on the gel (**Figure 1C**). Besides the 2 kb band extra bands were detected, indicating that additional T-DNAs were randomly integrated in the genome of the *parp3* mutant. The band around 3.5 kb indicated a random T-DNA integration as a tandem repeat.

Another protein that seems important for b-NHEJ is XRCC1 (Charbonnel *et al.* 2010). The homozygous mutant was isolated and characterized. The results showed that two T-DNAs were inserted in intron 5 in inverted orientation, with the LBs flanking the plant DNA and 3 bp of the plant DNA missing as well as the LB sequence of one of the T-DNAs (**Figure 1B**). For Southern blotting genomic DNA was again digested with *HindIII*. Two bands of 2173 bp and 3343 bp were expected diagnostic for the connection between the T-DNAs and the *XRCC1* gene. These can indeed be seen on the gel (**Figure 1C**). Besides, extra bands were detected, indicating that additional T-DNAs were randomly integrated in the genome of the *xrcc1* mutant.

DNA damage response

In order to study whether Parp3 functions in DNA repair, the *parp3* mutant was tested for sensitivity to the genotoxic agents bleomycin and MMS. As no difference was seen with the wildtype, *p1p3* double and *p1p2p3* triple mutants were obtained by crossing the individual mutants and assayed in the same way for hypersensitivity to genotoxic agents as this may reveal gene redundancy. Without any treatments, *p1p2* (Jia *et al.* 2013), *p1p3* double and *p1p2p3* triple mutants grew the same as wild type. When treated with bleomycin, there were again no clear differences in growth seen between the *parp* mutants and wild type plants (data not shown). However, as can be seen in **Figure 2A**, when treated with MMS, the *p1p2p3* triple and *p1p3* double mutants showed somewhat more sensitivity than each of the single mutants, especially when treated with 0.005% MMS. This result suggests that Parp3 is indeed involved in DNA repair and even may play a more important role than Parp2 in the repair of MMS damage.

The *xrcc1* mutant was treated in a similar way to test its function in DNA repair (**Figure 2B**). As no difference was observed with the wild type, a *ku80xrcc1* double mutant was

obtained by crossing the single mutants and assayed in the same way for hypersensitivity to genotoxic agents (**Figure 2B**). In normal growth conditions, the *ku80xrcc1* double mutant grew the same as the wild type and single mutants. When treated with bleomycin, the *ku80xrcc1* double mutant showed the same sensitivity as the *ku80* single mutant. However, the *ku80xrcc1* mutant showed more sensitivity to MMS than the wildtype and each of the single mutants. These results indicate that XRCC1 and Ku80 repair MMS induced damage by different pathways.

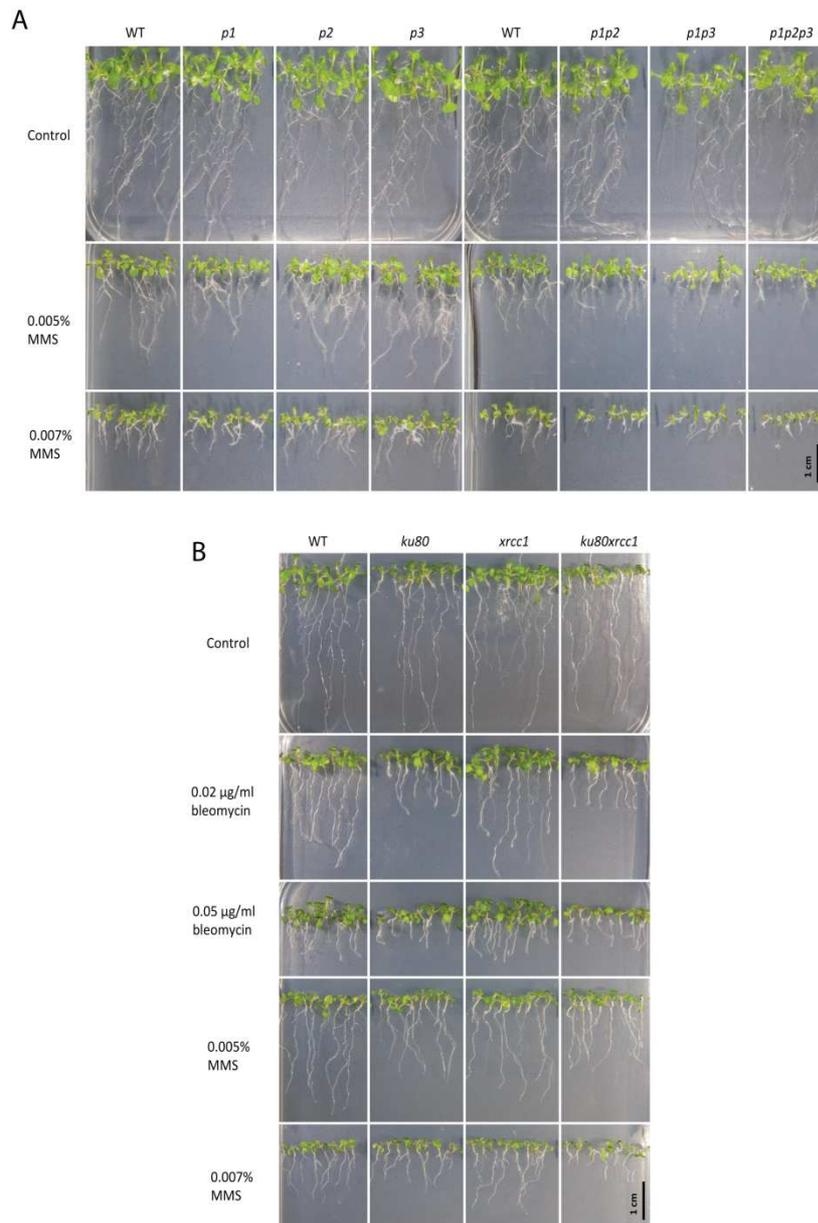


Figure 2. NHEJ mutants response to DNA damaging treatments. (A) Phenotypes of wild-type plants and *parp1*, *parp2*, *parp3*, *p1p2*, *p1p3* and *p1p2p3* mutants germinated on ½ MS medium (control) or ½ MS medium containing 0.005% and 0.007% MMS photograph taken 2 weeks after germination. (B) Phenotypes of wild-type plants and *ku80*, *xrcc1* and *ku80xrcc1* mutants germinated on ½ MS medium (control) or ½ MS medium containing 0.02 µg/ml, 0.05 µg/ml bleomycin or 0.005%, 0.007% MMS photograph taken 2 weeks after germination.

Root transformation

In order to determine the effects of the mutations on *Agrobacterium* T-DNA integration in somatic plant cells, we performed a root transformation assay using *Agrobacterium* strain LBA1100 harboring the binary vector (pCambia3301) and selected for callus formation in the presence of phosphinothricin (ppt). Roots of WT, *ku80*, *ku70*, *parp1*, *parp2*, *parp3*, *xrcc1*, *lig4*, *lig6*, *ku80xrcc1*, *lig4lig6*, *p1p2*, *p1p2p3* and *ku80p1p2* mutant lines were co-cultivated with the *Agrobacterium* strain, and numbers of green calli were counted after 4 weeks. Longer culture led to green shoot formation from these calli. Data from three independent tests and more than 300 root segments of each genotype are presented in **Figure 3**. Roots from *ku80*, *ku70*, *parp1*, *parp2*, *parp3*, *xrcc1*, *lig6*, *p1p2* and *p1p2p3* mutants were transformed as well as the wild type. The double mutant *ku80xrcc1* and triple mutant *ku80p1p2*, which were supposed to be disturbed in both c-NHEJ and b-NHEJ repair pathways, gave significantly less transformed calli than the wild type roots, indicating that NHEJ repair pathways are partly responsible for the T-DNA integration process. Interestingly, roots from the *lig4* mutant produced significantly more transformed calli than roots from wild type plants, indicating that Lig4 is not required for *Agrobacterium* T-DNA integration in plants and may even be somewhat inhibitory to the T-DNA integration process. However, the roots of the *lig4lig6* double mutant produced the same number of transformed calli as roots from wild-type plants.

Transient transformation is not affected in NHEJ mutants

We expected that the mutation of NHEJ genes would not affect the entry of T-DNA into the plant cell nucleus. To test this, we infected root segments of wild type and *Arabidopsis* mutant lines with the same *Agrobacterium* strain LBA1100 harboring the binary vector pCambia3301 which contains a CaMV 35S promoter *gusA*-intron gene. This *gusA* gene allows expression of GUS activity in plant cells, but not bacteria. After inoculation with *Agrobacterium* and co-cultivation for 3 days, the root segments were stained with X-Gluc to reveal transient transformation. As shown in **Figure 4A**, most root segments were stained with blue color, but this was difficult to quantify. In order to quantify the GUS activity, we used the fluorescent β -glucuronidase substrate MUG (4-methylumbelliferyl- β -D-galactopyranoside) instead. Three days after infection, proteins were extracted from root segments and tested in this way for GUS activity. As seen in **Figure 4B**, no significant differences were observed in the GUS activity between wild type and NHEJ mutants roots co-cultivated with *Agrobacterium*. Thus, NHEJ mutations that increased or reduced stable transformation frequencies did not affect T-DNA transfer and transient expression of the T-DNA.

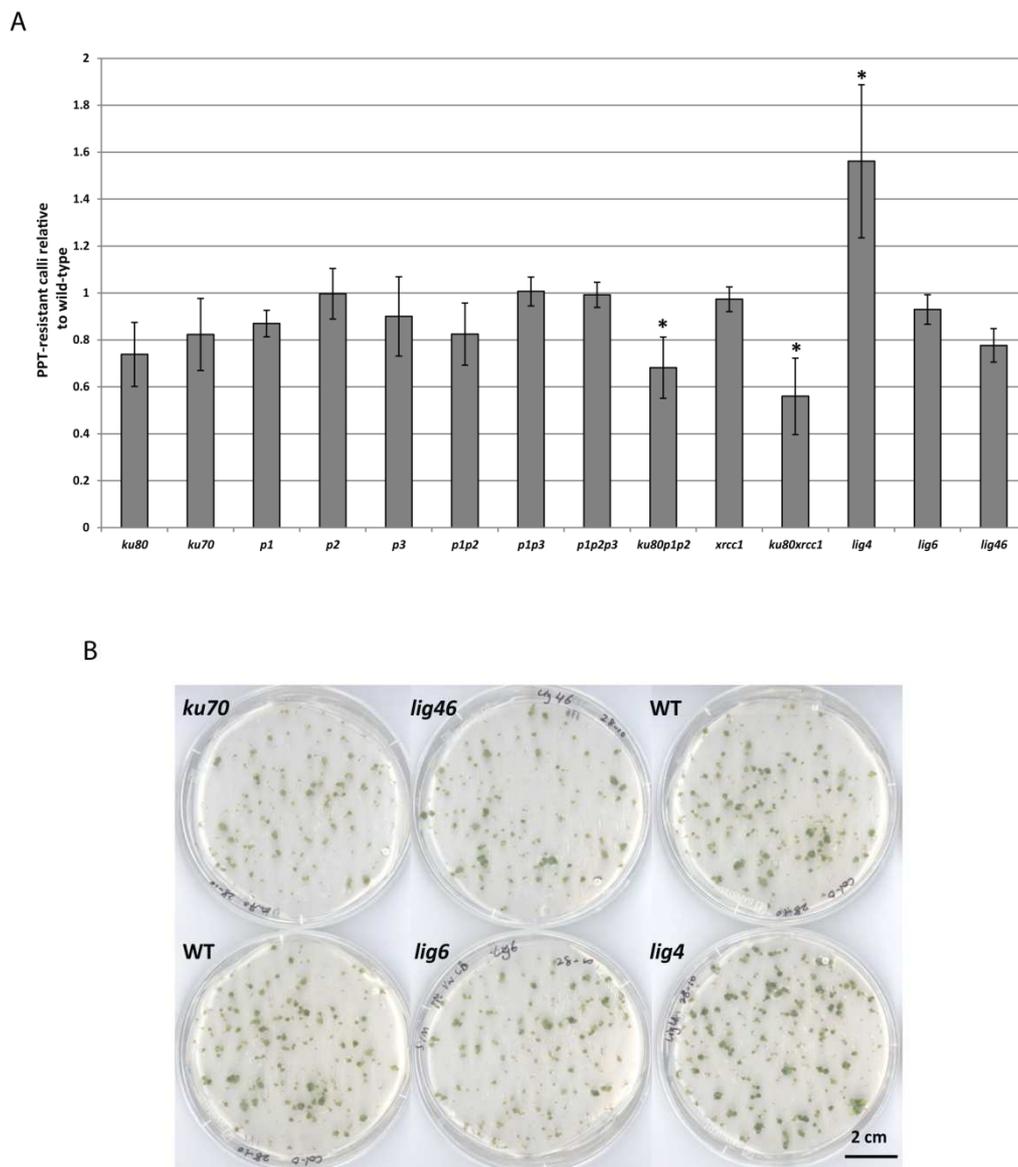


Figure 3. Stable root transformation of *Arabidopsis* NHEJ mutants. **(A)** Root segments from wild-type and mutant plants were co-cultivated with *Agrobacterium* strain LBA1100 harboring pCambia3301 for 48 hours, and transferred to selection induction medium. Efficiency of transformation as represented by the ratio of mean number of green calli per root segment relative to wild type. Asterisk indicates a significant difference relative to wild-type plants (ANOVA, $P < 0.05$). **(B)** Examples of root transformation assays from wild-type and *ku80*, *lig4*, *lig6*, *lig46* mutants. Photographs were taken 4 weeks after co-cultivation.

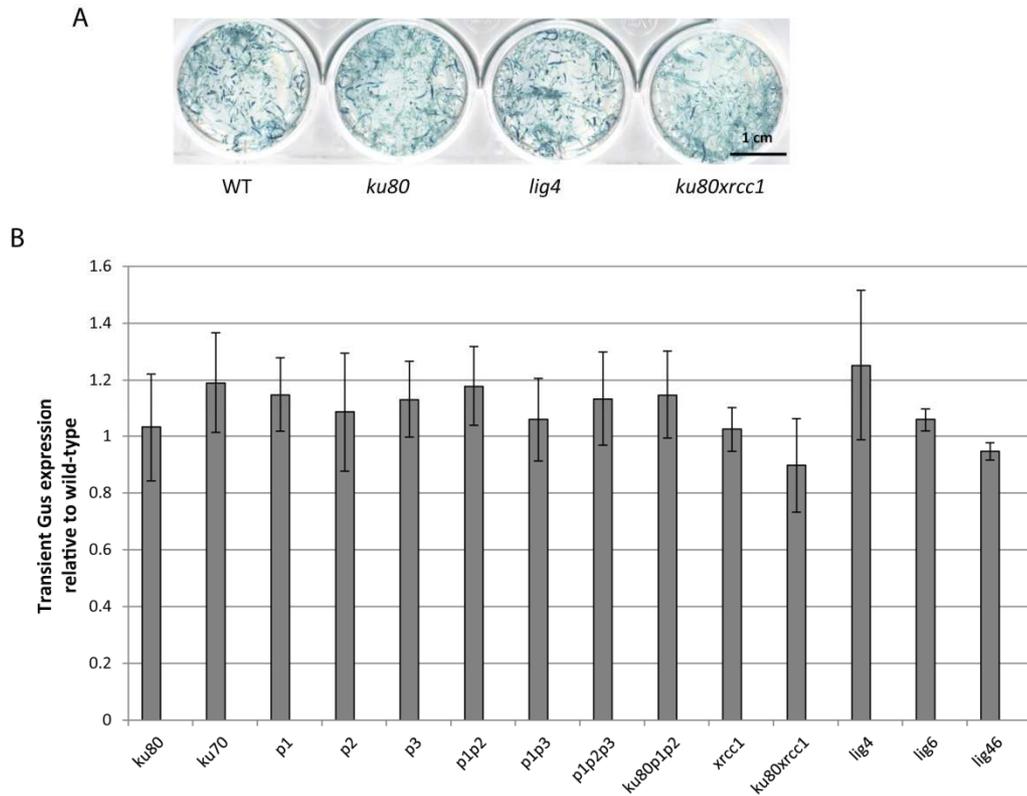


Figure 4. Transient GUS assay of Arabidopsis NHEJ mutants. **(A)** Root segments from wild-type and mutant plants were co-cultivated with *Agrobacterium* strain LBA1100 harboring pCambia3301 for 72 hours, and then stained with X-Gluc. **(B)** The transient GUS expression level is quantified by MUG assay and presented as the ratio of the level of MUG activity relative to wild type. Statistics analysis showed no significant differences.

Discussion

Our previous work showed that NHEJ can still occur in *parp1parp2* and *ku80parp1parp2* mutants (Jia *et al.* 2013), suggesting that there may be a third pathway of NHEJ in plants or that there may still be functionally redundant proteins taking over the function of Parp1 and Parp2 when they are absent. The recently discovered *PARP3* gene might represent such a redundant protein. The mammalian *PARP3* gene has been reported to be involved in the DNA damage response and to interact with different partners belonging to the c-NHEJ pathway. Therefore, a T-DNA insertion mutant of *PARP3* was obtained and characterized. There was no phenotypical difference between the *parp* single mutants and the wild-type plants under normal growth conditions or after genotoxic treatment. However, the *p1p3* double and *p1p2p3* triple mutants showed more sensitivity to MMS. It means that Parp3 is indeed involved in DNA repair in plants. Further work is needed to find out to what extent Parp1, Parp2 and Parp3 are functionally redundant in the same NHEJ repair pathway in plants. Recently, unexpectedly Parp1 was shown to be involved not only in b-NHEJ (Beck, Robert, *et al.* 2014), but also in c-NHEJ in mammalian cells as a recruitment factor for the chromatin remodeler CHD2 (Luijsterburg *et al.* 2016).

A

	AGL1/pSDM3834								AGL1/pSDM3900			
	<i>Col-o</i>	<i>ku80</i>	<i>parp1</i>	<i>xrcc1</i>	<i>ku80parp1</i>	<i>Ws</i>	<i>mre11-2(Ws)</i>	<i>ku70(Ws)</i>	<i>Col-o</i>	<i>ku80</i>	<i>parp1</i>	<i>xrcc1</i>
Mean	10.2	2.2	10.8	13.8	4	20.6	4.4	5	30.3	5.5	28.3	41.9
S.D	4.02	1.47	3.76	3.43	2.28	3.44	0.80	2.68	6.18	1.91	3.49	8.98
N.	5	5	5	5	5	5	5	5	10	10	10	10

B

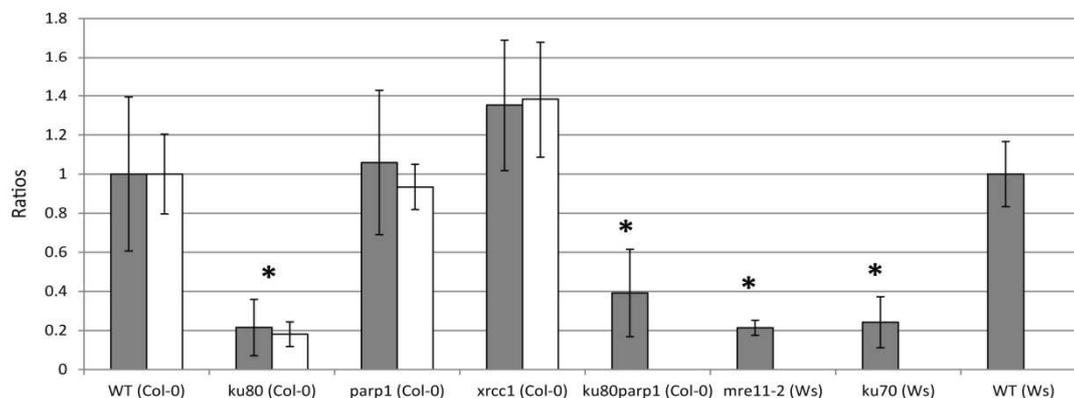


Figure 5. Transformation frequencies using floral dip assay. One or half gram of seeds from the wild-type (*Col-0* and *Ws*) and the NHEJ mutants obtained after floral dip transformations were selected on hygromycin (for pSDM3834) or phosphinothricin (for pSDM3900). The number of selection-marker-resistant seedlings was scored 2 weeks after germination. Plates with contamination were excluded, and mean numbers of resistant seedling (per plate), standard errors and numbers of measurements (N) are shown in (A). (B) The transformation efficiency is presented as the ratio of the number of selection-marker-resistant seedlings in the mutants and wild-type. The grey bar indicates the data for pSDM3834, and the white bar represents pSDM3900. Asterisk indicates a significant difference relative to wild-type plants (ANOVA, $P < 0.001$).

Table 1. Sequences of primers used for characterization of T-DNA insertion lines.

Name	Locus	Sequence
LBa1	T-DNA left end	5'-TGGTTCACGTAGTGGGCCATCG-3'
RB	T-DNA right end	5'-TTTGGAACCTGACAGAACCGC-3'
Sp170	<i>Atxrcc1</i>	5'-GACACTCTAAAGAAACGTTCC-3'
Sp171	<i>Atxrcc1</i>	5'-GAATCTCCGTTTAAACCATCC-3'
Sp271	pROK2	5'-CCCGTGTCTCTCCAAATG-3'
Sp272	pROK2	5'-CAGGTCCCCAGATTAGCC-3'
Sp539	<i>Atparp3</i>	5'-GTGAGTGGTGCAGTTGCGTGT-3'
Sp544	<i>Atparp3</i>	5'-CTTCGGCATTAGGGTCATCTC-3'

In this work also a mutant with a T-DNA insertion in the *XRCC1* gene was isolated and characterized. The Arabidopsis *xrcc1* mutant has been shown to be hypersensitive to γ -radiation and even more sensitive than *ku80* mutant plants (Charbonnel *et al.* 2010). However, we saw no phenotypical difference between *xrcc1* mutant and wild-type plants after treated with MMS or bleomycin agents. Double mutant *ku80xrcc1* plants showed similar or increased sensitivity compared to the *ku80* single mutant to a number of different genotoxic agents, confirming that XRCC1 is involved in DNA repair pathways in plants, but the precise role in the NHEJ pathways remain to be determined. In human cells, XRCC1 has been shown to play an important role together with DNA Lig3 in b-NHEJ pathways (Della-Maria *et al.* 2011). Due to lack of Lig3 in plants, XRCC1 must act in a different manner during the b-NHEJ repair in plants, which may depend on poly (ADP-ribose) synthesis (Breslin *et al.* 2015).

Agrobacterium T-DNA molecules integrate into plant DNA double strand break sites (Tzfira *et al.* 2004). Thus, double strand break repair mechanisms are hypothesized to be involved in the integration of *Agrobacterium* T-DNA in plants. Previously, our group had shown that non-homologous T-DNA is integrated by NHEJ in yeast, and that NHEJ proteins including the Ku70/Ku80 and Lig4 play an essential role in T-DNA integration in yeast. Several studies have investigated the role of NHEJ in *Agrobacterium*-mediated T-DNA integration in plants, but the results obtained so far are variable (Friesner and Britt 2003; van Attikum *et al.* 2003; Gallego *et al.* 2003; Li *et al.* 2005; Jia *et al.* 2012; Nishizawa-Yokoi *et al.* 2012; Mestiri *et al.* 2014; Park *et al.* 2015). The inconsistency of these findings is probably due to using different mutant lines or different experimental procedures. In order to test whether NHEJ proteins are involved in T-DNA integration, we performed root transformation assays. Although a lower T-DNA integration frequency was observed in the *ku80p1p2* and *ku80xrcc1* mutants, these mutants were still able to stably integrate T-DNA. In addition the *ku80p1p2* mutant can still repair nuclease-induced DSBs (**Chapter 5**) indicating that either another redundant protein is present or other repair pathways are active.

The root transformation frequency in *ku80* and *ku70* c-NHEJ mutants did not show significant differences compared to the wild type. However, our previous results from floral dip transformations with NHEJ mutants showed that the transformation frequency is highly reduced in *ku80*, *ku70* and *mre11-2* mutants (**Figure 5**). This may be due to differences of target tissues. Floral dip results with *parp1*, *xrcc1* and *ku80parp1* mutants showed that PARP1 and XRCC1 has no essential function in floral dip transformation (**Figure 5**). Roots and other somatic cells are the natural target for *Agrobacterium*-mediated transformation, whereas the target cells in floral dip transformation are the female gametophytes.

Our results of root transformation showed that mutations in either the c-NHEJ or b-NHEJ pathway did not significantly impact T-DNA integration, but mutations in both pathways together significantly reduced root transformation efficiency. Other reports similarly showed that T-DNA integration was not completely suppressed or not decreased at all in *ku80parp1* and *ku80xrcc1* double mutants (Mestiri *et al.* 2014; Park *et al.* 2015). It could be that c-NHEJ and b-NHEJ are functionally redundant in T-DNA integration. Mestiri *et al.* (2014) observed an about three-fold decrease in both floral dip transformation and root transformation frequencies of single mutants in the b-NHEJ pathway. However, the results

from Park *et al.* (2015) indicated that c-NHEJ and b-NHEJ proteins do not positively contribute to transformation susceptibility and may even limit stable transformation. One possibility is that different experimental conditions (such as *Agrobacterium* inoculation concentrations) caused different observations. Besides, *Arabidopsis* mutant lines may show altered growth or developmental characteristics, which may also affect the outcomes of transformations.

Although conflicting results were obtained from several research groups including our own investigating the role of NHEJ proteins in *Agrobacterium*-mediated plant transformation, all these findings show at least that disruption of one or multiple NHEJ repair pathways does not eliminate transformation, suggesting that another DNA repair pathway is involved in T-DNA integration. A recent study showed that the *Arabidopsis* Pol θ ortholog *Tebichi* (*Teb*) is essential for T-DNA integration (Kregten *et al.*, 2016). Since mutations in c-NHEJ components did show decreased transformation frequencies, they probably function together with Pol θ during the integration process.

Acknowledgements

This work was supported by a grant from the China Scholarship Council (CSC grant no. 2011699040).

References

- Alonso, J. M., M. Alonso, A. N. Stepanova, T. J. Leisse, D. K. Stevenson *et al.*, 2003 Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science*. 301: 653–657.
- van Attikum, H., P. Bundock, and P. J. Hooykaas, 2001 Non-homologous end-joining proteins are required for *Agrobacterium* T-DNA integration. *EMBO J*. 20: 6550–6558.
- van Attikum, H., P. Bundock, R. M. Overmeer, L. Y. Lee, S. B. Gelvin *et al.*, 2003 The *Arabidopsis* AtLIG4 gene is required for the repair of DNA damage, but not for the integration of *Agrobacterium* T-DNA. *Nucleic Acids Res*. 31: 4247–4255.
- van Attikum, H., and P. J. Hooykaas, 2003 Genetic requirements for the targeted integration of *Agrobacterium* T-DNA in *Saccharomyces cerevisiae*. *Nucleic Acids Res*. 31: 826–832.
- Audebert, M., B. Salles, and P. Calsou, 2004 Involvement of poly(ADP-ribose) polymerase-1 and XRCC1/DNA ligase III in an alternative route for DNA double-strand breaks rejoining. *J. Biol. Chem*. 279: 55117–55126.
- Beck, C., C. Boehler, J. Guirouilh Barbat, M. E. Bonnet, G. Illuzzi *et al.*, 2014 PARP3 affects the relative contribution of homologous recombination and nonhomologous end-joining pathways. *Nucleic Acids Res*. 42: 5616–5632.
- Beck, C., I. Robert, B. Reina-San-Martin, V. Schreiber, and F. Dantzer, 2014 Poly(ADP-ribose) polymerases in double-strand break repair: Focus on PARP1, PARP2 and PARP3. *Exp. Cell Res*. 329: 18–25.
- Boehler, C., L. R. Gauthier, O. Mortusewicz, D. S. Biard, J.-M. Saliou *et al.*, 2011 Poly(ADP-ribose) polymerase 3 (PARP3), a newcomer in cellular response to DNA damage and mitotic progression. *Proc. Natl. Acad. Sci. USA*. 108: 2783–2788.

- Breslin, C., P. Hornyak, A. Ridley, S. L. Rulten, H. Hanzlikova et al., 2015 The XRCC1 phosphate-binding pocket binds poly (ADP-ribose) and is required for XRCC1 function. *Nucleic Acids Res.* 43: 6934–6944.
- Bundock, P., and P. Hooykaas, 2002 Severe developmental defects, hypersensitivity to DNA-damaging agents, and lengthened telomeres in *Arabidopsis* MRE11 mutants. *Plant Cell* 14: 2451–2462.
- Charbonnel, C., M. E. Gallego, and C. I. White, 2010 Xrcc1-dependent and Ku-dependent DNA double-strand break repair kinetics in *Arabidopsis* plants. *Plant J.* 64: 280–290.
- Cheng, Q., N. Barboule, P. Frit, D. Gomez, O. Bombarde et al., 2011 Ku counteracts mobilization of PARP1 and MRN in chromatin damaged with DNA double-strand breaks. *Nucleic Acids Res.* 39: 9605–9619.
- Chilton, M.-D. M., and Q. Que, 2003 Targeted integration of T-DNA into the tobacco genome at double-stranded breaks: new insights on the mechanism of T-DNA integration. *Plant Physiol.* 133: 956–965.
- Della-Maria, J., Y. Zhou, M. S. Tsai, J. Kuhnlein, J. P. Carney et al., 2011 Human Mre11/human Rad50/Nbs1 and DNA ligase III α /XRCC1 protein complexes act together in an alternative nonhomologous end joining pathway. *J. Biol. Chem.* 286: 33845–33853.
- Friesner, J., and A. B. Britt, 2003 Ku80- and DNA ligase IV-deficient plants are sensitive to ionizing radiation and defective in T-DNA integration. *Plant J.* 34: 427–440.
- Gallego, M. E., J.-Y. Bleuyard, S. Daoudal-Cotterell, N. Jallut, and C. I. White, 2003 Ku80 plays a role in non-homologous recombination but is not required for T-DNA integration in *Arabidopsis*. *Plant J.* 35: 557–565.
- Gelvin, S. B., 2010 Plant proteins involved in *Agrobacterium*-mediated genetic transformation. *Annu. Rev. Phytopathol.* 48: 45–68.
- Gheysen, G., R. Villarroel, and M. Van Montagu, 1991 Illegitimate recombination in plants: A model for T-DNA integration. *Genes Dev.* 5: 287–297.
- Gibson, B. A., and W. L. Kraus, 2012 New insights into the molecular and cellular functions of poly(ADP-ribose) and PARPs. *Nat. Rev. Mol. Cell Biol.* 13: 411–424.
- Jia, Q., P. Bundock, P. J. J. Hooykaas, and S. de Pater, 2012 *Agrobacterium tumefaciens* T-DNA integration and gene targeting in *Arabidopsis thaliana* non-homologous end-joining mutants. *J. Bot.* 2012: 1–13.
- Jia, Q., A. Den Dulk-Ras, H. Shen, P. J. J. Hooykaas, and S. de Pater, 2013 Poly(ADP-ribose) polymerases are involved in microhomology mediated back-up non-homologous end joining in *Arabidopsis thaliana*. *Plant Mol. Biol.* 82: 339–351.
- Jia, Q., 2011 DNA repair and gene targeting in plant end-joining mutants. PhD thesis. Leiden University, Leiden, The Netherlands.
- Kooistra, R., P. J. J. Hooykaas, and H. . Y. Steensma, 2004 Efficient gene targeting in *Kluyveromyces lactis*. *Yeast* 21: 781–792.
- van Kregten M., S. de Pater, R. Romeijn, R. van Schendel, P. Hooykaas and M. Tijsterman, 2016 T-DNA integration in plants results from Polymerase Theta-mediated DNA repair. *Nature plants* 2:1-6.
- Li, J., M. Vaidya, C. White, A. Vainstein, V. Citovsky et al., 2005 Involvement of KU80 in T-DNA integration in plant cells. *Proc. Natl. Acad. Sci. U. S. A.* 102: 19231–19236.

- Luijsterburg, M. S., I. de Krijger, W. W. Wiegant, R. G. Shah, G. Smeenk et al., 2016 PARP1 links CHD2-mediated chromatin expansion and H3.3 deposition to DNA repair by non-homologous end-joining. *Mol. Cell* 61: 547–562.
- Mayerhofer, R., Z. Koncz-kalman, C. Nawrath, G. Bakkeren, A. Cramer et al., 1991 T-DNA integration : in plants mode of illegitimate recombination. *EMBO J.* 10: 697–704.
- Mestiri, I., F. Norre, M. E. Gallego, and C. I. White, 2014 Multiple host-cell recombination pathways act in *Agrobacterium*-mediated transformation of plant cells. *Plant J.* 77: 511–520.
- Nishizawa-Yokoi, A., S. Nonaka, H. Saika, Y. I. Kwon, K. Osakabe et al., 2012 Suppression of Ku70/80 or Lig4 leads to decreased stable transformation and enhanced homologous recombination in rice. *New Phytol.* 196: 1048–1059.
- Park, S.-Y., Z. Vaghchhipawala, B. Vasudevan, L.-Y. Lee, Y. Shen et al., 2015 *Agrobacterium* T-DNA integration into the plant genome can occur without the activity of key non-homologous end-joining proteins. *Plant J.* 81: 934–946.
- Rissel, D., J. Losch, and E. Peiter, 2014 The nuclear protein Poly(ADP-ribose) polymerase 3 (AtPARP3) is required for seed storability in *Arabidopsis thaliana*. *Plant Biol.* 16: 1058–1064.
- Rodenburg, K. W., M. J. de Groot, R. a Schilperoort, and P. J. Hooykaas, 1989 Single-stranded DNA used as an efficient new vehicle for transformation of plant protoplasts. *Plant Mol. Biol.* 13: 711–719.
- Salomon, S., and H. Puchta, 1998 Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. *EMBO J.* 17: 6086–6095.
- Schreiber, V., F. Dantzer, J.-C. Ame, and G. de Murcia, 2006 Poly(ADP-ribose): novel functions for an old molecule. *Nat. Rev. Mol. Cell Biol.* 7: 517–528.
- Stolarek, M., D. Gruszka, A. Braszewska-Zalewska, and M. Maluszynski, 2015 Alleles of newly identified barley gene HvPARP3 exhibit changes in efficiency of DNA repair. *DNA Repair* 28: 116–130.
- Tzfira, T., L. R. Frankman, M. Vaidya, and V. Citovsky, 2003 Site-specific integration of *Agrobacterium tumefaciens* T-DNA via double-stranded intermediates. *Plant Physiol.* 133: 1011–1023.
- Tzfira, T., J. Li, B. Lacroix, and V. Citovsky, 2004 *Agrobacterium* T-DNA integration: molecules and models. *Trends Genet.* 20: 375–383.
- Vergunst, A. C., B. Schrammeijer, A. den Dulk-Ras, C. M. T. de Vlaam, T. J. G. Regensburg-Tuïnk et al., 2000 VirB/D4-dependent protein translocation from *Agrobacterium* into plant cells. *Sci.* 290: 979–982.
- Weijers, D., M. Franke-van Dijk, R. J. Vencken, A. Quint, P. Hooykaas et al., 2001 An *Arabidopsis* minute-like phenotype caused by a semi-dominant mutation in a ribosomal protein S5 gene. *Development* 128: 4289–4299.

Chapter 3

Characterization of an *Arabidopsis* gene encoding a putative DNA ligase

Hexi Shen, Qi Jia, Paul J. J. Hooykaas, Sylvia de Pater

Department of Molecular and Developmental Genetics, Institute of Biology, Leiden University,
Leiden, 2333 BE, The Netherlands

Abstract

During DNA replication and DNA repair, broken DNA ends are ligated together by an ATP-dependent DNA ligase. In eukaryotic cells several DNA ligases are present. DNA ligase I has a function in DNA replication and various DNA repair pathways, whereas DNA ligase 4 is involved in non-homologous end-joining (NHEJ) and the plant specific ligase 6 functions only during seed germination. In an *in silico* search for homologs of ATP-dependent DNA ligases in Arabidopsis, we found another candidate gene At1g49250 (Lig1a). A homozygous mutant of *Atlig1a* was isolated and crossed with the *Atlig4* and *Atlig4lig6* mutants to obtain the double and triple mutants (*Atlig4lig1a* and *Atlig4lig6lig1a*). Under both normal growth conditions and under genotoxic stress, the *Atlig1a* mutant behaved like the wild type, but an additional phenotype was observed in *Atlig4lig1a*, *Atlig4lig6lig1a* mutants compared to *Atlig4* and *Atlig4lig6*. This suggests that Ligase 1a does play a role in break repair, when Ligase 4 and Ligase 6 are mutated.

Introduction

DNA ligases are very important enzymes of DNA metabolism. They have indispensable functions in DNA replication and DNA repair pathways, where DNA ligases are required for repair by both Homologous Recombination (HR) and Non-Homologous End-Joining (NHEJ) repair pathways to seal the broken ends. Eukaryotes possess several ATP-dependent DNA ligases, which are essential for the replication and repair of the nuclear and organelle genomes. So far, four different DNA ligases, termed LIG1, LIG3, LIG4 and LIG6 have been identified and functionally analysed in eukaryotes (Timson *et al.* 2000). LIG1 and LIG4 are conserved in all eukaryotes, while LIG3 is unique to vertebrates and LIG6 is specific for plants.

DNA LIG1 is an essential enzyme, responsible both for joining Okazaki fragments during DNA replication and for ligating the single-strand nicks formed during excision repair, and also functions in nucleotide excision repair and homologous recombination (Barnes *et al.* 1992; Levin *et al.* 2000; MacNeill 2001; Arakawa and Iliakis 2015). There is functional redundancy between LIG1 and LIG3 and they both might function in backup non-homologous end-joining (b-NHEJ) (Simsek *et al.* 2011; Arakawa *et al.* 2012; Paul *et al.* 2013). The vertebrate-specific LIG3 has two splice variants, LIG3 α and LIG3 β . LIG3 α has a function in single-strand break repair (SSBR) and base-excision repair (BER) repair pathways. Besides the DNA binding domain and catalytic domain present in all ligases, it has a N-terminal zinc finger domain and a C-terminal BRCT domain, which can also be found in LIG4. The BRCT domain is necessary for interaction with XRCC1, which is involved in BER/SSBR (Taylor *et al.* 1998). The N-terminal zinc finger domain is necessary for ligation activity and is thought to have a function in targeting LIG3 to nicks in DNA (Taylor *et al.* 1998; Mackey *et al.* 1999). Ligase 3 β is expressed only in the testis and may have a role in meiotic recombination (Chen *et al.* 1995; Mackey *et al.* 1997). Human DNA ligase 3 also encodes a mitochondrial form, probably produced by translation initiation at an upstream translation start site, producing a protein with a mitochondrial targeting sequence (Lakshmiathy and Campbell 1999). LIG6 is a DNA ligase, with a domain structure unique to plant species and different from LIG1, LIG3 and LIG4, and seems to function only during seed germination to promote seed storability and rapid germination (Waterworth *et al.* 2010).

LIG4 is conserved in yeast, fungi, mammals and plants and plays an important role in the classical non-homologous end-joining (c-NHEJ) pathway of DSB repair (Critchlow *et al.* 1997; Wilson *et al.* 1997; van Attikum *et al.* 2003), which is independent of homologous sequences and is the predominant mechanism for DSB repair in mammals and plants. During c-NHEJ DNA ends are directly joined, which is mediated by several protein complexes, including the LIG4/XRCC4/XLF complex (Ahnesorg *et al.* 2006). In the absence of LIG4, cells are more sensitive to genotoxic treatments such as ionizing radiation. Increasing evidence suggests that there is a backup NHEJ (b-NHEJ) repair pathway, which is activated when c-NHEJ is blocked, and LIG3 may play an important role in b-NHEJ in animal cells. However, a low level of LIG3 or LIG4 is sufficient for efficient NHEJ in mammalian cells (Windhofer *et al.* 2007). In plants *lig4lig6* mutants are still able to perform end-joining (Jia, 2011), suggesting that there must still be another ligase to take over their function. In order to find other putative ligases in the Arabidopsis genome, we analyzed the genome sequence for the

presence of sequences that shared homology to the DNA *lig1*, *lig4* and *lig6*. In this way we found a putative DNA ligase gene, *Atlig1a* (At1g49250). In order to analyze its role in DNA repair, we characterized a T-DNA insertion mutant in the At1g49250 gene and homozygous *Atlig1a*, *Atlig4*, *Atlig4lig1a* and *Atlig4lig6lig1a* mutants were tested for sensitivity to DNA damaging agents and DNA repair.

Materials and methods

Plant materials

The At1g49250 T-DNA insertion line was obtained from the SALK T-DNA collection (SALK_026361) (Alonso *et al.* 2003). Information is available at <http://signal.salk.edu/cgi-bin/tdnaexpress>. The homozygous mutant was crossed with *Atlig4* (At1g16970; SALK_123114) and *Atlig4lig6* mutants (At1g66730; SALK_065307) previously isolated in our lab (Jia, 2011) and the *Atlig4lig1a* double mutant and the *Atlig4lig6lig1a* triple mutant were obtained.

Characterization of the Atlig1a mutant

DNA was extracted from individual plants using the CTAB DNA isolation protocol (de Pater *et al.* 2006). The T-DNA insertion site of the mutant was mapped with a gene-specific primer (Sp416 or Sp417) and a T-DNA specific primer (LBa1 or RB) and PCR products were cloned and sequenced. Pairs of gene-specific primers around the insertion site were used to determine whether the plants were homozygous or heterozygous for the T-DNA insertion. The sequences of all the primers are listed in Table 1. For Southern blot analysis, DNA from the *Atlig1a* mutant was digested with *Hind*III. DNA (10 µg) was ran on a 0.7% agarose gel and transferred onto a positively charged Hybond-N membrane (Amersham Biosciences). The hybridization and detection procedures were done according to the DIG protocol from Roche Applied Sciences.

Assays for sensitivity to bleomycin and methyl methane sulfonate (MMS)

Seeds of wild-type plants and *Atlig4*, *Atlig6*, *Atlig1a*, *Atlig4lig1a*, and *Atlig4lig6lig1a* mutants were surface-sterilized as described (Weijers *et al.* 2001) and germinated on solid ½ MS medium without additions or containing 0.02 µg/ml bleomycin (Sigma), 0.05 µg/ml bleomycin, 0.005 % (v/v) MMS (Sigma) or 0.008 % MMS and photographed after 1 week.

Comet assay

DSBs were detected in plant cells by a neutral comet assay (N/N protocol) as described before (Menke *et al.* 2001), but with minor modifications. Two-week-old seedlings were submerged in liquid ½ MS containing 1 µg/ml bleomycin for 24h. These seedlings were subsequently recovered and submerged in liquid ½ MS for 2h, 6h and 24h. Plant nuclei were isolated and embedded in 1% low melting point Ultrapure™ agarose-1000 (Invitrogen) to make a mini gel on microscopic slides according to the protocol. Nuclei were subjected to lysis in high salt (2.5 M NaCl, 10 mM Tris-HCl, pH 7.5, 100 mM EDTA) for 20 min at room temperature (N/N protocol). Equilibration for 3 times 5 min in 1X TBE buffer (90 mM Tris-borate, 2 mM EDTA, pH 8.4) on ice was followed by electrophoresis at 4°C in TBE buffer for 15 min at 30 V (1 V/cm), 15-17 mA. Dry agarose gels were stained with 15 µl ethidium bromide (5 µg/ml)

and evaluated with a microscope using the DsRed channel (excitation at ~510 nm, emission at ~595 nm). Images of comets were captured at a 40-fold magnification by a digital camera. Comet scoring software CometScore™ (Tritek corporation) was used to analyze the comets. The percentage of tail DNA (%DNA in tail) was used as a measure of DNA damage. At least 100 comets from 6 independent gel replicas were analyzed for each treatment.

Floral dip transformation and GUS staining

The *Atlig1a* promoter sequence was obtained by PCR using Hp1 and Hp2 primers (Table 1). The *Atlig1a::gus* reporter plasmid was constructed by cloning the 1.271 kb *Atlig1a* promoter upstream of the GUS coding sequence in pCambia1391xa and subsequently introduced in *Agrobacterium* strain AGL1 by electroporation. Floral dip (Clough and Bent 1998) was performed to obtain *Atlig1a::gus* reporter plants. Transformants were selected on solid MA medium without sucrose containing 15 µg/ml hygromycin, 100 µg/ml timentin (to kill *Agrobacterium* cells) and 100 µg/ml nystatin (to prevent growth of fungi).

Two weeks old T2 seedlings were incubated in ½ MS or ½ MS supplemented with 0.5 µg/ml bleomycin for 24h, followed by GUS staining. The GUS staining was examined under a Leica MZ12 microscope (Leica microsystems) and photographed with a Leica DC 500 digital camera (Leica microsystems).

Results

Characterization of the Atlig1a mutant

By *in silico* analysis of the *Arabidopsis thaliana* genome, we found a novel gene encoding a putative DNA ligase. As can be seen in **Figure 1**, the encoded protein has all the essential domains of DNA ligases and shares a high homology with *Arabidopsis* ligase 1 (73% of homology). In order to study the function of AtLIG1a, a T-DNA insertion mutant from the Salk collection was analyzed. The homozygous T-DNA insertion mutant was identified by PCR analysis. When two gene-specific primers flanking the insertion site were used, PCR products were amplified for the wild type and heterozygous mutants. However, no PCR product was obtained from homozygous mutants as the PCR products would be more than 10 kb in size due to the T-DNA insertion and would not be detected by the PCR conditions used here. When T-DNA-specific left border primer (LBa1) was used in combination with the sense gene-specific primer (SP416), a PCR product was amplified from the T-DNA insertion mutants. No PCR products were obtained using LBa1 or the right border (RB) primer in combination with the antisense gene-specific primer (SP417). The T-DNA right border primer set (SP223 and SP224) was used to detect the T-DNA RB and putative flanking vector sequences. A PCR product was amplified from the T-DNA insertion mutants, indicating that the T-DNA indeed had been integrated with extra vector sequences. The insertion point was mapped by sequencing of the LBa1-SP416 PCR product. The results indicated that the T-DNA was inserted in exon 3 and had 7 bp filler DNA at the left border end. The RB was linked to vector pBin-ROK DNA and therefore, the RB integration site could not be recovered by PCR with the RB primer. The structure of the T-DNA insertion in *Atlig1a* is shown in **Figure 2A**.

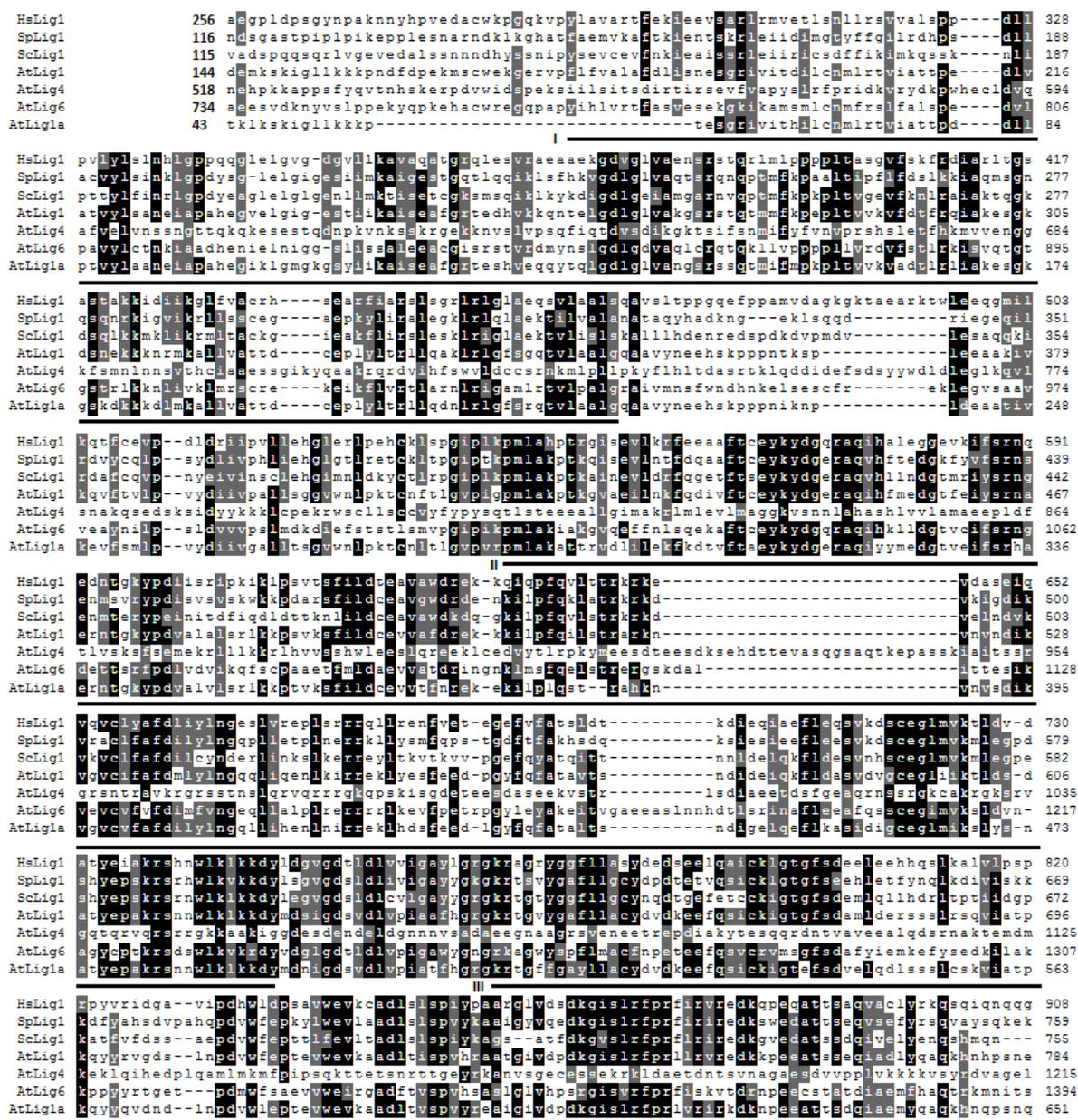


Figure 1. Alignment of central region of Ligase proteins from *Homo sapiens* (Hs), *Schizosaccharomyces pombe* (Sp), *Saccharomyces cerevisiae* (Sc) and *Arabidopsis thaliana* (At). The N terminal DNA binding domain (motif I), the catalytic adenylation domain (motif II) and the Oligonucleotide/oligosaccharide-binding (OB) fold domain (motif III) are indicated with black lines under the sequence and conserved amino acids are indicated with black boxes.

Genomic DNA of the homozygous mutant was digested by *Hind*III for Southern blotting (**Figure 2B**). As expected, a band with the size of 2.2kb was detected on the blot, representing the left border-fragment connecting the T-DNA to the host *AtLig1a* gene. Besides the 2.2kb band, there were additional bands on the blot, indicating that additional T-DNAs had been inserted in the genome. In order to reduce the number of the additional T-DNA insertions, we also analyzed the next generation of this mutant by Southern blotting. All progeny plants showed the same bands, suggesting that the T-DNA insertions are linked.

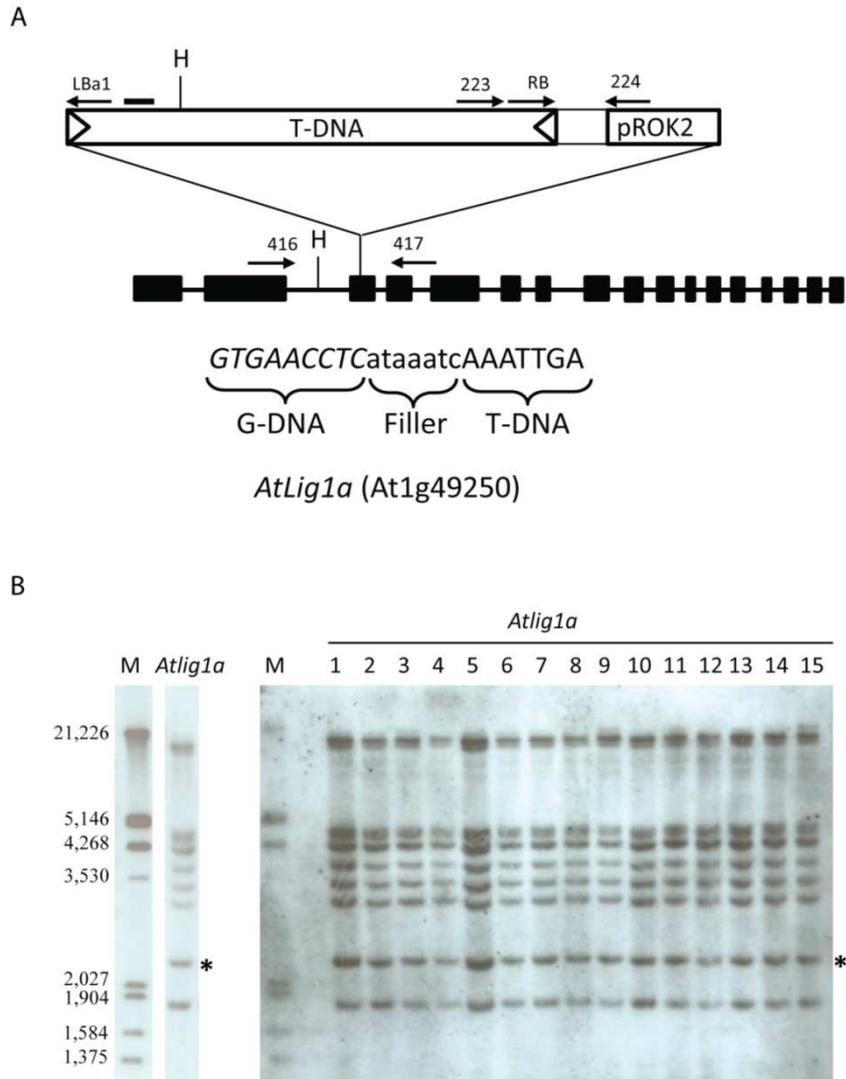


Figure 2. Molecular analysis of the T-DNA insertion in the *Atlig1a* locus. **(A)** Genomic organization of the *Atlig1a* locus is indicated with the positions of the inserted T-DNA. Exons are shown as black boxes and introns are shown as lines. The probe (—) and the restriction enzyme digestion site (H: *Hind*III) used for Southern blot analysis are indicated. Genomic DNA sequence (G-DNA) flanking the T-DNA insertion is shown in italics and filler DNA in small letters. T-DNA border indicated with triangle. **(B)** Southern blot analysis of the *Atlig1a* T-DNA insertion mutant. The genomic DNA was digested with *Hind*III. M: Lambda DNA/ *Eco*RI+*Hind*III marker. The expected band was indicated by Asterisk. 1 to 15 indicate individual progeny plants of *Atlig1a*.

DNA damage response

No obvious differences were observed between the *Atlig4*, *Atlig6*, *Atlig4lig6* and *Atlig1a* mutants and the wildtype under standard growth conditions (**Figure 3**). In order to study whether *Atlig1a* functions in DNA repair, the mutant was tested for sensitivity to the genotoxic agents bleomycin and MMS and compared to *Atlig4* and *Atlig6* mutants. As expected, the *Atlig4* and *Atlig4lig6* mutants turned out to be hypersensitive to bleomycin and MMS, whereas the *Atlig6* and *Atlig1a* mutants did not show hypersensitivity. In order to study whether the role of *Atlig1a* becomes more prominent in the absence of the other ligases, the *Atlig1a* mutant was crossed with the *Atlig4* and *Atlig4lig6* mutants and homozygous *Atlig4lig1a* and *Atlig4lig6lig1a* mutants were obtained. The *Atlig4lig1a* double and *Atlig4lig6lig1a* triple mutants, were somewhat smaller compared to the *Atlig4* and *Atlig4lig6* mutants. When *Atlig4lig1a* double and *Atlig4lig6lig1a* triple mutants were treated with genotoxic agents (bleomycin or MMS), they showed a more sensitive phenotype compared to the *Atlig4* and *Atlig4lig6* mutants similar to the difference under standard growth conditions (**Figure 3**). However, *Atlig6* did not show extra phenotypes in *Atlig4lig6lig1a* triple mutant compare to *Atlig4lig1a* double mutant.

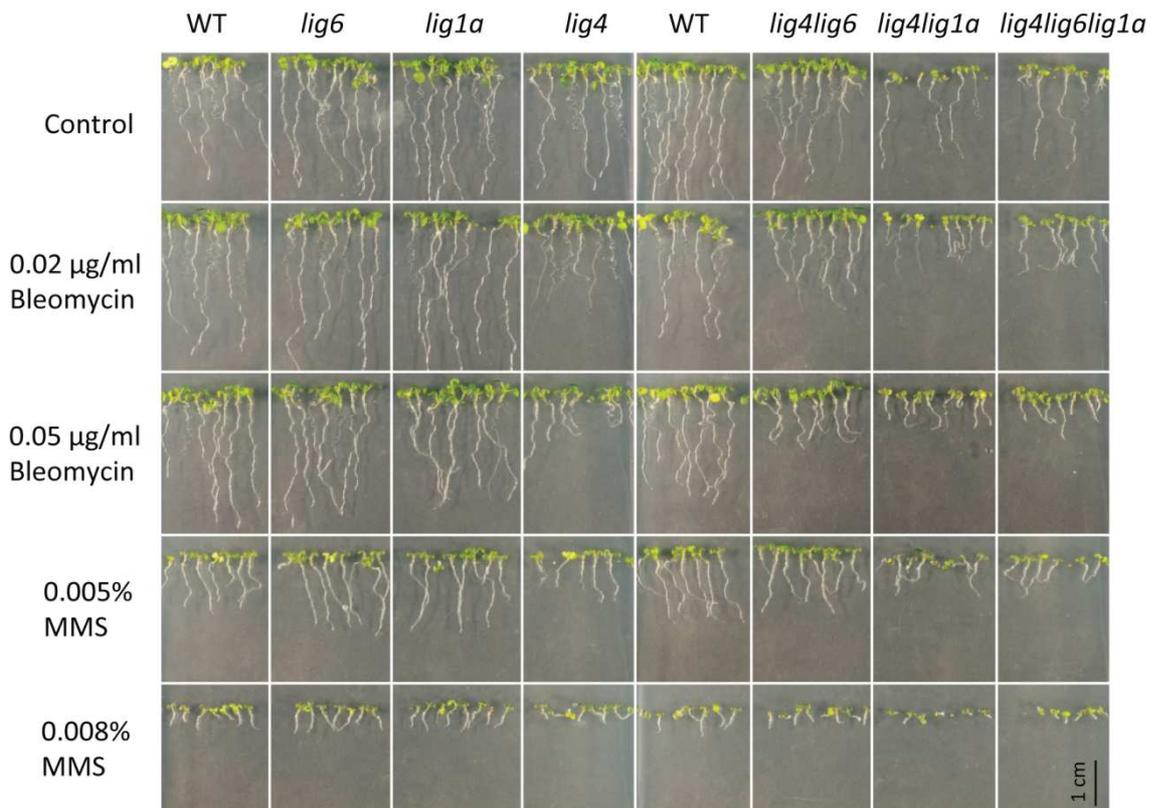


Figure 3. Response to DNA-damaging treatments. Phenotypes of wild-type plants and *Atlig4*, *Atlig1a*, *Atlig6*, *Atlig4lig1a*, *Atlig4lig6lig1a* mutants germinated on $\frac{1}{2}$ MS medium (control) or $\frac{1}{2}$ MS medium containing 0.02, 0.05 $\mu\text{g/ml}$ bleomycin or 0.005 %, 0.008 % MMS and photographed 1 week after germination.

In order to quantify the DNA damage and repair in *Atlig4*, *Atlig4lig1a* and *Atlig4lig6lig1a* mutants, comet assays (N/N protocol) were performed after bleomycin treatment. For each treatment around 100 randomly chosen nuclei from 6 independent mini gel replicas were analyzed by using CometScore™. The percentage DNA in the tail is related to the amount of DNA damage. The *Atku80* mutant was used as positive control. Without any treatment, the genomic DNA of *Atku80* already had a bit more DNA damage than that of the wild type as expected, while *Atlig4*, *Atlig4lig1a* and *Atlig4lig6lig1a* mutants showed nearly the same amount of DNA damage as the wild type (**Figure 4A**). All mutants had a higher level of nuclear DNA damage than the wild type after 24h treatment. The DNA damage of the wild type and the *Atku80*, *Atlig4* and *Atlig4lig1a* mutants was slowly repaired following the recovery step. The wild type had 68% damage remaining after 24h and the mutants somewhat more. The *Atlig4lig6lig1a* mutant however hardly showed any repair after 24h recovery (**Figure 4B**). This means that NHEJ is still functional in the *Atku80*, *Atlig4* and *Atlig4lig1a* mutants but it progresses slower than in wild type. However, simultaneous mutation of the 3 ligase genes *AtLig4*, *AtLig6* and *AtLig1a* inhibited NHEJ such that no DNA repair could be detected within 24 hours of recovery.

Discussion

Here we characterized the *Atlig1a*, *Atlig4lig1a* and *Atlig4lig6lig1a* mutants and compared these with the wild type and *Atlig4* and *Atlig4lig6* mutants. No phenotypical difference was detected in the *Atlig1a* mutants compared to the wild type. Since *AtLig1*, *AtLig4* and *AtLig6* are still present in this mutant, the function of AtLIG1a may be concealed or suppressed. In order to check if the *Atlig1a* gene is involved in NHEJ DNA repair pathways, we also isolated homozygous *Atlig1a* double and triple mutants by crossing with the *Atlig4* and *Atlig4lig6* mutants. It is impossible to obtain *Atlig1* homozygous mutants because mutation of the *AtLig1* gene is lethal (Babiychuk *et al.* 1998). There was a phenotypical difference under both normal growth conditions and under genotoxic stress between the *Atlig4lig1a* and *Atlig4lig6lig1a* mutants and the *Atlig4* and *Atlig4lig6* mutants. It is difficult to determine if treatment with genotoxic chemicals had a more severe effect on the *Atlig1a* mutants, compared to the difference under standard growth conditions. This could imply that AtLIG1a is involved in a process that occurs under standard growth conditions, such as DNA replication, which also generates a low level of DSBs. The results obtained from the comet assay analyses, showed an effect of the mutation of *Atlig6* in the triple mutant: no DNA repair was observed in the *Atlig4lig6lig1a* triple mutant within the 24 hours recovery period, although AtLig1, reported to be involved in replication and several DNA repair pathways (Arakawa and Iliakis 2015), was still present. This probably also true for *Atlig1a* since it gave an more severe phenotype compared to *Atlig4lig6* (**Figure 3**).

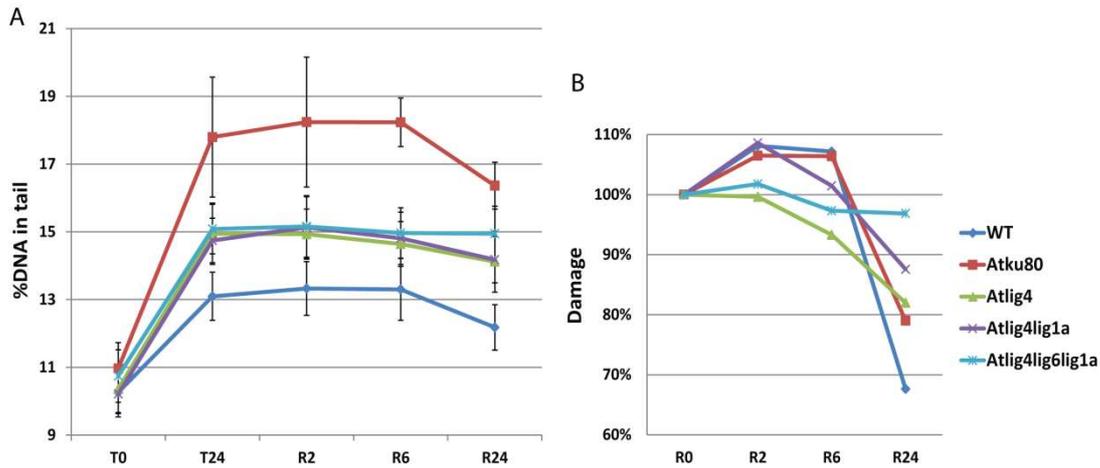


Figure 5. Quantification of DNA damage by comet assays. **(A)** The fraction of DNA in comet tails (%DNA in tail) was used as a measure of DNA damage in wild-type plants and *Atku80*, *Atlig4*, *Atlig4lig1a* and *Atlig4lig6lig1a* mutants. Around 100 nuclei for each treatment were analyzed. The means of %tail-DNA after bleomycin treatment are shown. T0: no treatment; T24: treated 24 hours. **(B)** The means of damage remaining after recovery are shown. R0: no recovery; R2, R6, R24: treated cell 24 hours followed by 2, 6 or 24 hours recovery. Damage at R0 was set as 100%.

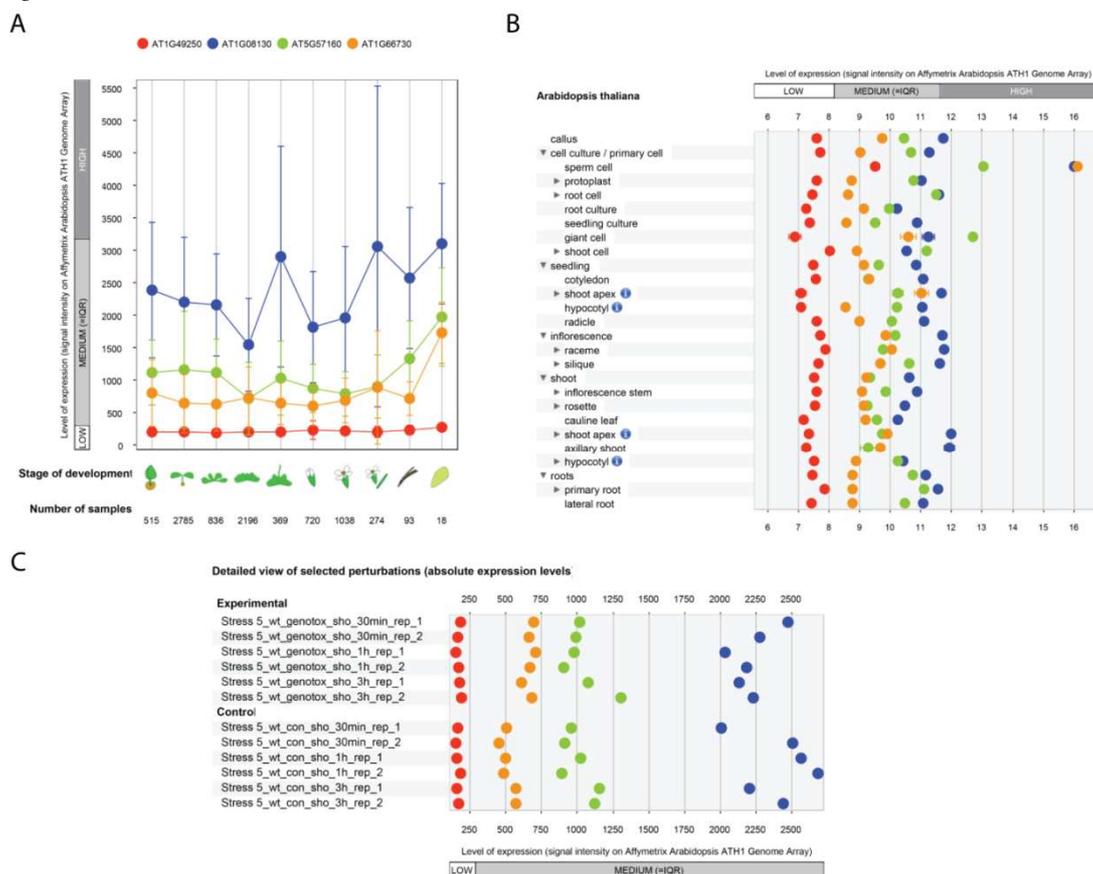


Figure 4. Expression pattern of the ligase genes *AtLig1a* (*At1g49250*), *AtLig1* (*At1g08130*), *AtLig4* (*At5g57160*), *AtLig6* (*At1g66730*) in Arabidopsis. Data were retrieved using Genevestigator analysis tools of development **(A)**, anatomy **(B)** and stimulus **(C)**. Scatterplot outputs of developmental **(A)** and tissue-specific **(B)** expression patterns are shown. *AtLig1a* is indicated as red dot, *AtLig1* as blue dot, *AtLig4* as green dot and *AtLig6* as yellow dot. The expression levels in response to genotoxic treatments for the four ligase genes are shown in **(C)**. “Medium” indicates the interquartile range (IQR).

By using the genevestigator tool (<https://genevestigator.com/gv/index.jsp>), we compared the expression of the ligase genes of *Arabidopsis thaliana*. This revealed that the expression levels of the *AtLig1a* gene was lowest compared to the other known ligase genes: *AtLig1*, *AtLig4* and *AtLig6*, independent of cell type or stage of development and not induced by genotoxic stress (**Figure 5**). We tried to clone the full length cDNA by normal cloning strategies, but we did not succeed, possibly due to the low abundance of the mRNA. We also tried to check the expression of *Atlig1a* gene by Northern blotting using a short part of *AtLig1a* as a probe, but we could not detect any signal on the blot. Finally, we cloned the *Atlig1a* promoter upstream of the GUS reporter and analysed the expression of this construct. Although the presence of the intact construct was verified by PCR, no GUS staining at all was seen in three transgenic plant lines. The results obtained by RNA seq also confirmed that the expression of *AtLig1a* gene is extremely low and more than 100 fold lower than the expression level of the *AtLig1* gene (data not shown). Therefore, we concluded that the expression level of the *Atlig1a* gene is very low in wild type seedlings under standard growth conditions. During our studies the supposedly low expression of *Atlig1a* was recently also reported by others (Li *et al.* 2015). Although we could not detect any expression of *AtLig1a*, a developmental effect was observed when in addition to *AtLig1a* both *AtLig4* and *AtLig6* were also mutated. Possibly *AtLig1a* expression is induced when these ligases are both mutated.

In summary, our results indicate that AtLIG1a probably has a function in DNA replication or NHEJ repair, when other ligase genes are mutated. Since the mechanisms of the b-NHEJ repair pathways in plants remains unclear, further work is needed to find out if AtLIG1a is indeed involved in the final ligation step of plant b-NHEJ repair.

Acknowledgements

This work was supported by a grant from the China Scholarship Council (CSC grant no. 2011699040).

Table 1. Sequences of primers used for characterization of T-DNA insertion lines and plasmid construction.

Name	Locus	Sequence
LBa1	T-DNA left border	5'-TGGTTCACGTAGTGGGCCATCG-3'
RB	T-DNA right border	5'-TTTGGAAGTACAGAACCGC-3'
Sp223	T-DNA right border	5'-TTCAACGTTGCGGTTCTGTC-3'
Sp224	T-DNA right border	5'-TGTGGTTGGCATGCACATAC-3'
Sp271	pROK2	5'-CCCGTGTCTCTCCAAATG-3'
Sp272	pROK2	5'-CAGGTCCCCAGATTAGCC-3'
Sp416	<i>Atlig1a</i>	5'-GCCGTTCTCTCAGACTATG-3'
Sp417	<i>Atlig1a</i>	5'-TACAGCAGCCTGTCCCAAGG-3'
Sp268	<i>Atlig4</i>	5'-ATGCTGAGGACTTGTTTAATG-3'
Sp269	<i>Atlig4</i>	5'-ACCAACATTTACCATCAAGG-3'
Sp264	<i>Atlig6-1</i>	5'-GTCAACTCTGTCAATGGTCC-3'
Sp265	<i>Atlig6-1</i>	5'-AATATCAAACACGAAGACGCAGAC-3'
Hp1	<i>Atlig1a</i> promoter	5'-CCAAGCTTTTTTAATTCACCTTGAATTTAC-3'
Hp2	<i>Atlig1a</i> promoter	5'-AAGGATCCGAGAACGTCAAATGCAGATAT-3'

References

- Ahnesorg, P., P. Smith, and S. P. Jackson, 2006 XLF interacts with the XRCC4-DNA Ligase IV complex to promote DNA nonhomologous end-joining. *Cell* 124: 301–313.
- Alonso, J. M., M. Alonso, A. N. Stepanova, T. J. Leisse, D. K. Stevenson *et al.*, 2003 Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301: 653–657.
- Arakawa, H., T. Bednar, M. Wang, K. Paul, E. Mladenov *et al.*, 2012 Functional redundancy between DNA ligases I and III in DNA replication in vertebrate cells. *Nucleic Acids Res.* 40: 2599–2610.
- Arakawa, H., and G. Iliakis, 2015 Alternative okazaki fragment ligation pathway by dna ligase III. *Genes* 6: 385–398.
- van Attikum, H., P. Bundock, R. M. Overmeer, L. Y. Lee, S. B. Gelvin *et al.*, 2003 The *Arabidopsis* *AtLIG4* gene is required for the repair of DNA damage, but not for the integration of *Agrobacterium* T-DNA. *Nucleic Acids Res.* 31: 4247–4255.
- Babiychuk, E., P. B. Cottrill, S. Storozhenko, M. Fuangthong, Y. Chen *et al.*, 1998 Higher plants possess two structurally different poly(ADP-ribose) polymerases. *Plant J.* 15: 635–645.
- Barnes, D. E., A. E. Tomkinson, A. R. Lehmann, A. D. B. Webster, and T. Lindahl, 1992 Mutations in the DNA ligase I gene of an individual with immunodeficiencies and cellular hypersensitivity to DNA-damaging agents. *Cell* 69: 495–503.
- Chen, J., A. E. Tomkinson, W. Ramos, Z. B. Mackey, S. Danehower *et al.*, 1995 Mammalian DNA ligase III: molecular cloning, chromosomal localization, and expression in spermatocytes undergoing meiotic recombination. *Mol. Cell. Biol.* 15: 5412–5422.
- Clough, S. J., and A. F. Bent, 1998 Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16: 735–743.
- Critchlow, S. E., R. P. Bowater, and S. P. Jackson, 1997 Mammalian DNA double-strand break repair

- protein XRCC4 interacts with DNA ligase IV. *Curr. Biol.* 7: 588–598.
- Lakshmipathy, U., and C. Campbell, 1999 Double strand break rejoining by mammalian mitochondrial extracts. *Nucleic Acids Res.* 27: 1198–1204.
- Levin, D. S., A. E. McKenna, T. a. Motycka, Y. Matsumoto, and A. E. Tomkinson, 2000 Interaction between PCNA and DNA ligase I is critical for joining of Okazaki fragments and long-patch base-excision repair. *Curr. Biol.* 10: 919–922.
- Li, Y., C.-G. Duan, X. Zhu, W. Qian, and J.-K. Zhu, 2015 A DNA ligase required for active DNA demethylation and genomic imprinting in *Arabidopsis*. *Cell Res.* 25: 757–760.
- Mackey, Z. B., C. Niedergang, J. M. -d. Murcia, J. Leppard, K. Au *et al.*, 1999 DNA ligase III is recruited to DNA strand breaks by a zinc finger motif homologous to that of poly(ADP-ribose) polymerase: identification of two functionally distinct DNA binding regions within DNA ligase III. *J. Biol. Chem.* 274: 21679–21687.
- Mackey, Z. B., W. Ramos, D. S. Levin, C. A. Walter, J. R. McCarrey *et al.*, 1997 An alternative splicing event which occurs in mouse pachytene spermatocytes generates a form of DNA ligase III with distinct biochemical properties that may function in meiotic recombination. *Mol. Cell. Biol.* 17: 989–998.
- MacNeill, S. a, 2001 DNA replication: Partners in the Okazaki two-step. *Curr. Biol.* 11: R842–R844.
- Menke, M., I. Chen, K. J. Angelis, and I. Schubert, 2001 DNA damage and repair in *Arabidopsis thaliana* as measured by the comet assay after treatment with different classes of genotoxins. *Mutat. Res.* 493: 87–93.
- de Pater, S., M. Caspers, M. Kottenhagen, H. Meima, R. Ter Stege *et al.*, 2006 Manipulation of starch granule size distribution in potato tubers by modulation of plastid division. *Plant Biotechnol. J.* 4: 123–134.
- Paul, K., M. Wang, E. Mladenov, A. Bencsik-Theilen, T. Bednar *et al.*, 2013 DNA ligases I and III cooperate in alternative non-homologous end-joining in vertebrates. *PLoS One* 8: e59505.
- Simsek, D., E. Brunet, S. Y. W. Wong, S. Katyal, Y. Gao *et al.*, 2011 DNA ligase III promotes alternative nonhomologous end-joining during chromosomal translocation formation. *PLoS Genet* 7(6): e1002080.
- Taylor, R. M., B. Wickstead, S. Cronin, and K. W. Caldecott, 1998 Role of a BRCT domain in the interaction of DNA ligase III-alpha with the DNA repair protein XRCC1. *Curr. Biol.* 8: 877–880.
- Timson, D. J., M. R. Singleton, and D. B. Wigley, 2000 DNA ligases in the repair and replication of DNA. *Mutat. Res. Repair* 460: 301–318.
- Waterworth, W. M., G. Masnavi, R. M. Bhardwaj, Q. Jiang, C. M. Bray *et al.*, 2010 A plant DNA ligase is an important determinant of seed longevity. *Plant J.* 63: 848–860.
- Weijers, D., M. Franke-van Dijk, R. J. Vencken, A. Quint, P. Hooykaas *et al.*, 2001 An *Arabidopsis* Minute-like phenotype caused by a semi-dominant mutation in a ribosomal protein S5 gene. *Development* 128: 4289–4299.
- Wilson, T. E., U. Grawunder, and M. R. Lieber, 1997 Yeast DNA ligase IV mediates non-homologous DNA end joining. *Nature* 388: 495–498.
- Windhofer, F., W. Wu, and G. Iliakis, 2007 Low levels of DNA ligases III and IV sufficient for effective NHEJ. *J. Cell. Physiol.* 213: 475–483.

Chapter 4

Mre11 and Ku80 control different pathways of DNA repair and T-DNA integration in Arabidopsis

Hexi Shen, Paul J. J. Hooykaas, Sylvia de Pater

Department of Molecular and Developmental Genetics, Institute of Biology, Leiden University,
Leiden, 2333 BE, The Netherlands

Abstract

The Mre11 complex has a critical role in DNA damage signaling and DSB repair. The Arabidopsis Mre11 protein shares a highly conserved N-terminus with yeast and mammalian Mre11. T-DNA insertions in this conserved region, thus fully inactivating the gene, resulted in many developmental defects including sterility (*mre11-1*, *mre11-3*). The *mre11-4* mutant with T-DNA insertion in the C-terminal region exhibited a similar phenotype, but the *mre11-2* mutant with a slightly downstream insertion was fully fertile and grew normally. Nevertheless, also this latter mutant was hypersensitive to DNA damaging agents, indicating defective DNA repair. By yeast two-hybrid assays, we found that the Mre11-2 protein (truncated at amino acids 1 to 529) still interacted with Rad50, but the Mre11-4 protein (truncated at amino acids 1 to 499) did not, indicating that the area between amino acid 499 and 529 is important for interaction with Rad50. As many functions of Mre11 rely on complex formation with Rad50, this explains the large difference between *mre11-2* and *mre11-4* mutants. As the *mre11-2* mutant showed sensitivity to DNA damage, we tested whether this was due to a defect in c-NHEJ repair by generating *ku80mre11-2* double mutant plants. Interestingly, the *ku80mre11-2* double mutant plants were more sensitive to DNA damage stress than single mutant plants and moreover were resistant to *Agrobacterium*-mediated transformation. These results indicate that Mre11 is important for backup NHEJ pathways involved in T-DNA integration.

Introduction

Non-homologous end-joining (NHEJ) is believed to be the major pathway for the repair of DSBs in most higher eukaryote cells. The classical NHEJ (c-NHEJ) pathway embraces as critical factors Ku70/80 and Lig4, which are also found in plant cells. In the absence of Ku, efficient repair still occurs indicating that the cell has backup pathways for repair, which are called backup NHEJ (b-NHEJ) or sometimes also microhomology-mediated end-joining (MMEJ) or alternative NHEJ (a-NHEJ). Such repair pathways were discovered in yeast (*Saccharomyces cerevisiae*) (Boulton and Jackson 1996) and in mammalian cells (see review Deriano and Roth 2013). The c-NHEJ pathway is relatively precise but still sometimes causes small deletions or insertions. In the absence of Ku, resection of the ends becomes possible, which may lead to large deletions, insertions and translocation which are features of b-NHEJ. However, the molecular mechanism underlying b-NHEJ remain unclear. Also in plants, evidence was obtained that repair can occur by such backup pathways to maintain plant genome integrity (Charbonnel *et al.* 2010; Jia *et al.* 2013). Several factors that are involved in b-NHEJ have already been identified: PARP1/2, Ligase3, XRCC1, Pol θ , CtIP and Mre11.

Mre11 has been well defined in DNA damage response (DDR) and homologous recombination (HR). It combines with Rad50 and Nbs1 (Xrs2 in yeast) to form the MRN (MRX in yeast) complex and acts as a DSB sensor for activation of ATM and downstream signaling. Mre11 has both single-stranded DNA endonuclease activity and 3' to 5' exonuclease activity, and thus MRN plays a crucial role in promoting DSB end resection, which is likely to determine pathway choice between HR and NHEJ and repair outcome (see review Ceccaldi *et al.* 2016). However, the MRN/MRX complex, and Mre11 in particular, have been found to have also a direct role in NHEJ repair pathways. In yeast, MRX has been implicated in NHEJ repair and in the MMEJ pathway (Chen *et al.* 2001; Ma *et al.* 2003). Deficiency of the MRX complex resulted in a significant reduction in the frequency of T-DNA integration by NHEJ in yeast (van Attikum *et al.* 2001). In mammalian cells, Mre11 promotes both c-NHEJ and b-NHEJ pathways during the repair of I-SceI endonuclease-induced DSBs (Rass *et al.* 2009; Xie *et al.* 2009).

In plants, it has been shown that MRN mutants (Mre11, Rad50 and Nbs1) exhibit an increased sensitivity to DNA damage (Gallego *et al.* 2001; Bundock and Hooykaas 2002; Waterworth *et al.* 2007). Four *Arabidopsis thaliana* T-DNA insertion mutant lines of the *MRE11* gene were previously described: *mre11-1* and *mre11-2* (Bundock and Hooykaas 2002), *mre11-3* (Puizina *et al.* 2004), *mre11-4* (Šamanić *et al.* 2013). The *mre11-2* plants displayed normal growth and fertility, while *mre11-1*, *mre11-3* and *mre11-4* mutant lines had many developmental defects and were sterile. These reports demonstrated that Mre11 has essential roles in the response to irradiation-induced DSBs, meiotic recombination and maintenance of chromosomal stability.

In this chapter, we compared *mre11-4* and *mre11-2* both with insertions in the C-terminal part of Mre11 to determine which function is missing in *mre11-4* versus *mre11-2*. The yeast two-hybrid results showed that *mre11-2* still preserved interaction with Rad50, but *mre11-4* did not. In order to study the role of Mre11 in plant cells, including NHEJ and *Agrobacterium* T-DNA integration, we generated a homozygous double mutant *ku80mre11-2*, in which the c-NHEJ repair pathway is inactivated. More sensitivity to DNA damage stress

was found in this double mutant revealing a crucial role for Mre11 in DNA repair in absence of Ku80. Using yeast as a model it was shown in our lab that T-DNA integration in yeast is dependent on the c-NHEJ pathway of DSB repair (van Attikum et al. 2001). However, inactivation of c-NHEJ or of both c-NHEJ and factors supposedly involved in b-NHEJ pathways in plants (PARP1, PARP2, XRCC1) still allowed T-DNA integration by *Agrobacterium* (see Chapter 2), suggesting that there must be other important proteins and/or another NHEJ pathway mediating the T-DNA integration in plant cells. Remarkably, the *ku80mre11-2* double mutant was resistant to T-DNA integration, while the single mutants still were proficient in T-DNA integration. This suggests that Ku80 and Mre11 control different pathways of T-DNA integration, each of which functions independently in *Agrobacterium* T-DNA integration in plant cells as only in the absence of both of them integration of T-DNA was not observed by the root transformation assay.

Materials and methods

Plant material

The T-DNA insertional *Arabidopsis thaliana* mutants of *ku80* and *mre11-2*, have been described previously (Bundock and Hooykaas 2002; Jia *et al.* 2012). The double *ku80mre11-2* mutant was produced by crossing *ku80* and *mre11-2* homozygous mutants followed by PCR analysis of progeny for identification of double homozygous plants.

Assays for sensitivity to bleomycin and methyl methane sulfonate (MMS)

Seeds of wild type, *ku80*, *mre11-2* and *ku80mre11-2* mutants were surface-sterilized as described (Weijers *et al.* 2001) and germinated on solid ½ MS medium without additions or containing 0.02 µg/ml or 0.05 µg/ml bleomycin (Sigma), 0.005% (v/v), 0.007% or 0.01% MMS (Sigma) and photographed after 2 weeks. After 10 days growth, the fresh weight (compared with controls) was determined by weighing the seedlings in batches of 20 in duplicate. Statistical analyses were performed using Prism version 5 (GraphPad Software Inc.).

Cell Death Assay

Four days old seedlings on MS-agar plates were transferred to new MS-plates with DNA damaging agents. After one or four days treatments, 20 seedlings were placed in propidium iodide solution (PI, Sigma-Aldrich, 5 µg/ml in water) for 1 min and rinsed three times with water. Root tips were then transferred to slides in a drop of water and covered with a cover slip for observation under the fluorescence microscope.

Yeast two-hybrid Assays

Full-length Mre11 was cloned into pACT2, while full length Rad50 was cloned into pAS2.1 (James *et al.* 1996). Mre11₍₁₋₄₉₉₎, Mre11₍₁₋₅₁₀₎, Mre11₍₁₋₅₂₉₎, Mre11₍₅₃₀₋₇₂₀₎ were amplified by PCR and cloned into pACT2 and pAS2.1. Primer sequences are presented in Table1. All PCR fragments were verified by sequencing. Interaction assays were performed by co-transformation of bait and prey plasmids into yeast strain PJ69-4A as previously described (Gietz *et al.*, 1992), and plated on MY minimal medium, supplemented with methionine, uracil, adenine and histidine and lacking leucine and tryptophan (MY/-LW). Subsequently,

cells were incubated for 16 h in liquid MY/-LW and spotted onto selective solid MY medium lacking leucine, tryptophan and histidine (MY/-LWH) supplemented with increasing concentrations of 3-amino-1,2,4-triazole (3-AT, Sigma) ranging from 0 to 30 mM. Yeast cells were allowed to grow for 5 days at 30°C.

Root transformation

Root transformation was performed as described (Vergunst *et al.* 2000). Root segments were infected with *Agrobacterium* LBA1100 (pCambia3301). After co-cultivation on callus induction medium containing 100 µM acetosyringone for 48 hours, root segments were washed, dried and incubated on shoot induction medium supplemented with 30 µg/ml phosphinothricin, 500 µg/ml carbenicillin and 100 µg/ml vancomycin. After 3-4 weeks, plates were photographed and transformation efficiencies were scored as infected root segments that produced any form of green callus. Statistical analyses were performed using Prism version 5 (GraphPad Software Inc.).

Results

The C-terminal part of Mre11 contains the domain interacting with Rad50

The Arabidopsis Mre11 protein shared a large conserved domain in the N-terminus with Mre11 from other organisms, but there was less clear homology in the C-terminus (**Figure 1A**). Four Arabidopsis *MRE11* T-DNA insertion mutants have been characterized. The *mre11-2* mutant exhibited a normal growth phenotype and was fertile, whereas the *mre11-1*, *mre11-3* and *mre11-4* mutants were sterile and showed obvious morphological abnormalities. While the T-DNA insertions in *mre11-1* and *mre11-3* were localized in the N-terminal part of the gene, the insertions in *mre11-2* and *mre11-4* were located in the C-terminus. As a result of T-DNA insertion, 10 additional amino acids (NTQLKNVNMI) may form the C-terminus of the predicted protein in the *mre11-2* mutant, while 35 additional amino acids (ARRYRFSCLITFFN-SGLLFQTGTTLNPFSGYSFDL) are predicted at the C-terminus in the *mre11-4* mutant. Except for the extra amino acids formed at the C-terminus, putative truncated Mre11 proteins formed in the *mre11-2* and *mre11-4* mutants would differ only by 30 amino acids. *In silico* analysis of the predicted truncated Mre11 proteins showed that the *mre11-4* deletion removes a large region that is highly conserved in Mre11 proteins of different plants, while only a small part of this is deleted in the Mre11 protein encoded by the *mre11-2* allele (**Figure 1B**).

A

Human	-----mstadalldde	n	t	r	k	i	l	v	a	t	d	h	l	g	f	m	e	k	d	-	a	v	r	g	n	d	f	v	t	d	e	i	l	r	l	a	q	e	n	v	d	f	i	l	l	g	d	f	h	a	n	k	p	s	r	k	t	l	h	74
<i>S. cerevisiae</i>	-----mdypdpd	t	r	i	l	t	d	h	v	g	y	n	e	d	-	p	l	g	d	s	w	k	t	h	e	m	l	a	k	n	n	v	d	m	v	q	g	d	f	h	v	n	k	p	s	k	k	l	y	g	70									
<i>S. pombe</i>	mpndpdmnnelhne	n	t	r	i	l	s	d	p	h	y	g	e	k	-	p	v	r	g	n	d	f	v	s	f	n	e	i	e	a	r	e	r	d	v	d	m	l	l	g	d	f	h	a	n	k	p	s	r	k	l	y	79							
<i>p. furiosus</i>	-----mgikfahlad	h	l	g	y	e	q	f	h	k	p	e	e	f	a	e	f	k	n	e	i	a	v	q	e	n	v	d	f	i	l	l	g	d	f	h	a	n	k	p	s	r	k	l	y	g	65													
<i>Arabidopsis</i>	-----msredfsd	t	r	i	v	a	t	d	h	l	g	f	m	e	k	d	-	e	r	h	d	s	f	k	a	f	e	a	c	s	i	a	e	k	q	v	d	f	i	l	l	g	d	f	h	a	n	k	p	s	r	k	t	l	v	71				

B

Nicotiana	t	r	n	k	i	a	r	d	e	a	v	k	e	e	d	i	l	k	v	g	e	c	e	r	v	k	a	t	q	n	k	d	q	p	f	f	n	g	s	l	d	i	r	g	k	t	v	g	s	a	v	f	d	d	e	d	t	m	l	s	a	k	554						
Rice	t	r	k	i	n	s	e	a	d	k	f	k	e	e	e	d	i	i	k	v	g	e	c	m	e	r	v	k	a	t	l	r	s	k	e	d	s	r	f	t	s	q	n	l	d	t	g	r	v	t	a	g	n	l	n	s	f	d	d	e	d	t	r	e	m	l	l	g	551
Maize	t	r	n	k	i	s	e	a	d	k	f	k	e	e	e	d	i	i	k	v	g	e	c	m	e	r	v	k	a	t	l	r	s	k	e	d	s	r	f	t	s	q	n	l	d	t	g	r	v	t	a	g	n	l	n	s	f	d	d	e	d	t	r	e	m	l	l	g	559
<i>Arabidopsis</i>	t	r	n	k	i	a	r	d	e	a	v	k	e	e	d	i	l	k	v	g	e	c	e	r	v	k	a	t	l	r	s	k	e	d	s	r	f	t	s	q	n	l	d	t	g	r	v	t	a	g	n	l	n	s	f	d	d	e	d	t	r	e	m	l	l	g	551		

Figure 1. (A) Alignment of Mre11 proteins from *Homo sapiens*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Pyrococcus furiosus* and *Arabidopsis thaliana*. The conserved amino acids are indicated with black and gray boxes. The Rad50 binding domains (RBD) in *P. furiosus* are underlined. (B) Alignment of Mre11 C-terminus from *Nicotiana tabacum*, *Oryza sativa* (rice), *Zea mays* (maize) and *Arabidopsis thaliana*.

To get a better understanding of the functional differences between the Mre11-2 protein truncated at amino acid 529 and the Mre11-4 protein truncated at amino acid 499, yeast two hybrid assays were performed (**Figure 2**). Full-length Mre11 displayed a strong interaction with Rad50 and with itself, indicating that the Arabidopsis Mre11 can potentially homodimerize and form a Mre11-Rad50 complex as expected. To test if the truncated proteins formed in *mre11* mutants still could interact with Rad50 and homodimerize, each of four fragments of Mre11 (amino acids 1 to 529; 1 to 510; 1 to 499; 530 to 720) cloned in pAS2-1 and pACT2 were cotransformed into yeast. A full-length clone of Rad50 in pACT2 and each of four fragments of Mre11 cloned in pAS2-1 were cotransformed into yeast. We found that amino acids 1 to 499 of Mre11 were sufficient for self-interaction. Although Mre11 amino acids 1 to 529 displayed a strong interaction with Rad50 comparable to that observed for full-length Mre11, the smaller version (amino acids 1 to 499) did no longer interact with Rad50. Apparently, presence of the area from amino acids 499 to 529 overlapping with a region that is strongly conserved in plant Mre11 is essential for the interaction of Mre11 with Rad50. Indeed, an Mre11 construct (amino acids 1 to 510) preserving this conserved region, but only 11 amino acids larger than the Mre11-4 protein Mre11 (amino acids 1 to 499) was now able to interact with Rad50. The C-terminal part of Mre11 (aa 530 to 720) by itself did neither homodimerize nor interact with Rad50. These results revealed that the ability to form of Mre11 homodimers and Mre11-Rad50 complexes are preserved in the Mre11-2 protein, but that the Mre11-4 protein could no longer form a complex with Rad50. As Mre11 relies for most of its functions on the formation of a complex with Rad50 (and Nbs1), this explains the different phenotypes of the *mre11-2* mutant in comparison with *mre11-4* and the other T-DNA insertion mutants.

We further tested the interaction between Mre11 and Nbs1 which is the third component of the MRN complex. Although such interactions between plant Mre11 and Nbs1 were observed previously (Waterworth *et al.* 2007), no interaction was observed by our two-hybrid analysis. This may be due to the different yeast two-hybrid systems used. As it was been reported that Mre11 plays a role in b-NHEJ pathways, we studied whether Mre11 interacts with proteins known to be involved in the b-NHEJ repair pathway (such as PARP1, PARP2, LIG1, XRCC1, COM1 and RAD9) by two-hybrid analysis as well. However, no interaction between Mre11 and any of those proteins could be detected.

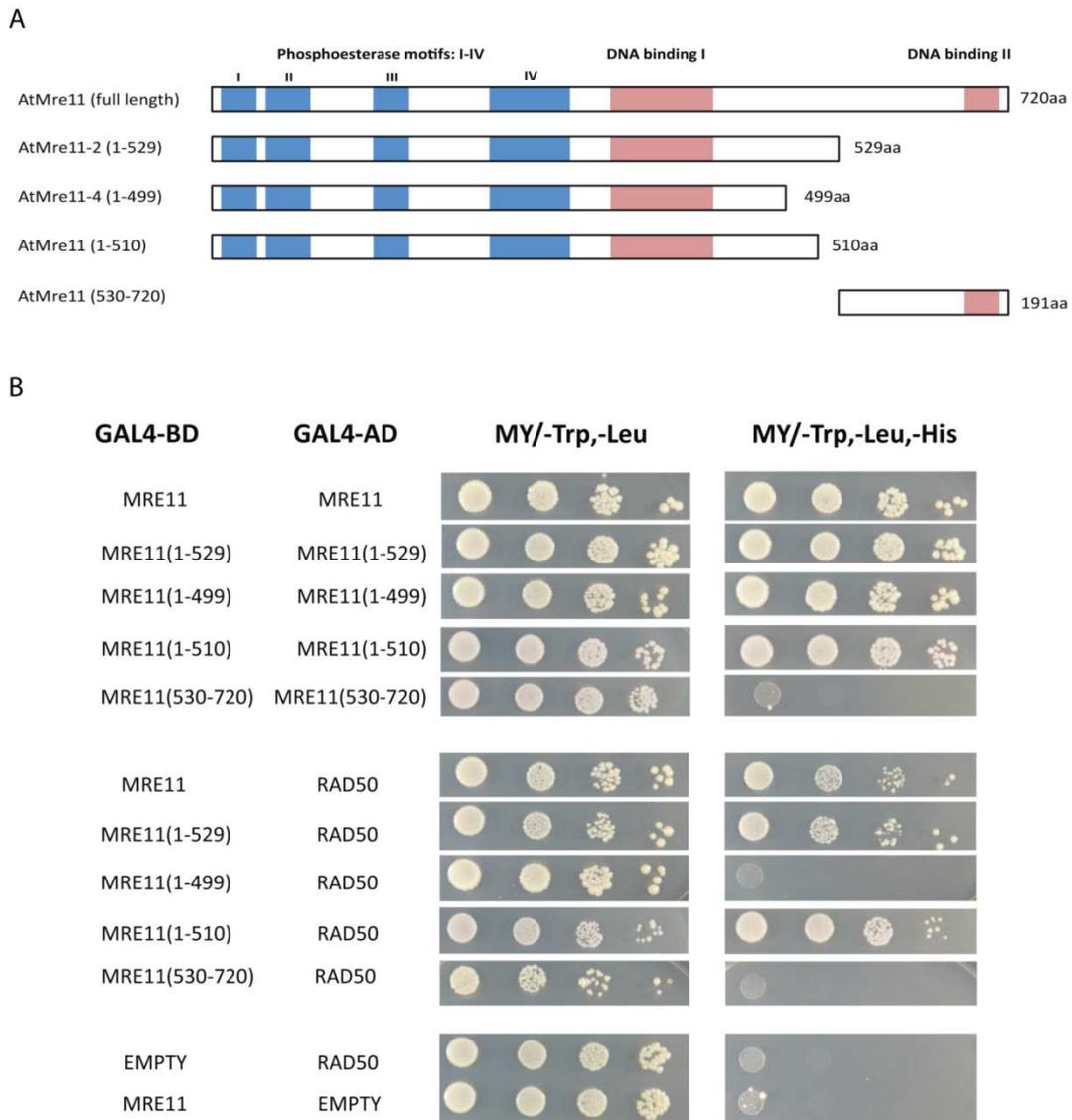


Figure 2. Yeast two-hybrid analysis of MRE11 interactions. **(A)** Schematic regions of MRE11 analyzed in the *in vitro* interaction studies and the constructs used in the yeast two-hybrid analysis. **(B)** Interaction of MRE11 full length protein and deletions with RAD50 protein. The interactions between MRE11 and RAD50 or MRE11 itself resulting in histidine autotrophy.

The $Ku80mre11-2$ double mutant is more sensitive to DNA damage stress than each of the single mutants

Previously, in our lab it was shown that the *ku80* and the *mre11-2* single mutants were hypersensitive to DNA damage stress (Bundock and Hooykaas 2002; Jia *et al.* 2013). In order to test whether the *mre11-2* mutation affects the same or a different repair pathway as *Ku80*, the *ku80mre11-2* double mutant was obtained by crossing the single mutants and tested for sensitivity to genotoxic agents (**Figure 3A**). Without any treatments, the double mutant grew similarly as the wild type. However, upon treatment with bleomycin or MMS, the double mutant grew more slowly and developed fewer true leaves in the assay period than each of the single mutants, indicating increased DNA damage sensitivity in double mutant plants. We also evaluated the effect of DNA damage on the growth of 10-days-old plate-grown plants by

quantification of the fresh weight (**Figure 3B**). We found that the fresh weight of double mutant plants was dramatically reduced in comparison with that of wild type and each of the single mutants, when treated with bleomycin or MMS.

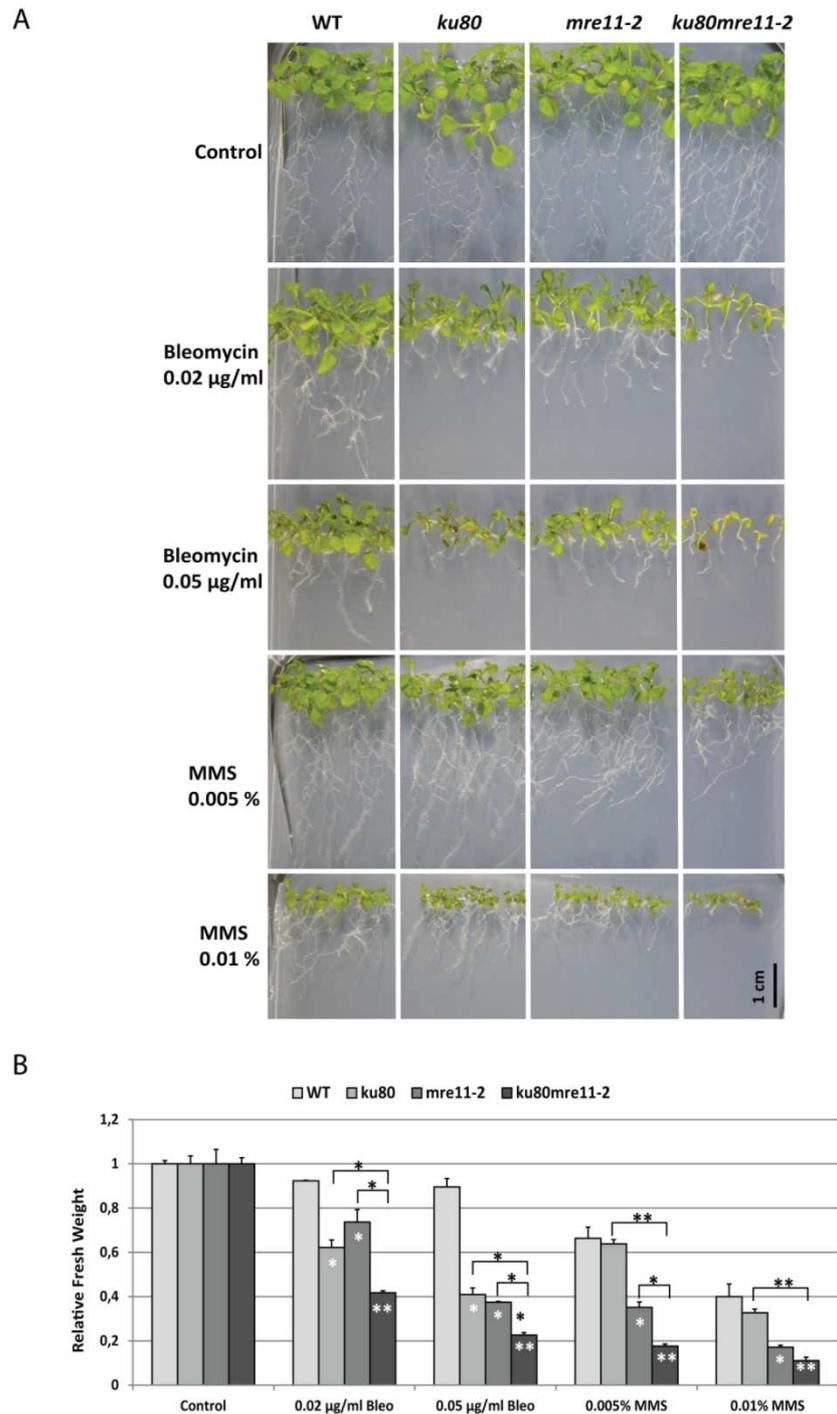


Figure 3. The *ku80mre11-2* double mutant is hypersensitive to DNA damaging treatments. **(A)** Phenotypes of wild-type plants and *ku80*, *mre11-2* and *ku80mre11-2* double mutants germinated on ½ MS medium (control) or ½ MS medium containing 0.02, 0.05µg/ml bleomycin or 0.005%, 0.01% MMS photographed 2 weeks after germination. **(B)** Fresh weight of 10-days-old wild-type and *ku80*, *mre11-2* and *ku80mre11-2* double mutants treated with 0.02, 0.05 µg/ml bleomycin or 0.005%, 0.01% MMS. For each treatment 20 seedlings were weighed in duplicate. Fresh weight of the plants grown 10 days without bleomycin or MMS was set on 1. Error bar represent SD. *Anova* test; for *, $p < 0.001$; **, $p < 0.0001$.

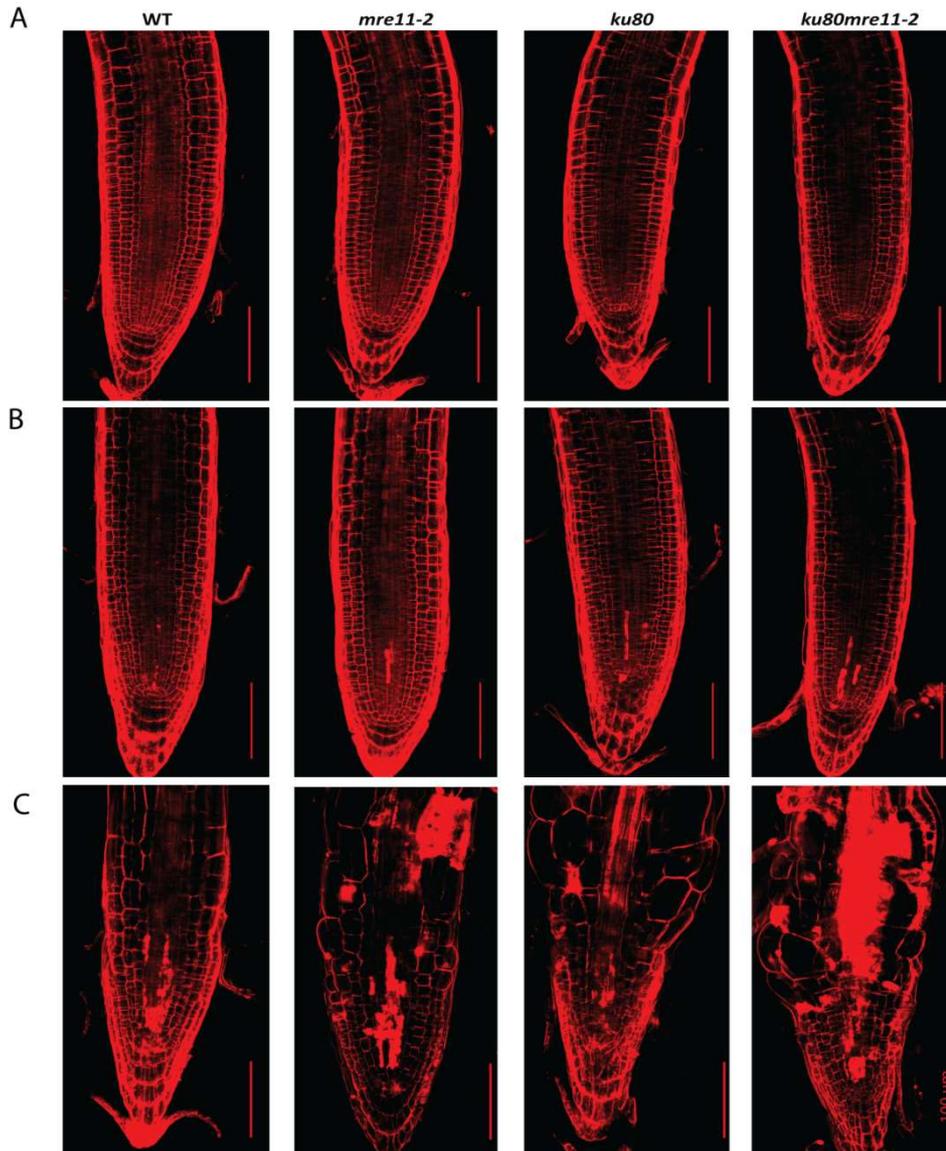


Figure 4. Cell death profile in root tips of wild type and mutants. PI-staining of root tips of 4-days-old wild-type, *ku80*, *mre11-2* and *ku80mre11-2* seedlings grown on MS medium and transferred to MS medium with 0.05 µg/ml bleomycin for 1 day (**B**) and 4 days (**C**), and to MS medium without bleomycin for 2 days as control (**A**).

As the *ku80mre11-2* double mutant plants might have difficulty in repairing DNA damage, this might lead to more cell death than in the wild type or each of the single mutants. To test this, we used propidium iodide (PI) staining to reveal cell death. As can be seen in **Figure 4**, no cell death was observed in wild type and mutants under the normal growth conditions. Bleomycin treatment caused cell death in the root meristem of both wild type and the mutants, and this increased over time. More cell death was seen in the roots of the double mutant than in that of the single mutants, which were more damages than the wild type (**Figure 4B**). When the time period of the genotoxic stress was extended, a much shorter meristematic zone was observed in the root tip of the double mutant (**Figure 4C**), which might be due to the enhanced cell death in the root meristem. Meanwhile, the epidermal cells

of the double mutant were also more deformed and enlarged than those of wild type and each of the single mutants (**Figure 4C**).

Taken together, the results showed that *ku80mre11* double mutant plants grow similarly as wild type plants under normal growth conditions, but they are more sensitive to DNA damage and grow slower in the presence of genotoxic agents than each of the single mutants. This indicates that Mre11 probably plays an important role in Ku-independent DNA repair pathways in somatic plant cells.

The Ku80mre11-2 double mutant is resistant to Agrobacterium mediated T-DNA integration

Previously, we found that both the *ku80* and *mre11-2* mutants were somewhat recalcitrant to transformation in the floral dip procedure (Jia *et al.* 2012), but other groups have shown that the *ku80* mutant is either not or little affected in floral dip or root transformation by *Agrobacterium* (Gallego *et al.* 2003; Park *et al.* 2015). In order to test whether the *ku80mre11-2* double mutant is affected more strongly in *Agrobacterium*-mediated T-DNA transformation, we used the *ku80mre11-2* double mutant together with wild type and single mutants in the root transformation procedure (**Figure 5**). To this end, roots were cocultivated with an *Agrobacterium* strain containing the pCambia3301 binary vector, and callus formation was selected on medium with phosphinothricin. The transformation frequencies of the single mutants were not significantly different from the wild type. The calli formed on the roots of the *mre11-2* mutant grew somewhat greener than those formed by the wild type. This was due to a difference in ecotype between the *mre11-2* mutant (ecotype Ws) and the *ku80* mutant (ecotype Col), but did not influence the transformation results. While wild type and single mutants were transformed at a similar frequency, surprisingly, the *ku80mre11-2* double mutant only produced very few green calli in the transformation procedure (**Figure 5A**). Upon longer incubation of these calli on selection plates no green shoots were formed and the calli eventually died (**Figure 5B, 6A**), indicating that the *ku80mre11-2* double mutant cannot be transformed by root transformation. One allele of Ku80 was sufficient for rescue as can be seen by the efficient transformation of roots with the *ku80+/-mre11-2-/-* genotype (Figure 5A). Also, calli and green shoots were formed normally on medium without selection, although calli of the double mutant grew a bit slower than those of wild type and single mutants (**Figure 6B**).

In order to test whether the lack of callus formation was due to the lack of T-DNA transfer or a deficiency in T-DNA integration, roots were transformed by an *Agrobacterium* strain with the pCambia3301 binary vector. Plant cells receiving T-DNA from this vector express the *GUS* gene before integration, which disappears in due time when the T-DNA is not stably integrated. The results from such transient GUS assay showed that transient transformation (and thus T-DNA transfer) is not affected in these mutants (**Figure 5C**).

Taken together, these results thus show that inactivating both Ku80 and Mre11 at the same time leads to resistance to *Agrobacterium*-mediated stable T-DNA transformation, but does not negatively affect T-DNA transfer. This suggests that Ku80 and Mre11 are involved in two different pathways of T-DNA integration, each of which is not essential, but after inactivation of both transformation is reduced to below detection level.

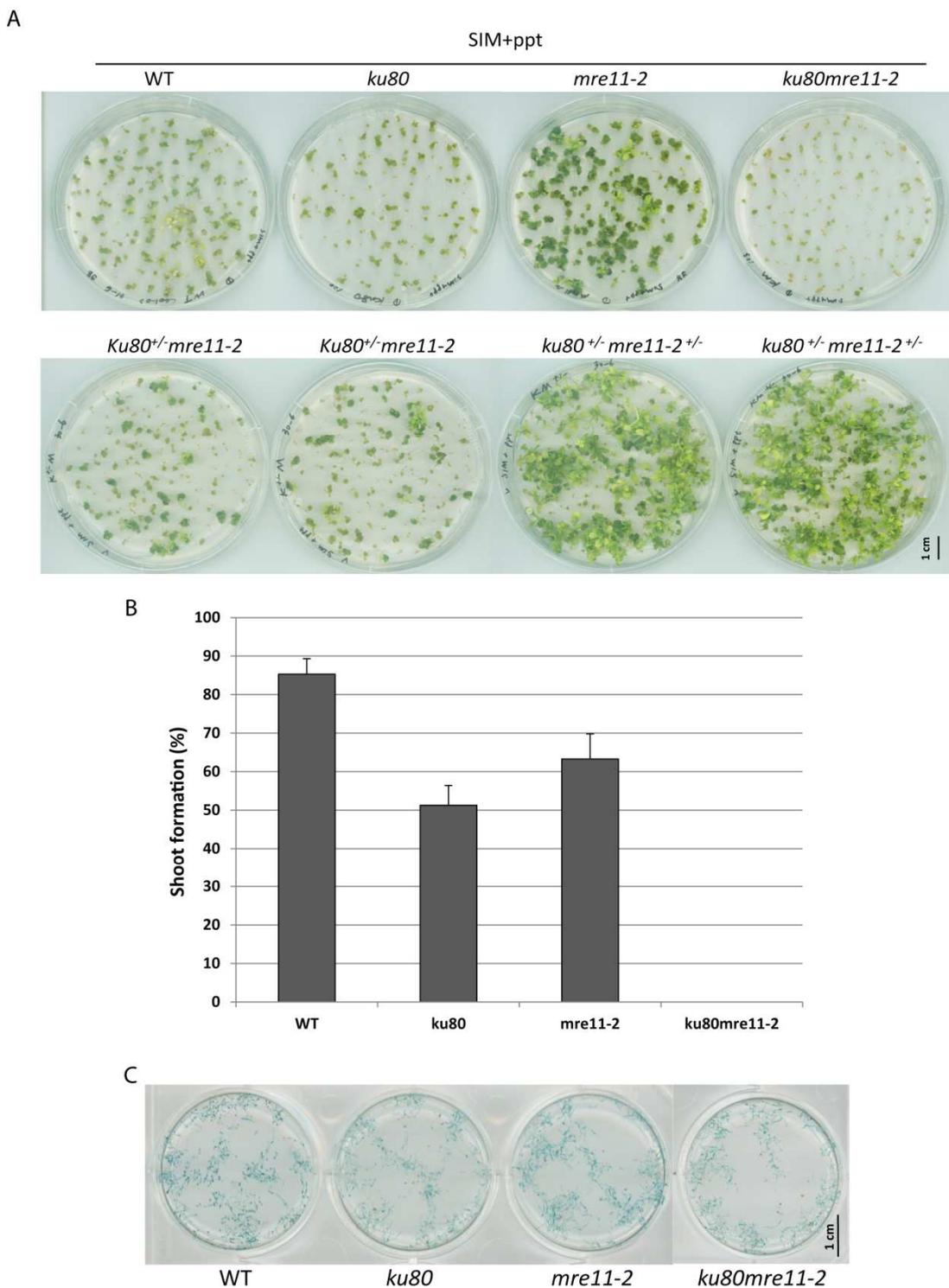


Figure 5. Root transformations of mutants. **(A)** Root segments from wild-type and mutant plants were co-cultivated with *Agrobacterium* strain LBA1100 (pCambia3301) for 48 hours, and transferred to shoot induction medium (SIM) with phosphinothricin selection (ppt). Photographs were taken 4 weeks after cocultivation. **(B)** shoot formation rate, represented by the percentage of green calli with shoot after 6 weeks cultivation on selection plates. **(C)** Root segments from wild-type and mutant plants were stained with X-Gluc overnight after 72 hours cocultivation.

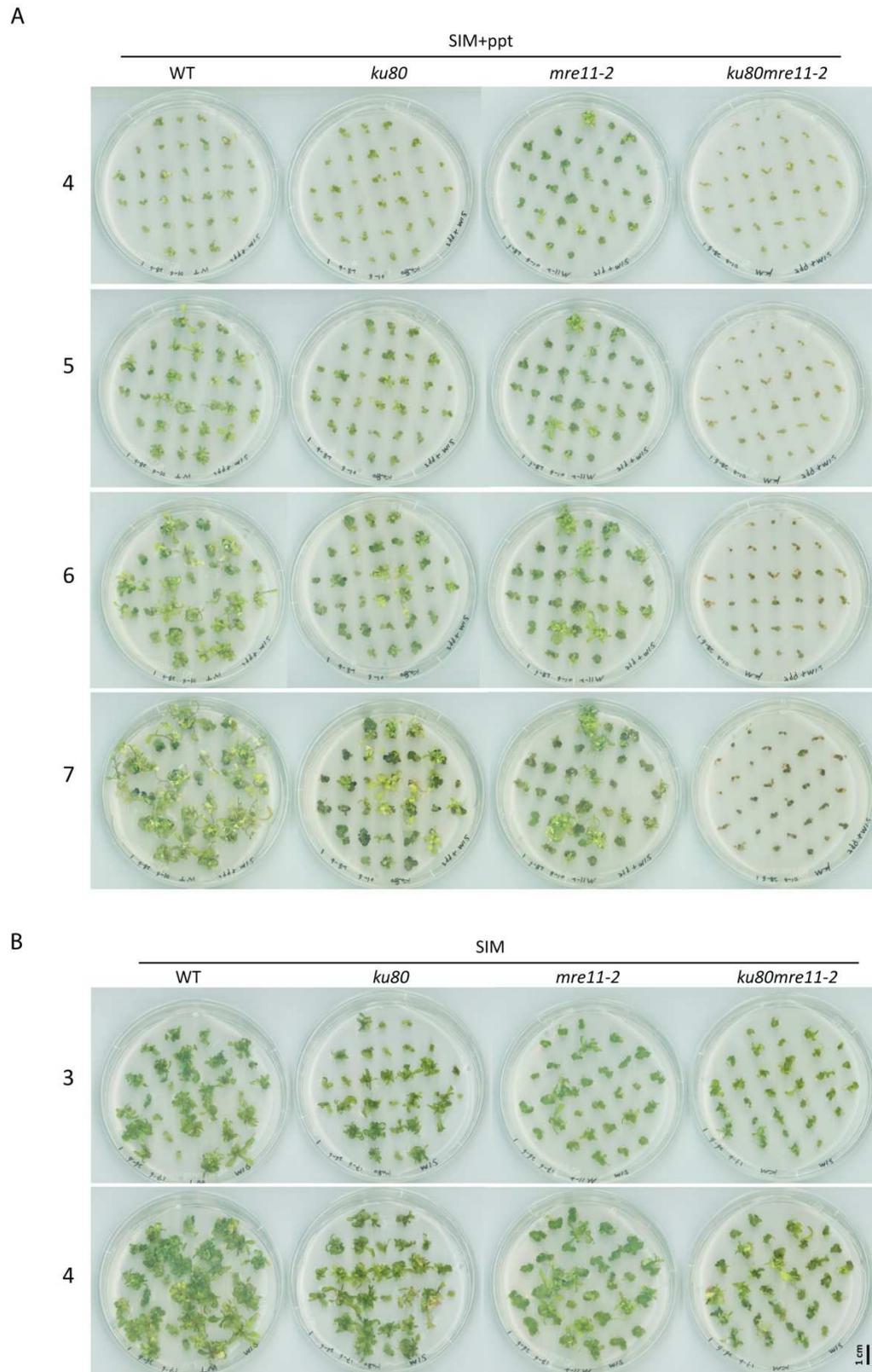


Figure 6. Shoot development on calli from root transformations. **(A)** After 3 weeks growth on shoot induction medium (SIM) with phosphinothricin (ppt) selection, green calli were transferred to fresh shoot induction medium with phosphinothricin. **(B)** After cocultivation with *Agrobacterium*, root segments were transferred to shoot induction medium without selection. After 3 weeks green calli were transferred to fresh shoot induction medium plates without selection. Photographs were taken every week after transfer to fresh plates. The number of 3, 4, 5, 6, 7 indicates the weeks of cultivation.

Discussion

The predicted protein sequence of the Arabidopsis Mre11 orthologue contains a highly conserved N-terminal nuclease domain which is essential for repair in yeast and human, but a less conserved region in the Mre11 C-terminus. Three different T-DNA insertion lines of Arabidopsis *MRE11* have many developmental defects and are sterile suggesting that the functions of Arabidopsis Mre11 are severely compromised in these mutants. The *mre11-2* mutant, however, is fertile and grows normally. The T-DNA insertions in the *mre11-4* and *mre11-2* mutant both leave the N-terminal nuclease domain of the encoded Mre11 protein intact. Our results from two-hybrid analysis demonstrated that the interaction domain necessary for the formation of a Mre11-Rad50 complex is still present in the Mre11-2 protein, but absent in Mre11-4. Mre11 relies for most of its functions on the formation of a Mre11-Rad50-Nbs1 complex explaining the phenotypic differences between the two mutants (Stracker and Petrini 2011). Rad50 binding domains (RBD) have been identified in the C-terminal part of *Pyrococcus.furiosus* Mre11 (Williams *et al.* 2011). Although this C-terminal region is less conserved, the C-terminal region of Mre11 in yeast and humans also have been indicated to be involved in the interaction with Rad50 (Chamankhah and Xiao 1999; Park *et al.* 2011). Our studies also point out that a region in the Arabidopsis Mre11 C-terminus likewise participates in Rad50 binding. Thus, these regions may form a similar three dimensional structure as the RBD of *P.furiosus* Mre11, although they share a low conservation in primary structure.

The *mre11-2* mutant preserves the Rad50 binding activities, grows normally but still shows hypersensitivity to DNA damaging agents. This suggests that either the deletion results in a less stable protein or the C-terminal part of Mre11 has a specific role in Arabidopsis DNA damage signaling and/or DNA repair. The MRN complex is responsible for the recruitment of ATM to sites of damage during the DNA damage response, which highly depends on the C-terminal part of Nbs1 binding to ATM (Falck *et al.* 2005). Arabidopsis *mre11-2atm-2* mutants are sterile probably because of a defect in meiotic repair (Šamanić *et al.* 2013), similar to the phenotype of *nbs1atm* double mutant plants (Waterworth *et al.* 2007), indicating that the MRN complex and ATM kinase have a redundant function in meiotic recombination and absence of the C-terminal part of Mre11 may impair the interaction between Mre11 (Mre11-Rad50 complex) and Nbs1. However, it has been shown that the Nbs1 binding domain of Mre11 localizes in its N-terminal region in yeast and humans. If this is the case in plants as well the deletion of C-terminus in *mre11-2* probably does not affect the formation of the MRN complex, but it may reduce its interaction activities, as the Mre11 C-terminus was reported to have important protein-protein interaction activities in other organisms.

A *ku80mre11-2* double mutant was generated by crossing the single mutants from different ecotypes. This may have affected the outcome of the results, although this is not likely because the heterozygote *ku80+/-mre11-2-/-* and *ku80+/-mre11-2+/-* progeny was used as controls experiments and behaved as expected. We found that the *ku80mre11-2* double mutant was more sensitive to DNA damage than each of the single mutants suggesting that Mre11 may also play a role in b-NHEJ pathway in plant cells and the C-terminal part of Mre11 is important for this function. The subsequent observation of more cell death events in

this double mutant when longer exposed to damage-inducing agents indicates that NHEJ repair pathways are severely disturbed in the concomitant absence of Ku80 and a fully functional Mre11. Beyond its function in activation of the DNA damage response, Mre11 has also been implicated in NHEJ, both in c-NHEJ and also in b-NHEJ that occurs in the absence of c-NHEJ (Zha *et al.* 2009). On the basis of our results we suggest that Mre11 plays an essential role in the b-NHEJ repair pathways in plant cells.

Other evidence supporting a role of Mre11 in b-NHEJ was that the *ku80mre11-2* double mutant was un-transformable by *Agrobacterium*. Using yeast as a model it was shown in our lab that T-DNA integration in yeast (*S.cerevisiae*) is dependent on the non-homologous end joining (NHEJ) pathway of DSB repair, and that proteins such as Ku70, Ku80 and DNA ligase IV are essential for T-DNA integration (van Attikum *et al.* 2001). Inactivation of c-NHEJ in yeast prevented random T-DNA integration, although there is a b-NHEJ pathway in yeast. Putative T-DNA integration by this b-NHEJ pathway may have been beyond the limits of detection. The double strand break repair mechanisms were supposed to be involved in the integration of *Agrobacterium* T-DNA in plants as well. In subsequent studies from other groups and ours on T-DNA integration with Arabidopsis NHEJ mutants T-DNA integrants were invariably obtained, demonstrating that other important factors must be involved in T-DNA integration. The Mre11 might represent such an important protein that plays a critical role in a b-NHEJ pathway involved in T-DNA integration as inactivating both c-NHEJ (Ku80) and b-NHEJ (Mre11) at the same time prevented T-DNA integration in plants. Therefore, Ku80 and Mre11 apparently control two different pathways, each of which is not essential, but together are responsible for all T-DNA integration in plants. A recent study showed that the Arabidopsis Pol θ ortholog Tebichi (Teb) is essential for T-DNA integration (van Kregten *et al.*, 2016). The plant pol θ is able to extend minimally paired 3' ends between the T-DNA left border and the plant genome. However, the mechanism of the attachment of 5' right border of T-DNA to the plant genome remains unclear. Combined with our results, Ku80 and Mre11 probably function together with Pol θ in T-DNA integration in plants. One possibility is that Ku80 or Mre11 are responsible for the capturing T-DNA right border to plant genome.

Acknowledgements

This work was supported by a grant from the China Scholarship Council (CSC grant no. 2011699040).

Table 1. Primers for yeast two-hybrid assays.

Primer	Gene description	Sequence ^a
M11ABDF	At5g54260, Mre11, full length, Fw	5'-TGGAATTCATGTCTAGGGAGGATTTTAGTGA-3'
M11ABDR	At5g54260, Mre11, full length, Rev	5'-CGCTGCAGTTATCTTCTTAGAGCTCCATAGTTC-3'
M11ABDF	At5g54260, Mre11, 1-529 aa, Fw	5'-TGGAATTCATGTCTAGGGAGGATTTTAGTGA-3'
M11NBDR	At5g54260, Mre11, 1-529 aa, Rev	5'-CGCTGCAGTGCTTCCTTTTGTCAAATTCCTCTG-3'
M11ABDF	At5g54260, Mre11, 1-499 aa, Fw	5'-TGGAATTCATGTCTAGGGAGGATTTTAGTGA-3'
M114BDR	At5g54260, Mre11, 1-499 aa, Rev	5'-CGCTGCAGCTCTAAGCACTCTCCCACTTTA-3'
M11ABDF	At5g54260, Mre11, 1-510 aa, Fw	5'-TGGAATTCATGTCTAGGGAGGATTTTAGTGA-3'
M11xBDR	At5g54260, Mre11, 1-510 aa, Rev	5'-CGCTGCAGAGTGGGTCGAGTGGACCTATCTTT-3'
M11CBDF	At5g54260, Mre11, 530-720 aa, Fw	5'-TGGAATTCAGTGGCATCGCAATGCTTCGTTTC-3'
M11ABDR	At5g54260, Mre11, 530-720 aa, Rev	5'-CGCTGCAGTTATCTTCTTAGAGCTCCATAGTTC-3'
R50BDF	At2g31970, Rad50, full length, Fw	5'-GGCCATGGAGATGAGTACGGTCGATAAAAATG-3'
R50BDR	At2g31970, Rad50, full length, Rev	5'-TTGGATCCCTCAATCAAAGATCTCTTGGGCCT-3'
M11AADF	At5g54260, Mre11, full length, Fw	5'-CGGAATTCGAATGTCTAGGGAGGATTTTAGTG-3'
M11AADR	At5g54260, Mre11, full length, Rev	5'-CGCTCGAGTTATCTTCTTAGAGCTCCATA-3'
M11AADF	At5g54260, Mre11, 1-529 aa, Fw	5'-CGGAATTCGAATGTCTAGGGAGGATTTTAGTG-3'
M11NADR	At5g54260, Mre11, 1-529 aa, Rev	5'-CGCTCGAGTGCTTCCTTTTGTCAAATTCCTCTG-3'
M11AADF	At5g54260, Mre11, 1-499 aa, Fw	5'-CGGAATTCGAATGTCTAGGGAGGATTTTAGTG-3'
M114ADR	At5g54260, Mre11, 1-499 aa, Rev	5'-CGCTCGAGCTCTAAGCACTCTCCCACTTTA-3'
M11AADF	At5g54260, Mre11, 1-510 aa, Fw	5'-CGGAATTCGAATGTCTAGGGAGGATTTTAGTG-3'
M11xADR	At5g54260, Mre11, 1-510 aa, Rev	5'-CGCTCGAGAGTGGGTCGAGTGGACCTATCTTT-3'
R50AADF	At2g31970, Rad50, full length, Fw	5'-CGGGATCCGAATGAGTACGGTCGATAAAAATGTTG-3'
R50AADR	At2g31970, Rad50, full length, Rev	5'-TGCTCGAGTCAATCAAAGATCTCTTGGGCCTCG-3'
N1AADF	At3g02680, Nbs1, full length, Fw	5'-TGGAATTCGAATGGTTTGGGGTCTCTTTCCCG-3'
N1AADR	At3g02680, Nbs1, full length, Rev	5'-CGCTCGAGTCAACTTCCAGAGAGAAACCCGCG-3'

^aRestriction sites are underlined.

References

- van Attikum, H., P. Bundock, and P. J. J. Hooykaas, 2001 Non-homologous end-joining proteins are required for *Agrobacterium* T-DNA integration. *EMBO J.* 20: 6550–6558.
- Boulton, S. J., and S. P. Jackson, 1996 *Saccharomyces cerevisiae* Ku70 potentiates illegitimate DNA double-strand break repair and serves as a barrier to error-prone DNA repair pathways. *EMBO J.* 15: 5093–5103.
- Bundock, P., and P. Hooykaas, 2002 Severe developmental defects, hypersensitivity to DNA-damaging agents, and lengthened telomeres in *Arabidopsis* MRE11 mutants. *Plant Cell* 14: 2451–2462.
- Ceccaldi, R., B. Rondinelli, and A. D. D'Andrea, 2016 Repair pathway choices and consequences at the double-strand break. *Trends Cell Biol.* 26: 52–64.
- Chamankhah, M., and W. Xiao, 1999 Formation of the yeast Mre11-Rad50-Xrs2 complex is correlated with DNA repair and telomere maintenance. *Nucleic Acids Res.* 27: 2072–2079.
- Charbonnel, C., M. E. Gallego, and C. I. White, 2010 Xrcc1-dependent and Ku-dependent DNA double-strand break repair kinetics in *Arabidopsis* plants. *Plant J.* 64: 280–290.
- Chen, L., K. Trujillo, W. Ramos, P. Sung, and A. E. Tomkinson, 2001 Promotion of Dnl4-Catalyzed DNA end-joining by the Rad50/Mre11/Xrs2 and Hdf1/Hdf2 complexes. *Mol. Cell* 8: 1105–

- 1115.
- Clough, S. J., and A. F. Bent, 1998 Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16: 735–743.
- Deriano, L., and D. B. Roth, 2013 Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. *Annu. Rev. Genet.* 47: 433–455.
- Falck, J., J. Coates, and S. P. Jackson, 2005 Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage. *Nature* 434: 605–611.
- Gallego, M. E., J.-Y. Bleuyard, S. Daoudal-Cotterell, N. Jallut, and C. I. White, 2003 Ku80 plays a role in non-homologous recombination but is not required for T-DNA integration in *Arabidopsis*. *Plant J.* 35: 557–565.
- Gallego, M. E., M. Jeanneau, F. Granier, D. Bouchez, N. Bechtold *et al.*, 2001 Disruption of the *Arabidopsis* RAD50 gene leads to plant sterility and MMS sensitivity. *Plant J.* 25: 31–41.
- James, P., J. Halladay, and E. A. Craig, 1996 Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* 144: 1425–1436.
- Jia, Q., P. Bundock, P. J. J. Hooykaas, and S. de Pater, 2012 *Agrobacterium tumefaciens* T-DNA integration and gene targeting in *Arabidopsis thaliana* non-homologous end-joining mutants. *J. Bot.* 2012: 1–13.
- Jia, Q., A. Den Dulk-Ras, H. Shen, P. J. J. Hooykaas, and S. de Pater, 2013 Poly(ADP-ribose)polymerases are involved in microhomology mediated back-up non-homologous end joining in *Arabidopsis thaliana*. *Plant Mol. Biol.* 82: 339–351.
- van Kregten M., S. de Pater, R. Romeijn, R. van Schendel, P. Hooykaas and M. Tijsterman, 2016 T-DNA integration in plants results from Polymerase Theta-mediated DNA repair. *Nature plants* 2:1-6.
- Ma, J.-L., E. M. Kim, J. E. Haber, and S. E. Lee, 2003 Yeast Mre11 and Rad1 proteins define a Ku-independent mechanism to repair double-strand breaks lacking overlapping end sequences. *Mol. Cell. Biol.* 23: 8820–8828.
- Park, Y. B., J. Chae, Y. C. Kim, and Y. Cho, 2011 Crystal structure of human Mre11: Understanding tumorigenic mutations. *Structure* 19: 1591–1602.
- Park, S. Y., Z. Vaghchhipawala, B. Vasudevan, L. Y. Lee, Y. Shen *et al.*, 2015 *Agrobacterium* T-DNA integration into the plant genome can occur without the activity of key non-homologous end-joining proteins. *Plant J.* 81: 934–946.
- Puizina, J., J. Siroky, P. Mokros, D. Schweizer, and K. Riha, 2004 Mre11 deficiency in *Arabidopsis* is associated with chromosomal instability in somatic cells and Spo11-dependent genome fragmentation during meiosis. *Plant Cell* 16: 1968–1978.
- Rass, E., A. Grabarz, I. Plo, J. Gautier, P. Bertrand *et al.*, 2009 Role of Mre11 in chromosomal nonhomologous end joining in mammalian cells. *Nat. Struct. Mol. Biol.* 16: 819–824.
- Šamanić, I., J. Simunić, K. Riha, and J. Puizina, 2013 Evidence for distinct functions of MRE11 in *Arabidopsis* meiosis. *PLoS One* 8 (10): e78760.
- Stracker, T. H., and J. H. J. Petrini, 2011 The MRE11 complex: starting from the ends. *Nat. Rev. Mol. Cell Biol.* 12: 90–103.
- Vergunst, A. C., B. Schrammeijer, A. den Dulk-Ras, C. M. T. de Vlaam, T. J. G. Regensburg-Tuïnk *et al.*, 2000 VirB/D4-dependent protein translocation from *Agrobacterium* into plant cells. *Science* 290: 979–982.

- Waterworth, W. M., C. Altun, S. J. Armstrong, N. Roberts, P. J. Dean *et al.*, 2007 NBS1 is involved in DNA repair and plays a synergistic role with ATM in mediating meiotic homologous recombination in plants. *Plant J.* 52: 41–52.
- Weijers, D., M. Franke-van Dijk, R. J. Vencken, A. Quint, P. Hooykaas *et al.*, 2001 An Arabidopsis Minute-like phenotype caused by a semi-dominant mutation in a RIBOSOMAL PROTEIN S5 gene. *Development* 128: 4289–4299.
- Williams, G. J., R. S. Williams, J. S. Williams, G. Moncalian, A. S. Arvai *et al.*, 2011 ABC ATPase signature helices in Rad50 link nucleotide state to Mre11 interface for DNA repair. *Nat. Struct. Mol. Biol.* 18: 423–431.
- Wood, R. D., and S. Doublé, 2016 DNA polymerase θ (POLQ), double-strand break repair, and cancer. *DNA Repair (Amst)*. 44: 22–32.
- Xie, A., A. Kwok, and R. Scully, 2009 Role of mammalian Mre11 in classical and alternative nonhomologous end joining. *Nat. Struct. Mol. Biol.* 16: 814–818.
- Zha, S., C. Boboila, and F. W. Alt, 2009 Mre11: roles in DNA repair beyond homologous recombination. *Nat. Struct. Mol. Biol.* 16: 798–800.

Chapter 5

Sequence-specific nuclease-induced double strand break repair in Arabidopsis non-homologous end-joining mutants

Hexi Shen*, Gary D Strunks*, Bart J. P. M. Klemann, Paul J. J. Hooykaas, Sylvia de Pater

Department of Molecular and Developmental Genetics, Institute of Biology, Leiden University, Leiden, 2333 BE, The Netherlands

Part of this work is published as

H. Shen*, G. D. Strunks*, B. J. P. M. Klemann, P. J. J. Hooykaas and S. de Pater. (2017) CRISPR/Cas9-induced double-strand break repair in Arabidopsis nonhomologous end-joining mutants. *G3 Genes Genomes genetics* doi: 10.1534/g3.116.035204.

Abstract

Double-strand breaks (DSBs) are one of the most harmful DNA lesions. Cells utilize two main pathways for DSB repair: homologous recombination (HR) and non-homologous end joining (NHEJ). NHEJ can be subdivided into the KU-dependent classical NHEJ (c-NHEJ) and the more error-prone KU-independent backup-NHEJ (b-NHEJ) pathways, involving the poly (ADP-ribose) polymerases (PARPs). However, in absence of these factors, cells still seem able to adequately maintain genome integrity, suggesting the presence of other b-NHEJ repair factors or pathways independent from KU and PARPs. The outcome of DSB repair by NHEJ pathways can be investigated by using artificial sequence-specific nucleases such as TALENs and CRISPR/Cas9 to induce DSBs at a target of interest. Here, we used TALEN and CRISPR/Cas9 for DSB induction at the *Arabidopsis cruciferin 3* (*CRU3*) and *protoporphyrinogen oxidase* (*PPO*) genes. DSB repair outcomes via NHEJ were analysed using footprint analysis in wild-type plants and plants deficient in the c-NHEJ pathway (*ku80*), the b-NHEJ pathway (*parp1parp2*) or both (*ku80parp1parp2*). We found that larger deletions of more than 20 bp predominated after DSB repair in *ku80* and *ku80parp1parp2* mutants, corroborating with a role of KU in preventing DSB end resection. Deletion lengths did not significantly differ between *ku80* and *ku80parp1parp2* mutants, suggesting that a KU and PARP-independent b-NHEJ mechanism becomes active in these mutants. Furthermore, microhomologies and templated insertions were observed at the repair junctions in the wild type and all mutants. Since these characteristics are hallmarks of Polymerase θ -mediated DSB repair, we suggest a possible role for this recently discovered polymerase in DSB repair in plants.

Introduction

Double-strand breaks (DSBs) are one of the most lethal forms of DNA damage. DSBs can occur during normal cellular metabolism or can be induced by external factors, and highly threaten genomic integrity and cell survival (Deriano and Roth 2013). To repair DSBs, cells have two main pathways: Homologous Recombination (HR) and Non-Homologous End-Joining (NHEJ). Both of them function together to maintain genome integrity. NHEJ is the predominant pathway in higher eukaryotes and repair may lead to mutations at break sites, such as deletions, insertions and substitutions. At least two NHEJ pathways have been identified: the classic NHEJ pathway (c-NHEJ) and the backup-NHEJ pathway (b-NHEJ) also called alternative-NHEJ (a-NHEJ) or microhomology-mediated end-joining (MMEJ). The c-NHEJ is initiated by the recognition and binding of the KU heterodimer, consisting of KU70 and KU80 subunits, to DSBs (Walker *et al.* 2001). Once bound to a DSB, the KU heterodimer serves as a scaffold to recruit other c-NHEJ factors to the broken ends and promotes end-joining. Because KU is the key component of the c-NHEJ pathway, this pathway is also called KU-dependent NHEJ. In the absence of KU, other factors gain entry to the DSB site for repair by backup pathways. Although the b-NHEJ pathway was defined by involving multiple components, including poly (ADP-ribose) polymerase 1 (PARP1) the precise mechanism is still not clear (Wang *et al.* 2006). Furthermore, recently PARP1 was shown to be involved in repair of DSBs also in the presence of KU (Luijsterburg *et al.* 2016).

Nowadays, DSBs can be induced artificially at specific sites in the genome by sequence-specific artificial nucleases, which can be used to study DSB repair. These induced DSBs will be mainly repaired via NHEJ, which may lead to targeted mutagenesis. When repair restores the target site for the nuclease, the sequence will be cut again in the continuous presence of the nuclease. This cycle of cutting and repair will continue until incorrect repair destroys the target site. When a homologous sequence such as a sister chromatid, is present, DSB repair may also occur via HR, but this will inevitably also lead to restoration of the target site. A repair template without the target site may be provided by transformation or pre-inserted in the genome, and, when used for repair, lead to gene targeting (GT) (Voytas 2013; Puchta and Fauser 2013a). The current genome editing tool kit comprises four classes of engineered nucleases: modified meganucleases, zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and the CRISPR/Cas9 (clustered regularly interspaced short palindromic repeat/CRISPR-associated 9) system (Voytas 2013; Puchta and Fauser 2013a), of which the TALENs and CRISPR/Cas9 system are the most easy and straightforward to use.

TALENs consist of a DNA binding domain, derived from proteins produced by plant pathogens of the genus *Xanthomonas*, fused to the *FokI* nuclease domain, and they can cleave DNA as a dimer (Christian *et al.* 2010). The DNA binding domain of TALENs consists of an array of 13 – 28 repeats. Each repeat consists of 34 highly conserved amino acids, of which only the amino acid residues at position 13 and 14, also called repeat variable diresidues (RVDs), vary and can specifically bind to one of each of the four DNA bases (Voytas 2013). In this way, an array of TAL effector repeats has a one to one correspondence with the DNA sequence it binds. A TALEN pair can recognize 26 – 56 bp, a sequence length which can be considered unique in higher eukaryotic genomes. The efficiency of DSB induction by

TALENs is also influenced by the spacer length between the binding sites of the two TAL arrays (Christian *et al.* 2010).

CRISPR/Cas9 is the most recent addition to the genome editing tool box (Jinek *et al.* 2012). It is derived from an adaptive immune system present in bacteria and archaea, where it serves in degrading invading foreign plasmid or viral DNA. The type II CRISPR genomic locus encodes the Cas9 ('CRISPR-associated 9') endonuclease, which can form a complex with two short RNA molecules: CRISPR RNA (crRNA) and trans-activating crRNA (tracrRNA), which guide the Cas9 protein to a DNA sequence of interest where it can induce a DSB through cleavage of the two DNA strands with its two nuclease domains (RuvC-like domain I and HNH motif). It was shown that the crRNA and tracrRNA can be fused into a chimeric single-guide RNA (sgRNA) comprising the functions of both precursor RNAs (Jinek *et al.* 2012). A sgRNA can be assembled to target any DNA sequence, with the prerequisite that a protospacer adjacent motif (PAM) sequence of NGG flanking the 3' end of the sgRNA target sequence is present, which interacts with the Cas9 PAM interacting domain (PI domain) (Nishimasu *et al.* 2014; Jinek *et al.* 2014). The direct RNA-DNA recognition of the CRISPR/Cas9 system has the advantage that only the sequence of the sgRNA needs to be changed if new loci have to be targeted, instead of the more laborious assembly of new TAL effector arrays when using TALENs. Furthermore, CRISPR/Cas9 conveniently allows DSB induction at multiple targets by simultaneously expressing multiple sgRNAs in combination with the Cas9 protein (Cong *et al.* 2013).

Previous studies already demonstrated the feasibility of DSB-mediated targeted mutagenesis at artificial and endogenous loci in plants using ZFNs, TALENs and the CRISPR/Cas9 system (Puchta and Fauser 2013b). Here, we used TALENs and CRISPR/Cas9 for DSB-mediated targeted mutagenesis at the *Arabidopsis cruciferin 3* (*CRU3*) and *protoporphyrinogen oxidase* (*PPO*) genes. CRISPR/Cas9 nucleases were expressed in mutants in the c-NHEJ or b-NHEJ DNA repair pathways and a combination of both. Footprint analysis in whole seedlings in the wild type and each of the three mutant genotype backgrounds (*ku80*, *parp1parp2* and *ku80parp1parp2* mutants) demonstrated that key factors of NHEJ can affect the outcomes of targeted mutagenesis.

Materials and Methods

Plant material

The *ku80* (SALK_016627), *parp1* (GABI-Kat Line 692A05) and *parp2* (SALK_140400) T-DNA insertion lines (ecotype Col-0), the *parp1parp2* double mutant and *ku80parp1parp2* triple mutant were described previously (Jia *et al.* 2013). More information about these lines can be found at <http://signal.salk.edu/cgi-bin/tdnaexpress> (Alonso *et al.* 2003).

TALEN and CRISPR/Cas9 vector construction and plant transformation

TALEN was designed and assembled with the Golden Gate Kit (AddGene) as described (Cermak *et al.* 2011). Arrays of repeats with the corresponding RVDs (**Table 1**) for the DNA binding sequences of the *CRU3* target were assembled together in vector pZHY500 (TALEN-CRU-1-left) or pZHY501 (TALEN-CRU-1-right). Individual TALE repeats were cloned into vector pZHY013 using *XbaI* and *BamHI* (TALEN-CRU-1-left) or *NheI* and *BglII* (TALEN-CRU-1-right). Subsequently, the DNA fragment containing the TALEN pair was cloned in the binary vector pMDC32 (Curtis and Grossniklaus 2003), a 35S T-DNA expression vector, via a Gateway LR reaction to create TALEN-CRU-1 (pSDM3906).

For the CRISPR/Cas9 constructs, oligo's SP509 and SP510 (*CRU3* target) and SP512 and SP513 (*PPO* target) (**Table 2**) were annealed and cloned in *BbsI* digested pEn-Chimera (Fauser *et al.* 2014). Subsequently, gene coding sgRNAs were cloned in pDE-pUbi-Cas9 (Fauser *et al.* 2014) by a Gateway LR reaction, resulting in Cas9-PPO (pSDM3905) and Cas9-CRU (pSDM3903), respectively.

Plant binary vectors were introduced into *Agrobacterium tumefaciens* strain AGL1 by electroporation. *Arabidopsis thaliana* plants of the Col-0 ecotype (wild type, *ku80*, *parp1parp2*, *ku80parp1parp2*) were transformed with T-DNAs containing nuclease expression cassettes, using the floral dip method (Clough and Bent 1998). T1 seeds were grown on MA solid medium without sucrose, supplemented with timentin (100 µg/mL), nystatin (100 µg/mL) and antibiotics for T-DNA selection: 15 µg/mL hygromycin for TALEN; 15 µg/mL phosphinothricin for CRISPR/Cas9.

DNA isolation and footprint analysis

T2 seeds were germinated on ½ MS with T-DNA selection, the seedlings were disrupted to a powder under liquid N₂ in a TissueLyser (Retch, Haan, Germany). Genomic DNA was extracted by the CTAB method (De Pater *et al.* 2009). For predigestion, one µg of genomic DNA was digested with *DdeI* (for TALEN-CRU-1 analysis), *PstI* (Cas9-CRU analysis) or *FauI* (for Cas9-PPO analysis) overnight and precipitated. Undigested or predigested DNA was used for PCR with Phusion polymerase (Thermo Scientific) to amplify the nucleases target sites, followed by digestion of the PCR products with *DdeI*, *PstI* or *FauI* and separated in agarose gels. PCR primers are shown in **Table 2**. Primers SP491 and SP492 were used for amplification of the TALEN-CRU-1 target region, primers SP245 and SP248 were used for the Cas9-CRU target region and primers SP392 and SP538 were used for the Cas9-PPO target region. The resistant fragments were isolated from gel and cloned into pJet1.2 (Thermo Scientific) and sequenced by MacroGen Europe (Amsterdam, The Netherlands). Identical sequences in the same line were considered as one mutagenesis event since they might have

been preselected from PCR amplification. Two-tailed Mann-Whitney tests were performed for statistical analysis of deletion- and insertion lengths .

Measurements of the mutation rate

To estimate the rate of TALEN- and Cas9-induced mutations the target sites were amplified using undigested genomic DNA. PCR products were digested with the appropriate restriction enzymes and analysed on agarose gels. The intensity of bands was quantified using ImageJ software. The mutation rate was calculated by dividing the intensity of the digest-resistant band by the total intensity of all bands in a given lane (Nekrasov *et al.* 2013).

High resolution melting

High resolution melting (HRM) analyses were performed on PCR clones from T2 seedlings of wild-type lines Cas9-CRU #2 and Cas9-PPO #7 using Precision Melt Supermix (Bio-Rad), containing EvaGreen saturated dye, and the Bio-Rad C1000 Touch thermal cycler (Bio-Rad). Melt curves were analyzed using the Bio-Rad Precision Melt Analysis software. For the CRU target primers SP492 and SP563 were used *and for the PPO target primers SP560 and SP561 (Table 2)*. Samples that showed melt curve differences were sequenced by Macrogen Europe (Amsterdam, The Netherlands).

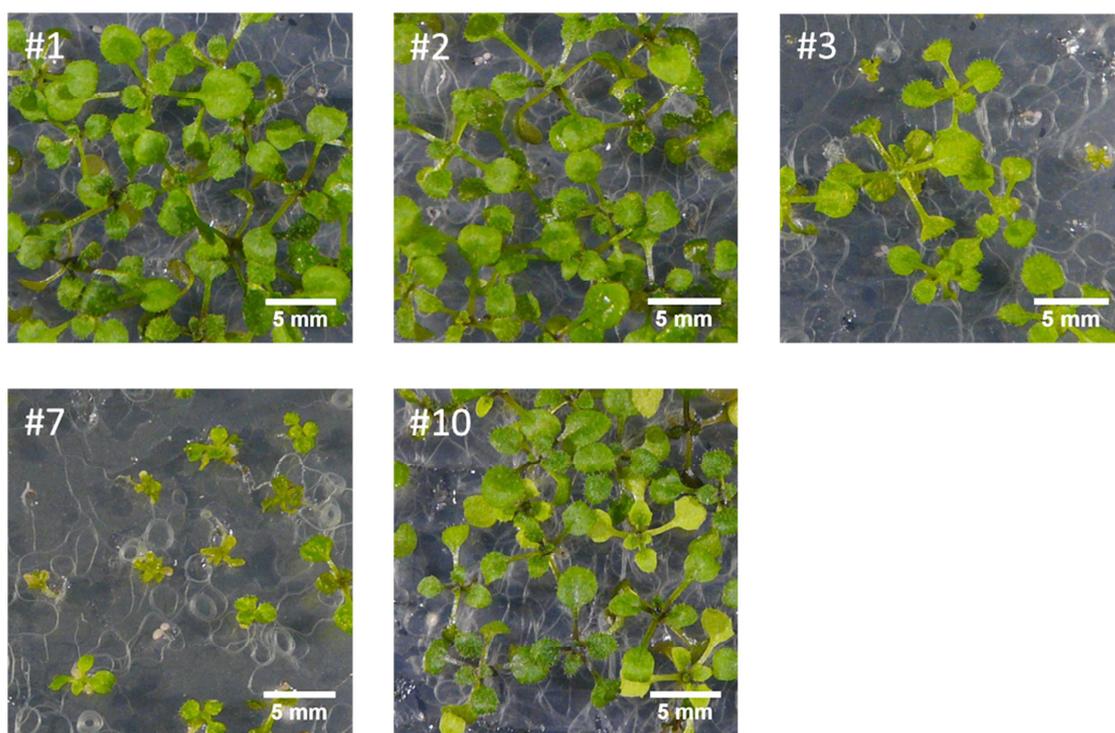


Figure 1. Phenotypes of 16 day-old T2 seedlings of 5 independent Cas9-*PPO* transformants. A stunted growth phenotype is observed in some seedlings of line #3 and #7. The other lines have a phenotype similar to wild-type.

Results

DSB-mediated mutagenesis by TALEN and CRISPR/Cas9 at the CRU3 and PPO loci

In order to investigate repair of induced DSBs, several sequence-specific nucleases were designed and expressed in *Arabidopsis*. Wild-type plants were transformed with TALEN or CRISPR/Cas9 expression constructs via the *Agrobacterium*-mediated floral dip method (Clough and Bent 1998) and T2 transformants were used for further analysis. Nuclease target sites in the *CRU3* and *PPO* genes were selected. The *CRU3* gene encodes a seed storage protein. The *PPO* gene encodes an essential enzyme that is involved in the final step of chlorophyll biosynthesis, and mutagenesis of the *PPO* gene is therefore toxic to plants. Plants expressing nucleases targeted at *CRU3* showed a phenotype similar to wild-type, but T2 seedlings of some plant lines expressing Cas9-*PPO* showed a stunted growth phenotype indicative of homozygous inactivation of the essential *PPO* gene in many cells (**Figure 1**). To detect mutagenesis caused by nuclease activity and subsequent erroneous NHEJ-mediated DSB repair at the molecular level, genomic DNA from T2 seedlings was analyzed for the presence of NHEJ-induced indels. In order to discriminate DNA molecules with a mutation, PCR products from the region containing the target site were digested with restriction enzymes having a recognition site overlapping or near the DSB site (*DdeI* for TALEN-CRU-1, *PstI* for Cas9-CRU, and *FauI* for Cas9-PPO) (**Figure 2**). Loss of the restriction site as a consequence of erroneous repair resulted in restriction digest-resistant PCR products. After gel electrophoresis, the relative band intensities were measured to estimate the mutation frequency in the target sites (**Figure 3**). Digestion of the PCR products from untransformed wild-type plants only left up to 3% of the material undigested, probably due to incomplete digestion. However, a distinguishably higher fraction of the PCR products from plant lines transformed with either TALEN or CRISPR/Cas9 nucleases were resistant for enzyme digestion. An average mutation rate of 5.5% was found in TALEN-CRU-1 lines, 6.2% in Cas9-CRU lines and 11.2% in Cas9-PPO lines (**Figure 3**).

To get a better insight into the mutations induced by the TALEN and CRISPR/Cas9 nucleases, the restriction digest-resistant PCR products of the *CRU3* and *PPO* targets were cloned and sequenced. Pre-digested genomic DNA was used for PCR to enrich for mutated sequences. Sequencing revealed mainly deletions and some insertions and substitutions in both TALEN and CRISPR/Cas9 lines (**Figure 3D**, **Figure 7**). TALEN-CRU-1 footprints mainly consisted of small deletions ranging from 1 bp to 15 bp. The main reason we detected only small deletions in the TALEN lines may have been the small size of the amplified PCR fragments, due to the presence of additional *DdeI* sites just outside the amplified region. CRISPR/Cas9 lines showed more frequent and generally larger deletions compared to the TALEN lines (**Figure 3D**). Short homologous sequences on either the left or right side flanking the deletion were often also present, suggesting MMEJ may have been involved in DSB repair.

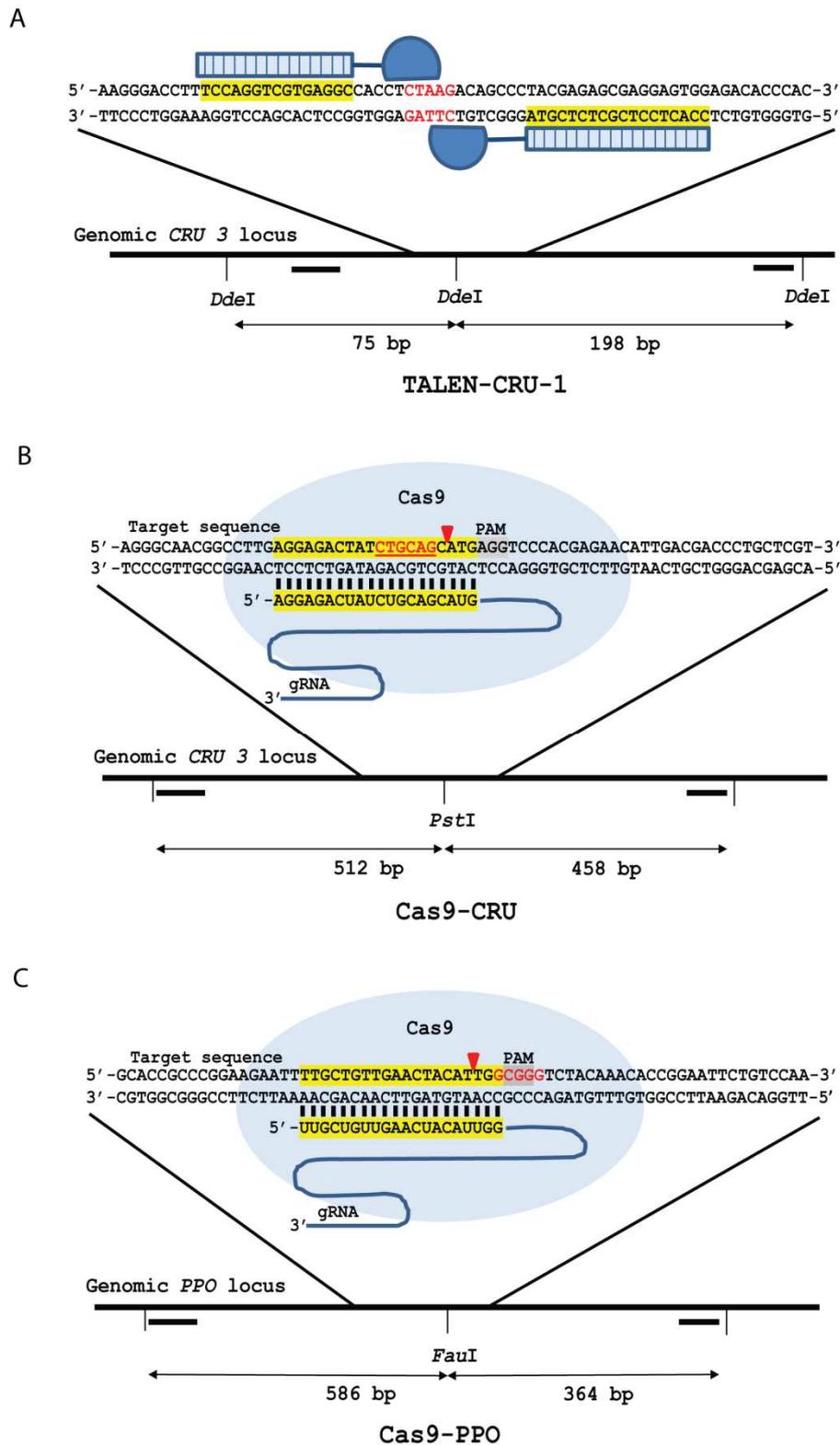


Figure 2. TALEN and CRISPR/Cas9 endonucleases for DSB induction in *CRU3* and *PPO*. TALEN-CRU-1 (A) and Cas9-CRU (B) with their target locus in the *CRU3* locus and Cas9-PPO (C) with its target in the *PPO* locus are shown. TALEN and sgRNA DNA binding sequences are highlighted with yellow, the PAM sequence is highlighted with gray and the *DdeI*, *PstI* and *FauI* restriction sites are shown in red lettering. The primers (—) used to amplify the target regions and the sizes are indicated. Red arrows indicated the position of DSB induction.

In our experimental design for both Cas9-CRU and Cas9-PPO, the *Pst*I and *Fau*I restriction sites are nearby but do not overlap the DSB site. Therefore, the preselection of loss of a restriction site has the caveat of neglecting mutations that occur outside of the restriction site. To get a more precise insight into the DSB-mediated mutations in *CRU3* and *PPO*, we used qPCR followed by high resolution melting (HRM) for footprint analysis. Then indeed, also footprints outside of the restriction site were detected (**Figure 4**). For Cas9-CRU, HRM was performed on 142 PCR clones, which revealed one deletion of 32 bp outside of the *Pst*I site. No footprints inside the *Pst*I site were detected in the remaining 141 clones, indicating a very low mutation frequency in this plant line. For Cas9-PPO, HRM was performed on 48 PCR clones from T2 seedlings of line Cas9-PPO #7, which has a severe phenotype (**Figure 1**). Four different footprints ranging from 1 bp insertion to 5 bp deletions were found outside the *Fau*I site. None of the 48 clones contained wild-type sequences, indicating a high mutation rate in this plant line.

Taken together, these results show that our constructs of TALENS and CRISPR/Cas9 are able to induce mutations at the target sites and that our CRISPR/Cas9 constructs are more efficient than our TALENs. Only four mutations outside the restriction site were detected by the HRM method, indicating that the restriction enzyme assay gives a good estimate of the mutation frequency at our target sequences.

Increased DSB-mediated mutagenesis by CRISPR/Cas9 in c-NHEJ deficient mutants

To compare mutagenesis in wild type and NHEJ-compromised genetic backgrounds, T-DNA insertion lines *ku80*, *parp1parp2* (*p1p2*), and *ku80parp1parp2* (*ku80p1p2*) as described previously (Jia *et al.* 2012; Jia *et al.* 2013) were transformed with TALEN-CRU-1, Cas9-CRU and Cas9-PPO nucleases, and several independent primary transformants were obtained. The target region was PCR amplified using total genomic DNA from several T2 plant lines as a template, followed by restriction enzyme digestion of the PCR product. The relative band intensity was measured to estimate the mutation frequency. The mutation frequency in the TALEN-CRU-1 and Cas9-CRU lines was very low and therefore we did not perform this semi-quantification with these lines. Clear resistant bands were, however, observed in Cas9-PPO lines (**Figure 3C**). Interestingly, the mean values of the mutation frequencies in *ku80* and *ku80p1p2* lines were higher than those in the wild-type and *p1p2* lines. These results indicated that knockout of the c-NHEJ repair pathway increases DSB-mediated mutagenesis by CRISPR/Cas9. Plants appeared normal. Apparently, DSB repair of the induced DSBs was still efficient (but less precise) even in the triple mutant *ku80p1p2*.

Larger deletions are predominant in c-NHEJ deficient mutants

To assess the outcomes of DSB repair at the nucleotide level in wild type and mutant lines, genomic DNA was pre-digested and the resistant bands were purified, cloned and sequenced as described for the wild type. The results showed that there were deletions, insertions and substitutions at the *CRU3* and *PPO* target sites in mutant lines (**Figure 7**). The majority of mutations recovered in the mutant lines were deletions. Substitutions seem to be very rare events based on the sequenced data and these might be PCR artefacts.

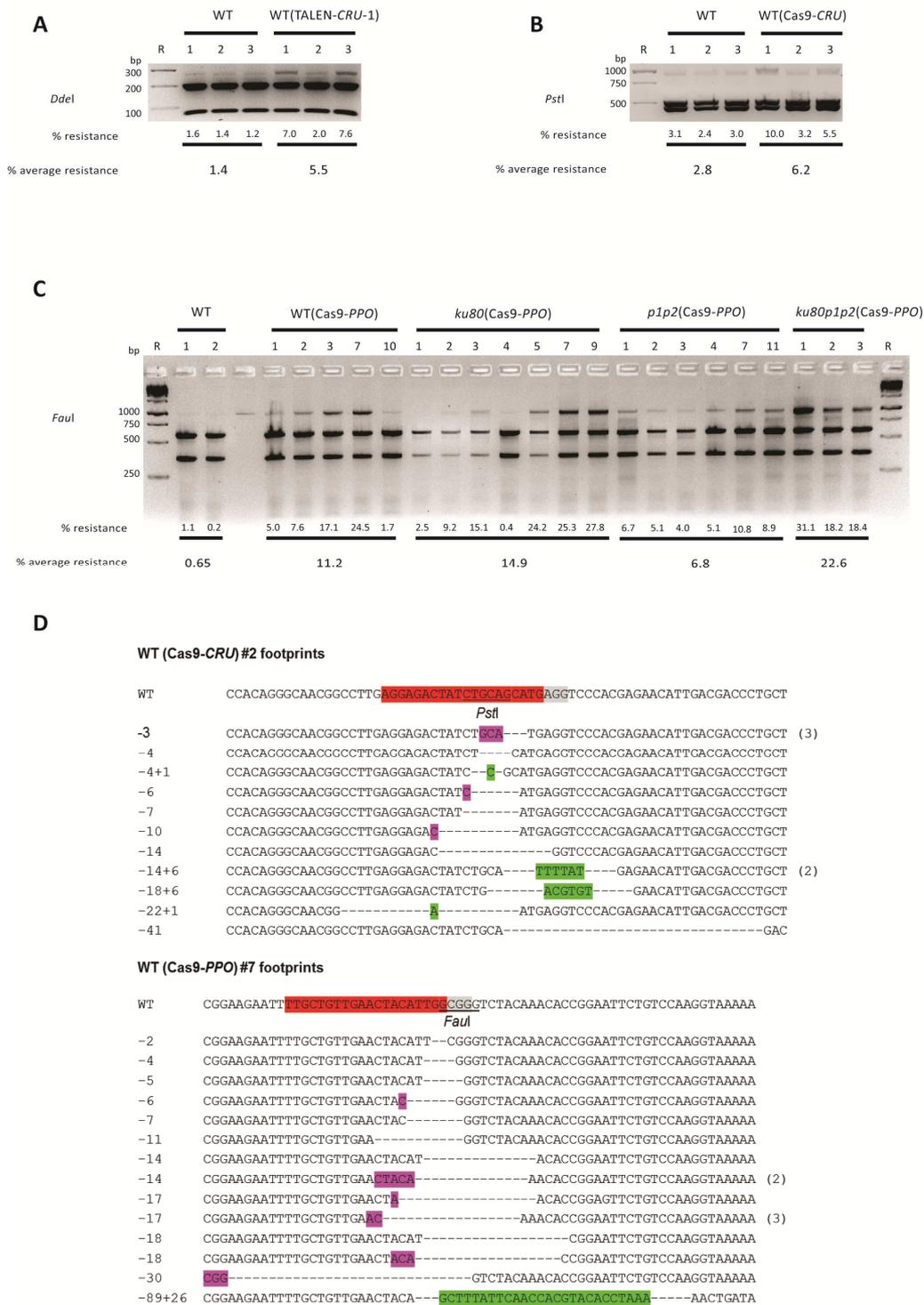
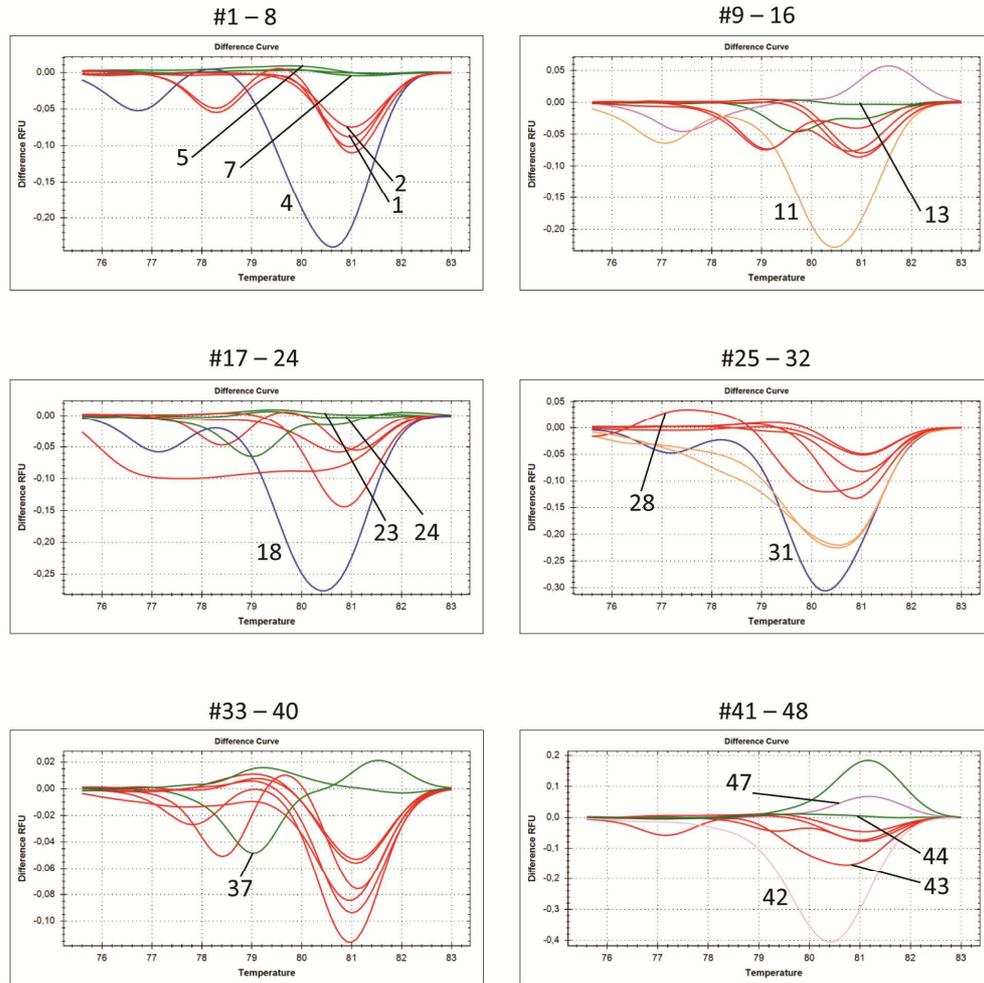


Figure 3. TALEN and CRISPR/Cas9 endonucleases induced mutagenesis. The *CRU3* target site was amplified using undigested genomic DNA from (A) untransformed wild type seedlings and TALEN-CRU-1 T2 transformants and digested with *DdeI* or (B) untransformed wild type seedlings and Cas9-CRU T2 transformants and digested with *PstI*. (C) The *PPO* target site was amplified from untransformed wild type seedlings and Cas9-PPO T2 transformants of wild type and *ku80*, *parp1parp2*, and *ku80parp1parp2* mutant plant lines and digested with *FauI*. R is the size reference (1 kb ladder) and the % resistant bands is shown below the lanes. (D) Sequences of *CRU3* and *PPO* targets from Cas9-CRU transformant #2 and Cas9-PPO transformant #7. The sgRNA protospacer is in red, PAM sequence is in grey, deletions are shown by dashes, insertions are in green, microhomologies used for repair are in purple. Number of multiple clones with the same sequence are shown at the right. Numbers are length of deletions (-) and insertions (+).

A



B

WT	GGAAGAATTTGCTGTTGAACTACATTGGCGGGTCTACAAACACCGGAATTCGTCCAAGGTAAAAACAGC
	<i>FauI</i>
1	GGAAGAATTTGCTGTTGAACTACATTGGCGGGTCTACAAACACCGGAATTCGTCCAAGGTAAAAACAGC +1
2	GGAAGAATTTGCTGTTGAACTACATTGGCGGGTCTACAAACACCGGAATTCGTCCAAGGTAAAAACAGC +1
4	GGAAGAATTTGCTGTTGAACTAC-----AAACACCGGAATTCGTCCAAGGTAAAAACAGC -17
5	GGAAGAATTTGCTGTTGAACTACAT-GGCGGGTCTACAAACACCGGAATTCGTCCAAGGTAAAAACAGC -1
7	GGAAGAATTTGCTGTTGAACTACAT-GGCGGGTCTACAAACACCGGAATTCGTCCAAGGTAAAAACAGC -1
11	GGAAGAATTTGCTGTTGAACTAC-----AAACACCGGAATTCGTCCAAGGTAAAAACAGC -14+1
13	GGAAGAATTTGCTGTTGAACTACAT-GGCGGGTCTACAAACACCGGAATTCGTCCAAGGTAAAAACAGC -1
18	GGAAGAATTTGCTGTTGAACTACATG-----ACCGGAATTCGTCCAAGGTAAAAACAGC -14
23	GGAAGAATTTGCTGTTGAACTACAT-GGCGGGTCTACAAACACCGGAATTCGTCCAAGGTAAAAACAGC -1
24	GGAAGAATTTGCTGTTGAACTACAT-GGCGGGTCTACAAACACCGGAATTCGTCCAAGGTAAAAACAGC -1
28	GGAAGAATTTGCTGTTGAACTACAT-GGCGGGTCTACAAACACCGGAATTCGTCCAAGGTAAAAACAGC -1
31	GGAAGAATTTGCTGTTGAACTACA-----AAACACCGGAATTCGTCCAAGGTAAAAACAGC -14
37	GGAAGAATTTGCTGTTGAACT-----GGCGGGTCTACAAACACCGGAATTCGTCCAAGGTAAAAACAGC -5
42	GGAAGAATTTGCTGTTGAACT-----AAACACCGGAATTCGTCCAAGGTAAAAACAGC -17
43	GGAAGAATTTGCTGTTGAACTACA-----CTACAAACACCGGAATTCGTCCAAGGTAAAAACAGC -9
44	GGAAGAATTTGCTGTTGAACTACAT-GGCGGGTCTACAAACACCGGAATTCGTCCAAGGTAAAAACAGC -1
47	GGAAGAATTTGCTGTTGAACT---TGGCGGGTCTACAAACACCGGAATTCGTCCAAGGTAAAAACAGC -4

Figure 4. HRM analysis of the PPO target. HRM analyses were performed on 48 PCR clones from pool of 10 seedlings of wild-type T2 seedlings of Cas9-*PPO* transformant #7. **(A)** Difference melt curves of samples 1 – 48 measured in relative fluorescence units (RFU). Numbers indicated in the graph refer to the sequences below. **(B)** Sequences of representative *PPO* targets. *PPO* sgRNA protospacer (red), the PAM sequence (gray) and *FauI* restriction site (underlined) are indicated in the WT sequence. Footprints included deletions (dashed lines), insertions (green) and substitutions (blue). Microhomologies used for repair are shown in purple.

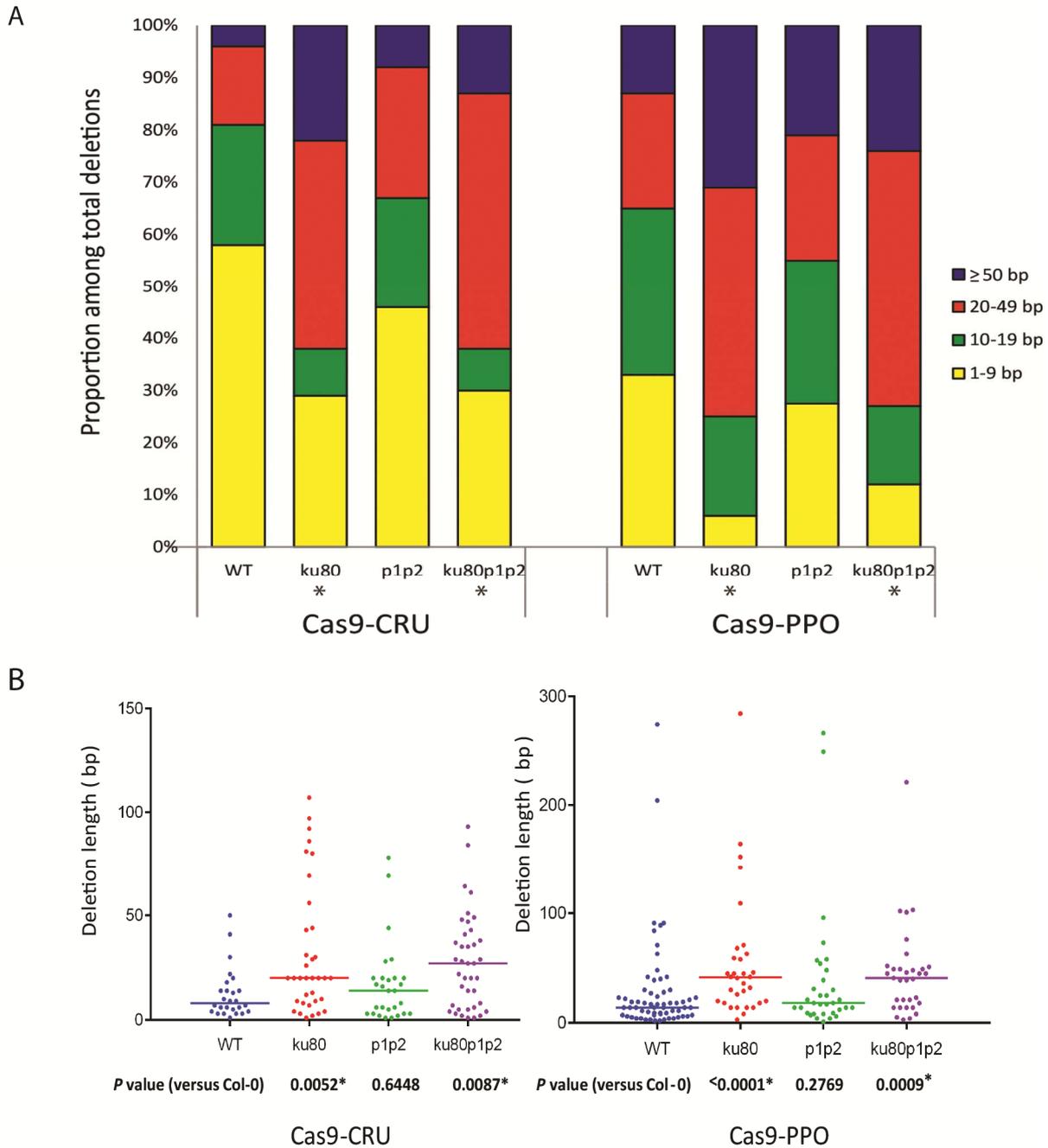


Figure 5. Analysis of deletion length. (A) Distribution of deletion lengths for the indicated genotypes with Cas9 nucleases. (B) Scatter plot of deletion lengths for the indicated genotypes. Median deletion length is indicated at the bar on the graph. P-values are derived from a two-tailed Mann-Whitney test. Asterisk indicates a statistically significant difference from wild-type ($P < 0.05$).

Due to the small size of the amplified target region in the TALEN-CRU-1 lines no large deletions could be found. Therefore we did not include the TALEN-CRU-1 lines in the analysis of deletion length and we examined the length of deletions only from all genotypes expressing Cas9-CRU or Cas9-PPO (Table 3, Figure 5A). In wild-type Cas9-CRU transformants, 57% of deletions were <10 bp and about 23% ranged from 10 to 19 bp, 15% ranged from 20 to 49 bp and 5% were ≥ 50 bp. The *p1p2* mutant lines showed somewhat

longer deletions; about 25% of deletions ranged from 20 to 49 bp and 7% were ≥ 50 bp. Larger deletions were predominant in the *ku80* and *ku80p1p2* mutant lines. In *ku80* lines 62% of the deletions were larger than 20 bp (22% were ≥ 50 bp), and in the *ku80p1p2* lines 61% of the deletions were larger than 20 bp (12% were ≥ 50 bp).

Deletion lengths in Cas9-PPO transformants were also examined (**Figure 5**). Similar to Cas9-CRU, there were no big differences in deletion length between the wild type and *p1p2* mutant lines. In the wild type about 33% of deletions were < 10 bp, 32% ranged from 10 to 19 bp, 22% ranged from 20 to 49 bp and 13% were ≥ 50 bp. In *p1p2* lines 27% of deletions were < 10 bp, 27% ranged from 10 to 19 bp, 24% ranged from 20 to 49 bp and 21% were ≥ 50 bp. Larger deletions of the *PPO* target were however, again predominant in *ku80* and *ku80p1p2* mutant lines. About 75% of deletions in *ku80* lines were larger than 20 bp, and about 73% of deletions in *ku80p1p2* lines were larger than 20 bp.

We performed statistical analysis using a two-tailed Mann-Whitney test, to find out whether the observed differences were significant. For the Cas9-CRU and Cas9-PPO nucleases, comparison of deletion lengths in wild type to *ku80* and *ku80p1p2* lines showed a statistically significant difference, whereas comparison of deletion lengths in wild type to *p1p2* lines did not (**Figure 5B**). These results indicate that imprecise end-joining after loss of the c-NHEJ key component KU80 resulted in substantial increases in deletion length and suggests a shift to a more error-prone repair mechanism of DSB repair in absence of KU80.

Templated insertions in wild type and NHEJ mutants

Insertion events, sometimes accompanied by deletions, were observed at the target loci in TALEN-CRU-1, Cas9-CRU or Cas9-PPO transformed wild type and mutant lines, although less frequently than deletions. The insertion frequency in mutants was comparable to or higher than the insertion frequency in the wild type. More than half of the insertions were smaller than 10 bp. A maximum insertion length of 60 bp was observed. Furthermore, insertion lengths in NHEJ mutant lines were not significantly different from wild type when insertion data of all nucleases was combined, indicating that the insertion mechanism may be independent of KU80 and PARP (**Figure 6A**). In addition, from the combined insertion data of all nucleases it can be deduced that the deletion length of junctions with insertions were significantly larger than junctions without insertions (**Figure 6B**).

Interestingly, many inserted sequences have at least one match to DNA within 100 bp of the repaired DSB. Some insertions have complex compositions with multiple stretches of identity, including reverse complementary homology. These results suggest that polymerase θ may be involved in the repair of these DSBs (Roerink *et al.* 2014). Another signature of Pol θ -mediated DSB repair is the presence of sequence identity between the 3' end that generated the junction (the primer) and the sequence immediately upstream of the template that is used for DNA synthesis. Such sequence identity is present in about 50 % of the inserted sequences (**Figure 6C**). The *ku80* and *ku80p1p2* mutant lines appeared to have more templated-insertion events than wild type and *p1p2* lines, although such insertions were found in all four genotypes (**Table 4**). Therefore, the templated insertions probably resulted from a KU80- and PARPs-independent alternative end-joining mechanism, such as that mediated by the recently discovered Pol θ (van Kregten *et al.* 2016).

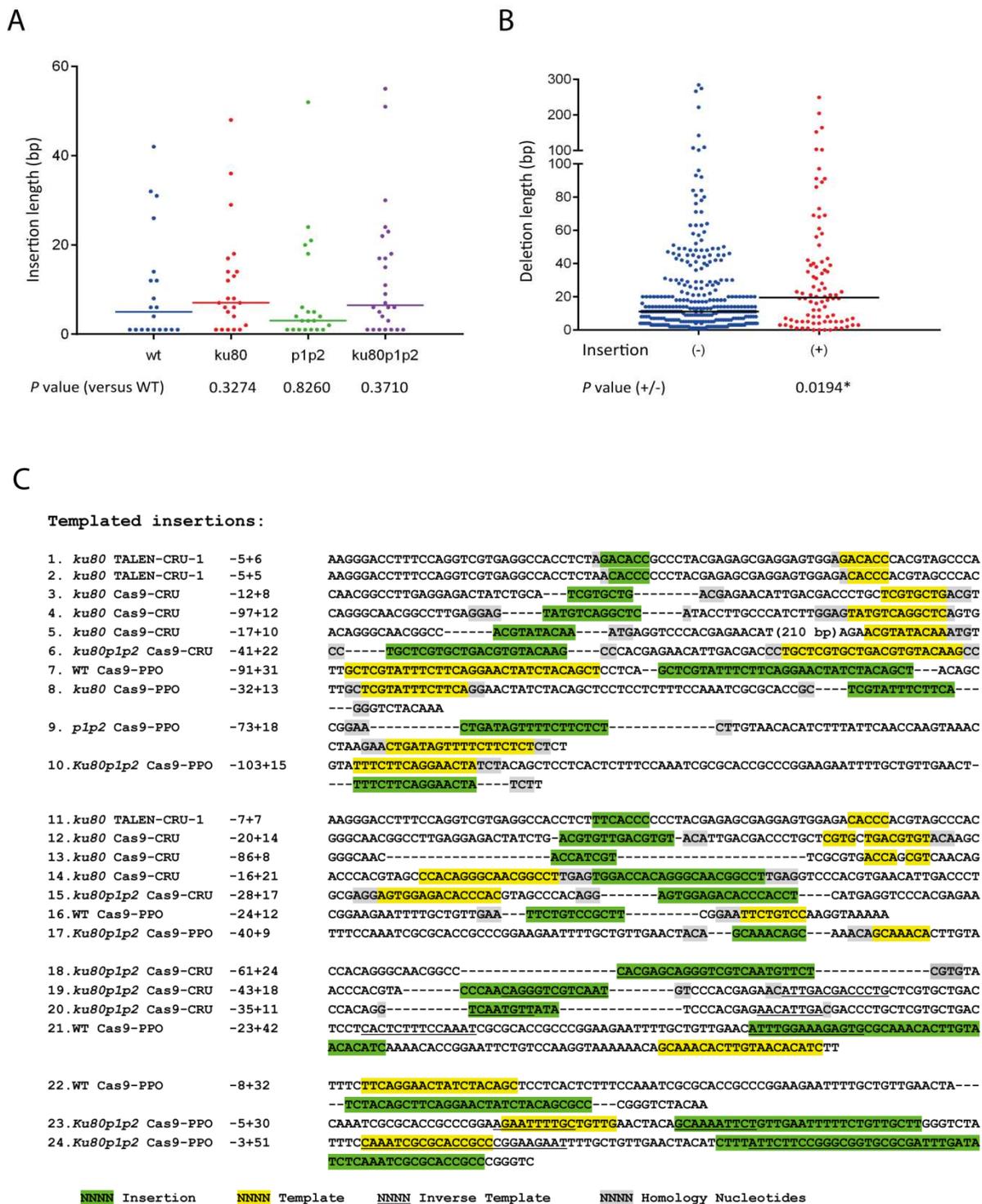


Figure 6. Analysis of insertions. (A) Scatter plot of insertion lengths for the indicated genotypes. Data are combined from all nucleases. (B) Scatter plot of deletion lengths with or without insertions. Data are combined from all genotypes and nucleases. Median insertion or deletion length is indicated by the bar in the graph. P-values are derived from a two-tailed Mann-Whitney test. Asterisk indicates the statistical significant difference ($P < 0.05$). (C) Footprints consisting of deletions (dashes) accompanied with insertions. Insertions are shown in green, template sources for the insertions are shown in yellow (direct strand) or underlined (reverse complement). Homologies between sequences flanking the template and the insertion and probably used as primer are shown in gray. Footprints from 1 to 10 are examples of perfectly matching the template, 11 to 17 are partially matching the template and 18 to 24 are reversely matching the template. Numbers are length of deletions (-) and insertions (+).

Sequence-specific nuclease-induced double strand break repair

1. TALEN-CRU			
WT TALEN-CRU-1			
plant line 3	-2	TCCAGGTCGTGAGGCCACCTCTA	-----CAGCCCTACGAGCGGAGTGG
	-11	TCCAGGTCGTGAGGCCACCTCTA	-----CGAGAGCGAGGAGTGG (2)
plant line 5	sub	TC-AGGTCGTGAGGCCACCTCTA	AGACAGCCCTACGAGCGGAGTGG
	sub	TCCAGGTCGTGAGGCCACCTCTA	AGACAGCCCTACGAGCGGAGTGG
	+1	TCCAGGTCGTGAGGCCACCTCTA	AGACAGCCCTACGAGCGGAGTGG
	-1	TCCAGGTCGTGAGGCCACCTCTA	-----ACAGCCCTACGAGCGGAGTGG
	-1	TCCAGGTCGTGAGGCCACCTCTA	-----GACAGCCCTACGAGCGGAGTGG
	+5+1	TCCAGGTCGTGAGGCCACCTCTA	-----ACAGCCCTACGAGCGGAGTGG (2)
	-2	TCCAGGTCGTGAGGCCACCTCTA	-----CAGCCCTACGAGCGGAGTGG
	-2	TCCAGGTCGTGAGGCCACCTCTA	-----CAGCCCTACGAGCGGAGTGG (6)
	-3	TCCAGGTCGTGAGGCCACCTCTA	-----CAGCCCTACGAGCGGAGTGG (3)
	-7	TCCAGGTCGTGAGGCCACCTCTA	-----AGCCCTACGAGCGGAGTGG
	-9	TCCAGGTCGTGAGGCCACCTCTA	-----CCTACGAGCGGAGTGG
	-11	TCCAGGTCGTGAGGCCACCTCTA	-----CCTACGAGCGGAGTGG (2)
plant line 7	-13	TCCAGGTCGTGAGGCCACCTCTA	-----ACGAGAGCGAGGAGTGG
	-2	TCCAGGTCGTGAGGCCACCTCTA	-----CAGCCCTACGAGCGGAGTGG (4)
	-3	TCCAGGTCGTGAGGCCACCTCTA	-----AGACAGCCCTACGAGCGGAGTGG
	-4	TCCAGGTCGTGAGGCCACCTCTA	-----ACAGCCCTACGAGCGGAGTGG (2)
	-11	TCCAGGTCGTGAGGCCACCTCTA	-----CGAGAGCGGAGTGG
plant line 9	-2	TCCAGGTCGTGAGGCCACCTCTA	-----CAGCCCTACGAGCGGAGTGG (3)
plant line 12	-13	TCCAGGTCGTGAGGCCACCTCTA	-----ACGAGAGCGGAGTGG (2)
	-2	TCCAGGTCGTGAGGCCACCTCTA	-----CAGCCCTACGAGCGGAGTGG
plant line 13	-13	TCCAGGTCGTGAGGCCACCTCTA	-----ACGAGAGCGGAGTGG
plant line 14	-3	TCCAGGTCGTGAGGCCACCTCTA	-----AGCCCTACGAGCGGAGTGG
	-2	TCCAGGTCGTGAGGCCACCTCTA	-----CAGCCCTACGAGCGGAGTGG (5)
	-14	TCCAGGTCGTGAGGCCACCTCTA	-----CAGCCCTACGAGCGGAGTGG (2)
plant line 17	-9	TCCAGGTCGTGAGGCCACCTCTA	-----AGCCCTACGAGCGGAGTGG (2)
	-5	TCCAGGTCGTGAGGCCACCTCTA	-----GCCCTACGAGCGGAGTGG
	-3	TCCAGGTCGTGAGGCCACCTCTA	-----CAGCCCTACGAGCGGAGTGG
	-2	TCCAGGTCGTGAGGCCACCTCTA	-----GACAGCCCTACGAGCGGAGTGG
plant line 19	-1	TCCAGGTCGTGAGGCCACCTCTA	-----ACAGCCCTACGAGCGGAGTGG
	-6	TCCAGGTCGTGAGGCCACCTCTA	-----ACAGCCCTACGAGCGGAGTGG
ku80 TALEN-CRU-1			
plant line 1	+1	TCCAGGTCGTGAGGCCACCTCTA	AGACAGCCCTACGAGCGGAGTGG
	-2	TCCAGGTCGTGAGGCCACCTCTA	-----CAGCCCTACGAGCGGAGTGG
	-3	TCCAGGTCGTGAGGCCACCTCTA	-----CAGCCCTACGAGCGGAGTGG
	-4	TCCAGGTCGTGAGGCCACCTCTA	-----GACAGCCCTACGAGCGGAGTGG
	-7	TCCAGGTCGTGAGGCCACCTCTA	-----CCTACGAGCGGAGTGG
	-11	TCCAGGTCGTGAGGCCACCTCTA	-----CGAGAGCGGAGTGG (5)
plant line 2	-2	TCCAGGTCGTGAGGCCACCTCTA	-----CAGCCCTACGAGCGGAGTGG (16)
	-4	TCCAGGTCGTGAGGCCACCTCTA	-----ACAGCCCTACGAGCGGAGTGG
	-11	TCCAGGTCGTGAGGCCACCTCTA	-----CGAGAGCGGAGTGG (2)
plant line 3	-2	TCCAGGTCGTGAGGCCACCTCTA	-----CAGCCCTACGAGCGGAGTGG (5)
	-3+1	TCCAGGTCGTGAGGCCACCTCTA	-----CAGCCCTACGAGCGGAGTGG
	-9	TCCAGGTCGTGAGGCCACCTCTA	-----AGCCCTACGAGCGGAGTGG
	-9+1	TCCAGGTCGTGAGGCCACCTCTA	-----CCCTACGAGCGGAGTGG
	-11	TCCAGGTCGTGAGGCCACCTCTA	-----CGAGAGCGGAGTGG (2)
	-13	TCCAGGTCGTGAGGCCACCTCTA	-----ACGAGAGCGGAGTGG
	-24	TCCAGGTCGTGAGGCCACCTCTA	-----GCAGGAGTGG
	-25	TCCAGGTCGTGAGGCCACCTCTA	-----GAGGAGTGG
plant line 5	-1	TCCAGGTCGTGAGGCCACCTCTA	-----GACAGCCCTACGAGCGGAGTGG
	-2	TCCAGGTCGTGAGGCCACCTCTA	-----CAGCCCTACGAGCGGAGTGG (2)
	-5	TCCAGGTCGTGAGGCCACCTCTA	-----GCCCTACGAGCGGAGTGG (2)
	-5	TCCAGGTCGTGAGGCCACCTCTA	-----ACAGCCCTACGAGCGGAGTGG
	-5+1	TCCAGGTCGTGAGGCCACCTCTA	-----ACAGCCCTACGAGCGGAGTGG (2)
	-5+6	TCCAGGTCGTGAGGCCACCTCTA	AGACAGCCCTACGAGCGGAGTGG
	-7	TCCAGGTCGTGAGGCCACCTCTA	-----AGCCCTACGAGCGGAGTGG
	-7+1	TCCAGGTCGTGAGGCCACCTCTA	-----CCCTACGAGCGGAGTGG
	-7+7	TCCAGGTCGTGAGGCCACCTCTA	AGACAGCCCTACGAGCGGAGTGG
	-7+7	TCCAGGTCGTGAGGCCACCTCTA	AGACAGCCCTACGAGCGGAGTGG
	-8	TCCAGGTCGTGAGGCCACCTCTA	-----CCCTACGAGCGGAGTGG
	-13	TCCAGGTCGTGAGGCCACCTCTA	-----GACAGCCCTACGAGCGGAGTGG
plant line 9	-2	TCCAGGTCGTGAGGCCACCTCTA	-----CAGCCCTACGAGCGGAGTGG (2)
	-9	TCCAGGTCGTGAGGCCACCTCTA	-----AGCCCTACGAGCGGAGTGG
	-11	TCCAGGTCGTGAGGCCACCTCTA	-----CGAGAGCGGAGTGG (8)
plant line 11	-2	TCCAGGTCGTGAGGCCACCTCTA	-----CAGCCCTACGAGCGGAGTGG (8)
	-5+5	TCCAGGTCGTGAGGCCACCTCTA	AGACAGCCCTACGAGCGGAGTGG
plant line 12	-6	TCCAGGTCGTGAGGCCACCTCTA	-----ACAGCCCTACGAGCGGAGTGG (2)
	-11	TCCAGGTCGTGAGGCCACCTCTA	-----CGAGAGCGGAGTGG (2)
	-18	TCCAGGTCGTGAGGCCACCTCTA	-----GAGGAGTGG
plant line 14	-2	TCCAGGTCGTGAGGCCACCTCTA	-----CAGCCCTACGAGCGGAGTGG
	-11	TCCAGGTCGTGAGGCCACCTCTA	-----CGAGAGCGGAGTGG (5)
plant line 15	-2	TCCAGGTCGTGAGGCCACCTCTA	-----CAGCCCTACGAGCGGAGTGG (2)
	-3	TCCAGGTCGTGAGGCCACCTCTA	-----CAGCCCTACGAGCGGAGTGG (2)
	-11	TCCAGGTCGTGAGGCCACCTCTA	-----CGAGAGCGGAGTGG (2)

Figure 7. Sequences of resistant target sites. DNA from several plant lines of wild type and NHEJ mutants with different nuclease constructs was predigested with the appropriate restriction enzyme (*DdeI* for TALEN-CRU-1, *PstI* for Cas9-CRU, and *FauI* for Cas9-PPO), used for PCR, digested with the same enzyme and resistant products were cloned and sequenced. Footprints included deletions (dashed lines), insertions (green) and substitutions (blue). Microhomologies used for repair are shown in purple. Numbers are length of deletions (-) and insertions (+).

p1p2 TALEN-CRU-1			
plant line 1	-5+1	TCCAGGTCGTGAGGCCACCT-- A ---ACAGCCCTACGAGAGCGAGGAGTGG	
	-11	TCCAGGTCGTGAGGCCACCT CTA -----CGAGAGCGAGGAGTGG	(12)
	-11	TCCAGGTCGTGAGGCCAC C -----CCTACGAGAGCGAGGAGTGG	(2)
plant line 2	-1+24	TCCAGGTCGTGAGGCCACCT CTAA ACAGCCCTACGAGAGCGAGGAGTGG	
		AGACACCCACGTAGCCACAGGGCAACGGCCTGAGGAGACTATCTGCA G	
		GAGAACATTGACGACCCCTGATC GAGGT	
	-2	TCCAGGTCGTGAGGCCACCT CTA ---ACAGCCCTACGAGAGCGAGGAGTGG	
	-4+3	TCCAGGTCGTGAGGCCACCT- TAC GACAGCCCTACGAGAGCGAGGAGTGG	
	-11	TCCAGGTCGTGAGGCCACCT CTA -----CGAGAGCGAGGAGTGG	(10)
plant line 3	-8	TCCAGGTCGTGAGGCCAC C -----AGCCCTACGAGAGCGAGGAGTGG	
	-11	TCCAGGTCGTGAGGCCACCT CTA -----CGAGAGCGAGGAGTGG	(8)
plant line 5	-2	TCCAGGTCGTGAGGCCACCT CTA ---ACAGCCCTACGAGAGCGAGGAGTGG	(3)
	-11	TCCAGGTCGTGAGGCCACCT CTA -----CGAGAGCGAGGAGTGG	
plant line 6	-3	TCCAGGTCGTGAGGCCACCT CTA ---ACAGCCCTACGAGAGCGAGGAGTGG	
	-11	TCCAGGTCGTGAGGCCAC C -----CCTACGAGAGCGAGGAGTGG	(3)
	-11	TCCAGGTCGTGAGGCCACCT CTA -----CGAGAGCGAGGAGTGG	(5)
plant line 7	-4	TCCAGGTCGTGAGGCCACCT---GACAGCCCTACGAGAGCGAGGAGTGG	
	-5	TCCAGGTCGTGAGGCCACCT---ACAGCCCTACGAGAGCGAGGAGTGG	(5)
	-5	TCCAGGTCGTGAGGCCACCT C ---ACAGCCCTACGAGAGCGAGGAGTGG	
	-6+1	TCCAGGTCGTGAGGCCACCT-- H ---ACAGCCCTACGAGAGCGAGGAGTGG	
	-9	TCCAGGTCGTGAGGCCACCT CT -----CTACGAGAGCGAGGAGTGG	
	-11	TCCAGGTCGTGAGGCCACCT CTA -----CGAGAGCGAGGAGTGG	(5)
	-11	TCCAGGTCGTGAGGCCAC C -----CCTACGAGAGCGAGGAGTGG	
plant line 8	-2	TCCAGGTCGTGAGGCCACCT CTA ---ACAGCCCTACGAGAGCGAGGAGTGG	(3)
	-4	TCCAGGTCGTGAGGCCACCT C ---ACAGCCCTACGAGAGCGAGGAGTGG	(3)
	-11	TCCAGGTCGTGAGGCCACCT CTA -----CGAGAGCGAGGAGTGG	(2)
plant line 9	-1	TCCAGGTCGTGAGGCCACCT CTA ---GACAGCCCTACGAGAGCGAGGAGTGG	
	+4	TCCAGGTCGTGAGGCCACCT CTAGAGA AAGACAGCCCTACGAGAGCGAGGA	(2)
	-3+1	TCCAGGTCGTGAGGCCACCT CTA - H ---ACAGCCCTACGAGAGCGAGGAGTGG	(3)
	-11	TCCAGGTCGTGAGGCCACCT CTA -----CGAGAGCGAGGAGTGG	(3)
plant line 10	-1	TCCAGGTCGTGAGGCCACCT CTA ---ACAGCCCTACGAGAGCGAGGAGTGG	
	-2	TCCAGGTCGTGAGGCCACCT CTA ---ACAGCCCTACGAGAGCGAGGAGTGG	
	-11	TCCAGGTCGTGAGGCCACCT CTA -----CGAGAGCGAGGAGTGG	(2)
plant line 11	-1+1	TCCAGGTCGTGAGGCCACCT TTAAG ACAGCCCTACGAGAGCGAGGAGTGG	
	-3	TCCAGGTCGTGAGGCCACCT CTA ---ACAGCCCTACGAGAGCGAGGAGTGG	
	-5	TCCAGGTCGTGAGGCCACCT---ACAGCCCTACGAGAGCGAGGAGTGG	
	-5	TCCAGGTCGTGAGGCCACCT CTA ---GCCCTACGAGAGCGAGGAGTGG	
	-11	TCCAGGTCGTGAGGCCACCT CTA -----CGAGAGCGAGGAGTGG	(5)
ku80p1p2 TALEN-CRU-1			
plant line 1	-2	TCCAGGTCGTGAGGCCACCT CTA ---ACAGCCCTACGAGAGCGAGGAGTGG	
	Subs3	TCCAGGTCGTGAGGCCACCT TTAAG ACAGCCCTACGAGAGCGAGGAGTGG	
	-5+1	TCCAGGTCGTGAGGCCACCT-- A ---ACAGCCCTACGAGAGCGAGGAGTGG	(4)
	-8	TCCAGGTCGTGAGGCCAC C -----AGCCCTACGAGAGCGAGGAGTGG	
	-11	TCCAGGTCGTGAGGCCACCT CTA -----CGAGAGCGAGGAGTGG	(8)
plant line 2	-2	TCCAGGTCGTGAGGCCACCT CTA ---ACAGCCCTACGAGAGCGAGGAGTGG	(2)
	-11	TCCAGGTCGTGAGGCCACCT CTA -----CGAGAGCGAGGAGTGG	(6)
plant line 3	+1	TCCAGGTCGTGAGGCCACCT CTAA GACAGCCCTACGAGAGCGAGGAGTGG	
	-3	TCCAGGTCGTGAGGCCACCT CTA ---ACAGCCCTACGAGAGCGAGGAGTGG	
	-8	TCCAGGTCGTGAGGCCAC C -----AGCCCTACGAGAGCGAGGAGTGG	
	-11	TCCAGGTCGTGAGGCCACCT CTA -----CGAGAGCGAGGAGTGG	(4)
	-11	T@CA GTCGTGAGGCCACCT CTA -----CGAGAGCGAGGAGTGG	
plant line 4	-11	TCCAGGTCGTGAGGCCAC C -----CCTACGAGAGCGAGGAGTGG	
	-11	TCCAGGTCGTGAGGCCACCT CTA -----CGAGAGCGAGGAGTGG	(5)
	-34+23	ACCT CTA -- GACAGCGAGGAGTGAACACCA -----CACGTAGCC	
plant line 5	-1+1	TCCAGGTCGTGAGGCCACCT TTAAG ACAGCCCTACGAGAGCGAGGAGTGG	
	-11	TCCAGGTCGTGAGGCCAC C -----CCTACGAGAGCGAGGAGTGG	(2)
	-11	TCCAGGTCGTGAGGCCACCT CTA -----CGAGAGCGAGGAGTGG	(5)
	-11	T@CA GTCGTGAGGCCACCT CTA -----CGAGAGCGAGGAGTGG	
plant line 7	-11	TCCAGGTCGTGAGGCCAC C -----CCTACGAGAGCGAGGAGTGG	(4)
plant line 8	-11	TCCAGGTCGTGAGGCCACCT CTA -----CGAGAGCGAGGAGTGG	(3)
	-11	TCCAGGTCGTGAGGCCAC C -----CCTACGAGAGCGAGGAGTGG	
plant line 9	-5	TCCAGGTCGTGAGGCCACCT---ACAGCCCTACGAGAGCGAGGAGTGG	(4)
	-20+4	TCCAGGTCGTGAGGCC----- FACT -----AGAGCGAGGAGTGG	
plant line 10	-5	TCCAGGTCGTGAGGCCACCT---ACAGCCCTACGAGAGCGAGGAGTGG	(2)
	-11	TCCAGGTCGTGAGGCCAC C -----CCTACGAGAGCGAGGAGTGG	(3)

Figure 7. (continued 1).

2. CRISPR/Cas9-CRU

WT Cas9-CRU

plant line 1
 -3 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCA---TGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC (2)
 -3 CCACAGGGCAACGGCCTTGAGGAGACTATC---AGCATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -6 CCACAGGGCAACGGCCTTGAGGAGACTATCT---TGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -7 CCACAGGGCAACGGCCTTGAGGAGACTATCTG---AGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -9 CCACAGGGCAACGGCCTTGAGGAGACTAT---GAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -13 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCA---CGAGAACATTGACGACCCCTGCTCGTGTGAC (2)
 -14 CCACAGGGCAACGGCCTTGAGGAGACTATC---CCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -20 CCACAGGGCAACGGCCTTGAGG---TCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC (2)
 -30 CCACAGGGCAAC---TCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -50 CCACAGGG---ACGACCCCTGCTCGTGTGAC (2)
 plant line 2
 -3 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCA---TGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC (3)
 -4 CCACAGGGCAACGGCCTTGAGGAGACTATCT---CATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -4+1 CCACAGGGCAACGGCCTTGAGGAGACTATC---GCATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -6 CCACAGGGCAACGGCCTTGAGGAGACTATC---ATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -7 CCACAGGGCAACGGCCTTGAGGAGACTAT---ATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -10 CCACAGGGCAACGGCCTTGAGGAGA---ATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -14 CCACAGGGCAACGGCCTTGAGGAGAC---GTTCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -14+6 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCA---TTTAT---GAGAACATTGACGACCCCTGCTCGTGTGAC (2)
 -18+6 CCACAGGGCAACGGCCTTGAGGAGACTATCTG---ACCTGT---GAACATTGACGACCCCTGCTCGTGTGAC
 -22+1 CCACAGGGCAACGG---G---ATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -41 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCA---GAC
 plant line 6
 -1 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCA---CATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC (2)
 -3 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCA---TGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC (2)
 -5 CCACAGGGCAACGGCCTTGAGGAGACTATCTGC---GAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -6 CCACAGGGCAACGGCCTTGAGGAGACTATC---ATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -9 CCACAGGGCAACGGCCTTGAGGAGACTATCTG---GTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC

ku80 Cas9-CRU

plant line 1
 -20 CCACAGGGCAACGGC---ATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC (2)
 -29 CCACAGGGCA---GGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -20+14 CCACAGGGCAACGGCCTTGAGGAGACTATCTG---ACGTGTGACCTGT---ACATTGACGACCCCTGCTCGTGTGAC
 -43 GCCAC---GAGAACATTGACGACCCCTGCTCGTGTGAC (2)
 -92 CCACAG---92---TCGCGTGACCAAGCTCAACAGTATACCT
 -86+8 ACAGGGCAAC---ACCATCGT---TCGCGTGACCAAGCTCAACAG
 plant line 2
 -2 CCACAGGGCAACGGCCTTGAGGAGACTATCTGC---CATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -3 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCA---TGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -4 CCACAGGGCAACGGCCTTGAGGAGACTATCT---CATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -8 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCA---TCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -10 CCACAGGGCAACGGCCTTGAGGAGA---ATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -12+8 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCA---TCGTGCTG---ACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -20 CCACAGGGCAACGGCCTTGAGG---TCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -56+36 CCTTGAGGAGACTATCT---AACCTTGACGACCCCTGCACAAGCTTGCCCTACACTGC---CCGAG
 -9 CCACAGGGCAACGGCCTTGAGGAGACTAT---GAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -20 CCACAGGGCAACGGC---ATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -20 CCACAGGGCAACGGCCTTGAGG---TCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -30 CCACAGGGCAACGGC---ACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -81 CCACAGGGCAACGGCCTTGAGGAGACT---81---GCGTCAACAGCTAT
 plant line 4
 -1 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCA---CATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -3 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCA---TGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -7 CCACAGGGCAACGGCCTTGAGGAGACTAT---ATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC (3)
 -20 CCACAGGGCAACGGC---ATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -31+6 CCAC---ETAGCG---CATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 plant line 6
 -26 CCACAGGGCAACGGCCTTGAGGAGACTATCTG---ACGACCCCTGCTCGTGTGAC (6)
 plant line 9
 -20 CCACAGGGCAACGGCCTTGAGG---TCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC (2)
 -61+48 GACTATCT---ACAGTCCATGAAGGTTCAATAAGGCTCTGATAGTTGGCCTTCATGAGCTA---GCGTGT (4)
 plant line 11
 -4 CCACAGGGCAACGGCCTTGAGGAGACTATCTGC---TGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -80 CGTGAGGCCACTC---80---ATGAGGTCCC
 -97+12 CCACAGGGCAACGGCCTTGAGGAG---FATGTCAGGCTC---ATACCTTG (3)
 plant line 12
 -9 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCA---CCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -13 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCA---CGAGAACATTGACGACCCCTGCTCGTGTGAC
 -20 CCACAGGGCAACGGC---ATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -44 CCACAGGGCAACGGCCTTGAGGAGACTATCTG---AC
 -107 TAAGGACCTTCC---107---CACGAGAACAT

Figure 7. (continued 2).

p1p2 Cas9-CRU

plant line 1
 -3 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCA---TGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -6 CCACAGGGCAACGGCCTTGAGGAGACTATG-----ATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -8 CCACAGGGCAACGGCCTTGAGGAGACTAT-----TGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -17 CCACAGGGCAACGGCCTTGAGGAGACTATCTG-----AGAACATTGACGACCCCTGCTCGTGTGAC
 -20 CCACAGGGCAACGGCCTTGAGG-----TCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC (3)
 -16+21 CCACAGGGCAACGGCCTTGAGTGGACCACAGGGCAACGGCCTTGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTG
 -69+52 TATCT-----ACAGTCCATGAAGGTTCAAAAAGGCTGTAGATAGTTGGCCTTCATGGACTA-----GCGTGA (3)

plant line 2
 -1 CCACAGGGCAACGGCCTTGAGGAGACTATCTG-AGCATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC (2)
 -2 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCA--ATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC (2)
 -20 CCACAGGGCAACGGCCTTGAGG-----TCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC (2)
 -17+10 CCACAGGGCAACGGC-----ACGTATACAA--ATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC (2)
 -78 ACCTCTAAGACAGCC-----78-----CACGAGAACATTGACGACCCCTGCTCGTGTGAC (2)

plant line 4
 -3 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCA--TGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC (2)
 -6 CCACAGGGCAACGGCCTTGAGGAGACTATG-----ATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC (3)
 -44 CCACAGGGCAACGGCCTTGAGGAGACTATCTG-----44-----ACGTGTAC (4)

plant line 8
 -3 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCA--TGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -6 CCACAGGGCAACGGCCTTGAGGAGACTATG-----ATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -20 CCACAGGGCAACGGCCTTGAGG-----TCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -28 CCACAGGGCAACGG-----TCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC (2)

plant line 11
 -1 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCA-CATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC (2)
 -3 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCA--TGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -5+2 CCACAGGGCAACGGCCTTGAGGAGACTAT-PT--GCATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -19+5 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCA-----TTTAT-----CATTGACGACCCCTGCTCGTGTGAC

plant line 15
 -2 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCA--CATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -14 CCACAGGGCAACGGCCTTGAG-----CATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -14 CCACAGGGCAACGGCCTTGAGGAGACTATCT-----CCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -20 CCACAGGGCAACGGCCTTGAGG-----TCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -29 CCACAGGGCAACGGCC-----CACGAGAACATTGACGACCCCTGCTCGTGTGAC

ku80p1p2 Cas9-CRU

plant line 1
 -3 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCA--TGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC (2)
 -4 CCACAGGGCAACGGCCTTGAGGAGACTATCT-----CATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC (3)
 -8 CCACAGGGCAACGGCCTTGAGGAGACTATCT-----AGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -22 CCACAGGGCAACGGCCTTGAGGAGACTAT-----GAACATTGACGACCCCTGCTCGTGTGAC (4)
 -29 CCACAGGGCAACGGCC-----CACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -35+11 CCACAGG-----TCAATGTTATA-----TCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC (2)
 -84 TAAGACAGCCCTACGAGAGCGA-----84-----CCCTGCTCGTGTGAC
 -93 TTTCAGGT-----93-----ATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC

plant line 2
 -1 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCA-CATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -2+1 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCA-ATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC (2)
 -5 CCACAGGGCAACGGCCTTGAGGAGACTATCT-----ATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC (2)
 -16 CCACAGGGCAACGGCCTTGAGGAGACTATG-----ACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -20 CCACAGGGCAACGGCC-----ATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -35+3 CCACAGGGCAACGGCCTTGAGGAGAC-----GAC-----ACCCTGCTCGTGTGAC
 -47 GGAGACACCCACG-----47-----CCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -49 CCACAGGGCAACGGCCTTGAGGAGACTATCTG-----49-----TACAAGCCACGCTAGGTCGCGTGA

plant line 3
 -4 CCACAGGGCAACGGCCTTGAGGAGACTATCTG-----ATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -6 CCACAGGGCAACGGCCTTGAGGAGACTATG-----ATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -27 CCACAGGGCAACGGCCTTGAGGAGACTAT-----TGACGACCCCTGCTCGTGTGAC
 -36 CCACAGG-----TCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC

plant line 6
 -5 CCACAGGGCAACGGCCTTGAGGAGACTATC-----CATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -14 CCACAGGGCAACGGCCTTGAGGAGACTATC-----CCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -20 CCACAGGGCAACGGCCTTGAGG-----TCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -41+22 CC-----TGCTCGTGTGACGTGTACAAG-----CCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -61+24 CCACAGGGCAACGGCC-----CACGAGCAGGGTCGTCAATGTTGT-----CGTGTA
 -1+5 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCAACGGCCATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTG
 -14 CCACAGGGCAACGGCCTTGAGGAGACTATC-----CCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -38+6 CCACAGGGCAACGGCCTTGAGGAGA-----TCTTGA-----CGTGTGAC (2)
 -48 CCACAGGGCAACG-----ACCTGCTCGTGTGAC

plant line 13
 -37 CCAC-----GTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC (5)

plant line 15
 -27 CCACAGGGCAACGGCCTTGAGGAGA-----ACATTGACGACCCCTGCTCGTGTGAC (2)
 -28+17 CCACAGG-----AGTGGAGACACCCACCT-----CATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -51+6 CCACAGGGCAACGGCCTTGAGGAGACTATCTG-----ACGTGT-----CAAGCCACGCC
 -64 GCGAGGAGTGSAG-----64-----ACATTGACGACCCCTGCTCGTGTGAC

plant line 16
 -2+1 CCACAGGGCAACGGCCTTGAGGAGACTATCTGCA-ATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -7 CCACAGGGCAACGGCCTTGAGGAGACTA-----CATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -20 CCACAGGGCAACGGCC-----ATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -29 CCACAGG-----ATGAGGTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC
 -43+18 ACCCACGTA-----CCCAACAGGGTCGTCAAT-----GTCCCACGAGAACATTGACGACCCCTGCTCGTGTGAC

Figure 7. (continued 3).

Sequence-specific nuclease-induced double strand break repair

3. CRISPR/Cas9-PPO

WT Cas9-PPO

plant line 1	-5	CGGAAGAATTTGCTGTTGAACACTACATG----	TCTACAACACC	CGGAATTCGTCCAAGTAAAAA
	-3	CGGAAGAATTTGCTGTTGAACACTACAT--	CGGGTCTACAACACC	CGGAATTCGTCCAAGTAAAAA (2)
	-7	CGGAAGAATTTGCTGTTGAACACT-----	CGGGTCTACAACACC	CGGAATTCGTCCAAGTAAAAA
	-10	CGGAAGAATTTGCTGTTGAACACT-----	GCTACAACACC	CGGAATTCGTCCAAGTAAAAA
	-14	CGGAAGAATTTGCTGTTGAACACTA-----	AACACCGGAATTCGTCCAAGTAAAAA (5)	
	-18	CGGAAGAATTTGCTGTTGAACACTA-----	CGGAATTCGTCCAAGTAAAAA	
	-19+4	CGGAAGAATTTGCTGTTGAACACTA-----	CACG	CGGAATTCGTCCAAGTAAAAA
	-19+1	CGGAAGAATTTGCTGTTGAACACTA-----	A	CGGAATTCGTCCAAGTAAAAA
	-24+12	CGGAAGAATTTGCTGTTGAA-----	TTCTGTCGGCT	CGGAATTCGTCCAAGTAAAAA
	-27	CGGAAGAATTTGCTGTTGAACTACAT-----	ACAACACC	CGGAATTCGTCCAAGTAAAAA
	-63	CGGAAGAATTTGCTGTTGAACTACAT-----	-----	CTTTA
	-84	GCTCGTATTTCTTCA-----	AACACCGGAATTCGTCCAAGTAAAAA	
plant line 2	-3	CGGAAGAATTTGCTGTTGAACACTACAT--	CGGGTCTACAACACC	CGGAATTCGTCCAAGTAAAAA
	-4	CGGAAGAATTTGCTGTTGAACACTA-----	CGGGTCTACAACACC	CGGAATTCGTCCAAGTAAAAA
	-6	CGGAAGAATTTGCTGTTGAACACTAC-----	GGGTCTACAACACC	CGGAATTCGTCCAAGTAAAAA (6)
	-9	CGGAAGAATTTGCTGTTGAACACTAC-----	TCTACAACACC	CGGAATTCGTCCAAGTAAAAA
	-14	CGGAAGAATTTGCTGTTGAACACTA-----	AACACCGGAATTCGTCCAAGTAAAAA (2)	
	-28	CGGAAGAATTTGCTGTTGAA-----	-----	ATTCTGTCCAAGTAAAAA
	-30	CGG-----	GCTACAACACC	CGGAATTCGTCCAAGTAAAAA (2)
	-40	CGGAAGAATTTGCTGTTGAACACTA-----	-----	AACAGCAACACTGTGTA
	-48	CGGAAGAATTTGCTGTTGAACACTA-----	-----	AACACTGTAAACACTCTTTA
	-71	CTCCTATTTCTTCAGGACT-----	71	GCTACAACACC
	-8+32	TTTCTCAGGAACCTACTACAGCTCCTCACTCTTCCAAATCGGCACCGCCGGGAAGAAATTTGCTGTTGAACTA-----	CTACAGCTTCAGGAACCTACTACAGCGCC	CGGGTCTACA (2)
	-23+42	TCCTCACTCTTCCAAATCGGCACCGCCGGGAAGAAATTTGCTGTTGAACTTTGGAAGAGTGCCG	AAACACTTGTAAACACATCAAAACACC	CGGAATTCGTCCAAGTAAAAA (7)
plant line 3	-3+1	CGGAAGAATTTGCTGTTGAACACTACATG--	GGTCTACAACACC	CGGAATTCGTCCAAGTAAAAA
	-4	CGGAAGAATTTGCTGTTGAACACTACAT--	GGGTCTACAACACC	CGGAATTCGTCCAAGTAAAAA
	-6	CGGAAGAATTTGCTGTTGAACACTACAT--	GCTACAACACC	CGGAATTCGTCCAAGTAAAAA (3)
	-9	CGGAAGAATTTGCTGTTGAA-----	GGGTCTACAACACC	CGGAATTCGTCCAAGTAAAAA
	-11	CGGAAGAATTTGCTGTTGAA-----	GGGTCTACAACACC	CGGAATTCGTCCAAGTAAAAA
	-11	CGGAAGAATTTGCTGTTGAA-----	GCTACAACACC	CGGAATTCGTCCAAGTAAAAA
	-14	CGGAAGAATTTGCTGTTGAACACTA-----	AACACCGGAATTCGTCCAAGTAAAAA (6)	
	-14	CGGAAGAATTTGCTGTTGAACACTACAT-----	ACACCGGAATTCGTCCAAGTAAAAA	
	-17	CGGAAGAATTTGCTGTT-----	CTACAACACC	CGGAATTCGTCCAAGTAAAAA
	-21	CGGAAGAATTTGCTGTTGAA-----	ACCGGAATTCGTCCAAGTAAAAA	
	-22	CGGAAGAATTTGCT-----	ACAACACC	CGGAATTCGTCCAAGTAAAAA
	-23	CGGAAGAATTTGCTGTTGAA-----	CGGAATTCGTCCAAGTAAAAA	
	-42	CGGAAGAATTTG-----	-----	TCCAAGTAAAAA
	-42+8	CGGAAGAATTTGCTGTTGAACACTAC-----	CGGAATTC	AACAGCAACACTTTGTA
	-91+31	CAGGAACCTACTACAGCTCCTCA-----	SCTCGTATTCTTCAGGAACCTACTACAGCT	ACAGCAA
	-204+14	CGGAAGAATTTGCTGTTGAACACTA-----	CGGAATTCGTCCAAGTAAAAA	GATCCACTT
	-274	TTTCCAAATC-----	AAGCATTCTCAGTTCTAGTTGGTCA	-----
plant line 7	-2	CGGAAGAATTTGCTGTTGAACACTACAT--	CGGGTCTACAACACC	CGGAATTCGTCCAAGTAAAAA
	-4	CGGAAGAATTTGCTGTTGAACACTACAT--	GGGTCTACAACACC	CGGAATTCGTCCAAGTAAAAA
	-5	CGGAAGAATTTGCTGTTGAACACTACAT--	GGTCTACAACACC	CGGAATTCGTCCAAGTAAAAA
	-6	CGGAAGAATTTGCTGTTGAACACTA-----	GGGTCTACAACACC	CGGAATTCGTCCAAGTAAAAA
	-7	CGGAAGAATTTGCTGTTGAACACTAC-----	GGTCTACAACACC	CGGAATTCGTCCAAGTAAAAA
	-11	CGGAAGAATTTGCTGTTGAA-----	GGTCTACAACACC	CGGAATTCGTCCAAGTAAAAA
	-14	CGGAAGAATTTGCTGTTGAACACTACAT-----	ACACCGGAATTCGTCCAAGTAAAAA	
	-14	CGGAAGAATTTGCTGTTGAACACTA-----	AACACCGGAATTCGTCCAAGTAAAAA (2)	
	-17	CGGAAGAATTTGCTGTTGAACTA-----	ACACCGGAATTCGTCCAAGTAAAAA	
	-17	CGGAAGAATTTGCTGTTGAA-----	AAACACC	CGGAATTCGTCCAAGTAAAAA (3)
	-18	CGGAAGAATTTGCTGTTGAACACTACAT-----	CGGAATTCGTCCAAGTAAAAA	
	-30	CGG-----	GCTACAACACC	CGGAATTCGTCCAAGTAAAAA
	-89+26	CGGAAGAATTTGCTGTTGAACACTA-----	SCTTTATTCACACC	CGGAATTCGTCCAAGTAAAAA
plant line 10	-2	CGGAAGAATTTGCTGTTGAACACTACAT--	CGGGTCTACAACACC	CGGAATTCGTCCAAGTAAAAA
	-4	CGGAAGAATTTGCTGTTGAACACTACAT--	GGGTCTACAACACC	CGGAATTCGTCCAAGTAAAAA
	-7	CGGAAGAATTTGCTGTTGAACACTACAT--	CTACAACACC	CGGAATTCGTCCAAGTAAAAA
	-9	CGGAAGAATTTGCTGTTGAACT-----	GGTCTACAACACC	CGGAATTCGTCCAAGTAAAAA
	-14	CGGAAGAATTTGCTGTTGAACACTA-----	AACACCGGAATTCGTCCAAGTAAAAA (4)	
	-19+1	CGGACGAATTT-----	CGGGTCTACAACACC	CGGAATTCGTCCAAGTAAAAA
	-39	CGGAAGAATTTGCTGTTGAA-----	-----	GGTAAAAA
	-91+12	CGTATTTCTTCAGGAAC-----	91	SAGACCGTGAG
				CTGTCCAAGTAAAAA
Ku80 Cas9-PPO				
plant line 1	-20	CGGAAGAATTTGCTGTTGAACACTAC-----	CGGAATTCGTCCAAGTAAAAA	
	-32+13	CTACAGCTCCTCCTTTCCAAATCGGCACCGC-----	TCGTATTCTTCA	GGGTCTACAAA
	-42	CGGAAGAATTT-----	-----	CTGTCCAAGTAAAAA
	-45	CGCCCGGAA-----	-----	TTCTGTCCAAGTAAAAA (6)
plant line 2	-14	CGGAAGAATTTGCTGTTGAACTA-----	AACACCGGAATTCGTCCAAGTAAAAA (6)	
	-20	CGGAAGAATTTGCTGTTGAACACTAC-----	CGGAATTCGTCCAAGTAAAAA	
	-26+2	CGGAAGAATTTT-----	TC	CAAACACC
	-45	CGCCCGGAA-----	-----	TTCTGTCCAAGTAAAAA
plant line 3	-8	CGGAAGAATTTGCTGTTGAACACTACAT-----	CTACAACACC	CGGAATTCGTCCAAGTAAAAA (2)
	-14	CGGAAGAATTTGCTGTTGAACACTACAT-----	ACACCGGAATTCGTCCAAGTAAAAA	
	-14	CGGAAGAATTTGCTGTTGAACACTA-----	AACACCGGAATTCGTCCAAGTAAAAA (2)	
	-18	CGGAAGAATTTGCTGTTGAACACTACAT-----	CGGAATTCGTCCAAGTAAAAA	
	-18	CGGAAGAATTTGCTGTT-----	CTACAACACC	CGGAATTCGTCCAAGTAAAAA
	-30	CGCCCGG-----	GCTACAACACC	CGGAATTCGTCCAAGTAAAAA (2)
	-41	CGGAAGAATTTGCTGTTGAACACTA-----	42	AACAGCAACCA
	-59	CGGAAGAATTTGCT-----	59	AACACTTGAAC (2)
	-109	TTTCTCAGGAACCTACTACA-----	-----	CTGTAAACACTCTTTA

Figure 7. (continued 4).

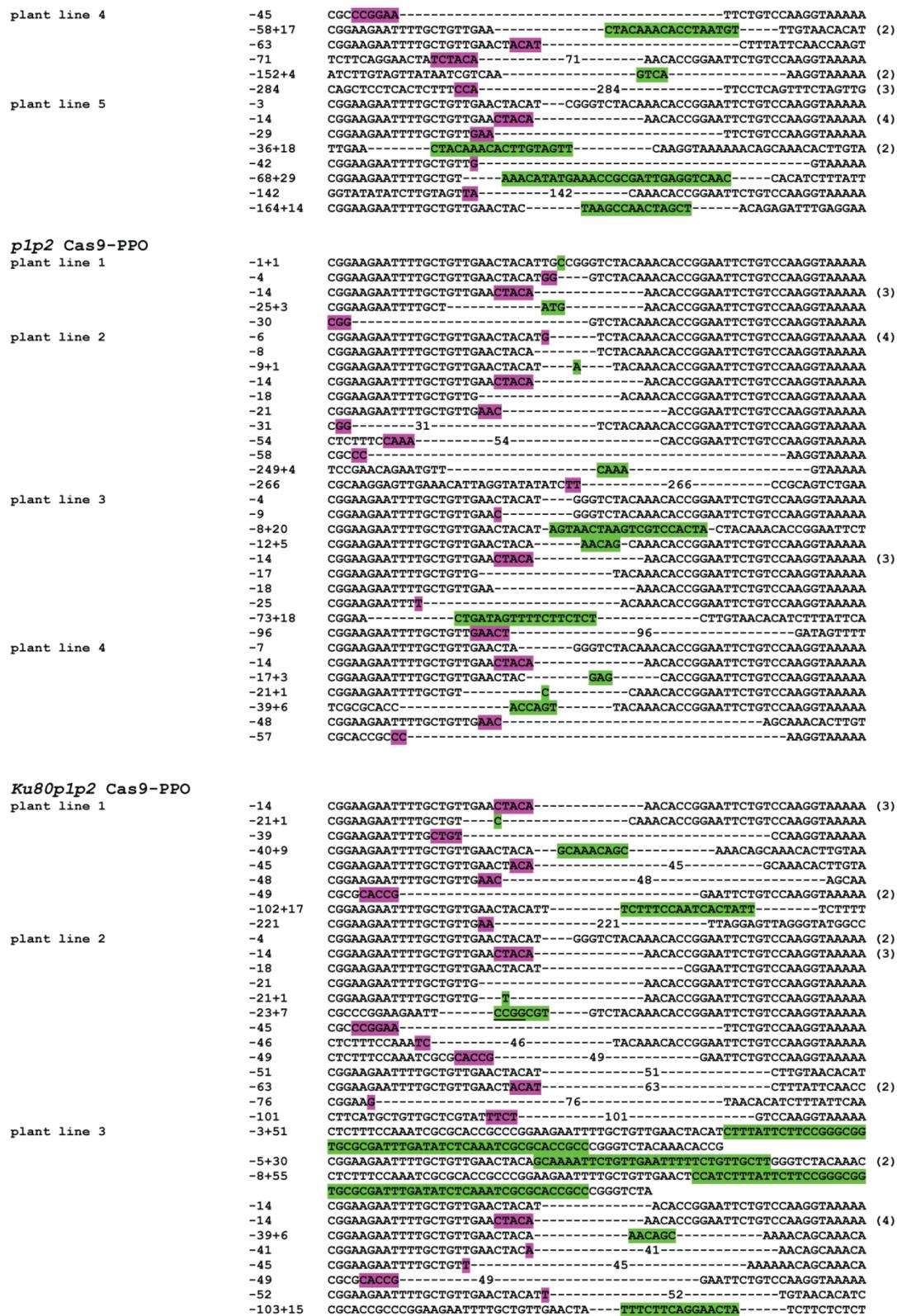


Figure 7. (continued 5).

Discussion

In this study, we demonstrated that TALEN- and CRISPR/Cas9-induced DSBs can be successfully repaired in *Arabidopsis*, even after loss of key components of the NHEJ repair pathway. In general plants expressing these nucleases look healthy and develop normally (unless an essential gene was targeted). As precise repair restores the break site leading to an intact substrate for the nuclease, a cycle of DSB induction and repair continues until mutations in the target sequence prevent the action of the nuclease. We showed that especially the Cas9-PPO nuclease was a very efficient tool for targeted mutagenesis, with close to 100% mutation frequency in one line. Since the *PPO* gene is an essential gene, this resulted in stunted growth of the seedlings (**Figure 1**). This phenotype was not observed in earlier targeted mutagenesis experiments of *PPO* using ZFNs (de Pater *et al.* 2013), indicating a higher activity of CRISPR/Cas9 on the *PPO* gene compared to the ZFNs. We noticed that the sgRNA recognized a sequence in the *PPO* gene with GG just 5' of the PAM sequence, which was recently shown to promote higher rates of mutagenesis (Farboud and Meyer 2015).

We analysed TALEN-induced and Cas9-induced mutations. Both nucleases are able to induce mutagenesis at the target sites and a variety of footprints (deletions and insertions) were found. However, much longer deletions were found in Cas9-CRU lines than in TALEN-CRU-1 lines with the same genetic background. The main reason for this observation may be the different lengths of amplified target sequence. In TALEN-CRU-1 lines a region of only 273 bp was amplified, due to the presence of additional *DdeI* sites just outside of this region. In Cas9-CRU lines a region of 970 bp long was amplified and analysed for footprints. This means that larger deletions could be missed more easily in TALEN-CRU-1 lines than in Cas9-CRU lines. Besides, different types of break ends may activate different end-resection mechanisms, which can affect end-joining outcomes. TALENs induce DSBs with 5' overhangs which are suitable for ligation, whereas Cas9 induces mostly blunt ends or incompatible ends (Jinek *et al.* 2012). Additionally, CRISPR-Cas9 might be more efficient than TALENs in plants.

As a result of imperfect end joining, various mutations in the target sites were found in each line. In the analysis of DSB repair outcome in NHEJ mutants, we observed a statistically significant increase in the median deletion length at the repair junction in the *ku80* and *ku80parp1parp2* mutants compared to wild type, but not in the *parp1parp2* mutant. This suggests that b-NHEJ is a more error-prone DSBs repair pathway than c-NHEJ. KU is known to competitively bind to DSB ends and protects break ends from end processing (Downs and Jackson 2004; Fell and Schild-Poulter 2014). Thus, when KU is absent, DNA ends are exposed to end resection proteins which would promote the generation of larger deletions. Similar results have been described previously with ZFNs-induced DSB repair in a *ku80* mutant and in *ku70* and *lig4* mutants (Osakabe *et al.* 2010; Qi *et al.* 2013)

We previously showed that PARP1 and PARP2 are involved in the MMEJ repair pathway by an *in vitro* end joining assay in *Arabidopsis* (Jia *et al.* 2013). In the *in vivo* end-joining experiments described here, however, we did not observe a role for PARP1 and PARP2 in MMEJ, and therefore there must be another repair pathway independent of PARP1 and PARP2 that uses microhomology. It is still elusive whether b-NHEJ is a single pathway or a category containing multiple mechanisms (Deriano and Roth 2013). The similar mutation

characteristics observed in the *parp1parp2* mutant and the wild type supports the dominant role of the KU-dependent c-NHEJ pathway rather than the PARP-dependent b-NHEJ pathway. Notably, we did not observe much differences in junction characteristics between *ku80* and *ku80parp1parp2* mutants, indicating other repair pathways (independent of PARP) with similar characteristics become active when c-NHEJ is absent. However, we cannot rule out that other factors, for example PARP3 (Rulten *et al.* 2011; Langelier *et al.* 2014), could slip into the b-NHEJ pathway without disturbing its outcome.

Insertions were found at both c-NHEJ-proficient and -deficient repair junctions, although most junctions were repaired without an insertion. *ku80* mutants had more insertion events than the wild type. However, the median insertion length in break junctions is a few base pairs and no statistically significant difference was observed among mutants and wild type. Besides, larger deletions (of more than 20 bp) were found with insertions compared to those in repair products without insertions during Cas9-induced repair. Templated insertions were observed both in c-NHEJ efficient and c-NHEJ deficient mutants in animal cells. The current models of templated mutagenesis are based on a MMEJ mechanism involving DNA Polymerase θ (McVey and Lee 2008; Koole *et al.* 2014; Roerink *et al.* 2014). In plants, templated insertions were also observed after DSB induced repair in *Arabidopsis* and tobacco (Shirley *et al.* 1992; Gorbunova and Levy 1997; Salomon and Puchta 1998; Lloyd *et al.* 2012; Vu *et al.* 2014). Furthermore, a recent study showed that the *Arabidopsis* Pol θ ortholog *Tebichi* (*Teb*) is essential for T-DNA integration (van Kregten *et al.* 2016). Templated insertions were found at the repair junctions of T-DNA inserts, and it was shown that *teb* mutants were resistant to T-DNA integration and very sensitive to the DNA damaging agents bleomycin and MMS. Our results indicate that nuclease-induced DSBs may be repaired by Ku-dependent NHEJ, or a backup pathway in the absence of Ku, leading to larger deletions in the latter case. Templated insertions, which have the hallmarks of Pol θ -mediated repair, may be formed in both cases, but in with a higher frequency in the absence of Ku, revealing a complex interplay of repair factors during DSB repair in *Arabidopsis*.

Acknowledgements

We thank the Puchta lab for supplying the pEn-Chimera and pDe-Cas9 vectors. We also like to thank BSc internship students Manoah van de Velde and Kübra Kontbay, and MSc internship student Kelly van Kooperen for their technical assistance. This work was financially supported by the Partnership Program STW-Rijk Zwaan of the Dutch Technology Foundation STW, which is part of the Netherlands Organisation for Scientific Research (NWO), and which is partly funded by the Ministry of Economic Affairs, Agriculture and Innovation (12428). HS was supported by a scholarship from the Chinese Scholarship Council (CSC) and PJJH by the fellowship associated with the appointment as KNAW Academy Professor.

Table 1. TALEN RVDs

TAL	RVDs
TAL-CRU-1-left	NI HD HD NI NN NN NG HD NN NG NN NI NN NN HD
TAL-CRU-1-right	HD HD NI HD NG HD HD NG HD NN HD NG HD NG HD NN NG NI

Table 2. Primers used for cloning and PCR reactions.

Primer	Sequence	Used for
SP509	ATTGAGGAGACTATCTGCAGCATG	sgRNA cloning CRU3
SP510	AAACCATGCTGCAGATAGTCTCCT	sgRNA cloning CRU3
SP512	ATTGTTGCTGTTGAACTACATTGG	sgRNA cloning PPO
SP513	AAACCCAATGTAGTTCAACAGCAA	sgRNA cloning PPO
SP491	GCTTCAGAACCAACAAGACAGC	CRU3 target sense (TALEN)
SP492	TGAGCCTGACATACTCCAAG	CRU3 target antisense (TALEN/HRM)
SP245	TGCCAACACTCCAGGCTCTG	CRU3 target sense (CRISPR/Cas9)
SP248	CAAGTGGTCAACGACAACGG	CRU3 target antisense (CRISPR/Cas9)
SP392	CACTTTGACAGATTAGGTAG	PPO target sense (CRISPR/Cas9)
SP538	CTTCCACTAACTCACCTTC	PPO target antisense (CRISPR/Cas9)
SP560	CTCCTCACTCTTTTCCAAATCG	PPO target sense (HRM)
SP561	AGATGTGTTACAAGTGTGCTG	PPO target antisense (HRM)
SP563	ACCTCTAAGACAGCCCTACG	CRU3 target sense (HRM)

Table 3. Distribution of deletion length for the target site derived from the indicated genotypes.

Nuclease mutants	Deletion length				Total deletions
	1-9 bp	10-19 bp	20-49 bp	≥50 bp	
WT Cas9-CRU	15 (57.7%)	6 (23.1%)	4 (15.4%)	1 (3.8%)	26
<i>ku80</i> cas9-CRU	10 (28.6%)	3 (8.6%)	14 (40%)	8 (22.8%)	35
<i>p1p2</i> Cas9-CRU	13 (46.4%)	6 (21.4%)	7 (25%)	2 (7.2%)	28
<i>ku80p1p2</i> Cas9-CRU	12 (30.8%)	3 (7.7%)	19 (48.7%)	5 (12.8%)	39
WT Cas9-PPO	21 (33.3%)	20 (31.7%)	14 (22.2%)	8 (12.8%)	63
<i>ku80</i> Cas9-PPO	2 (6.3%)	6 (18.7%)	14 (43.7%)	10 (31.3%)	32
<i>p1p2</i> Cas9-PPO	9 (27.3%)	9 (27.3%)	8 (24.2%)	7 (21.2%)	33
<i>ku80p1p2</i> Cas9-PPO	4 (12.1%)	5 (15.2%)	16 (48.5%)	8 (24.2%)	33

Table 4. Insertions and templated-insertions.

Nuclease mutants	Total mutations	Insertions	Templated-insertions >3 bp
WT TALEN-CRU-1	30	2 (6.6%)	0
<i>ku80</i> TALEN-CRU-1	42	9 (21.4%)	3 (7.1%)
<i>p1p2</i> TALEN-CRU-1	36	7 (19.5%)	1 (2.8%)
<i>ku80p1p2</i> TALEN-CRU-1	26	5 (19.2%)	0
WT Cas9-CRU	26	4 (15.4%)	0
<i>ku80</i> Cas9-CRU	35	7 (20.0%)	6 (17.1%)
<i>p1p2</i> Cas9-CRU	28	5 (17.9%)	0
<i>ku80p1p2</i> Cas9-CRU	39	11 (28.2%)	5 (12.8%)
WT Cas9-PPO	63	12 (19.1%)	4 (6.3%)
<i>ku80</i> Cas9-PPO	32	7 (21.9%)	4 (12.5%)
<i>p1p2</i> Cas9-PPO	33	10 (30.3%)	1 (3.0%)
<i>ku80p1p2</i> Cas9-PPO	33	10 (30.3%)	6 (18.2%)

References

- Alonso, J. M., M. Alonso, A. N. Stepanova, T. J. Leisse, D. K. Stevenson *et al.*, 2009 Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301: 653–657.
- Cermak, T., E. Doyle, and M. Christian, 2011 Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* 39: e82.
- Christian, M., T. Cermak, E. L. Doyle, C. Schmidt, F. Zhang *et al.*, 2010 Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* 186: 756–761.
- Clough, S. J., and A. F. Bent, 1998 Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16: 735–743.
- Cong, L., F. A. Ran, D. Cox, S. Lin, R. Barretto *et al.*, 2013 Multiplex genome engineering using CRISPR/Cas systems. *Science* 339: 819–823.
- Curtis, M. D., and U. Grossniklaus, 2003 A Gateway cloning vector set for high-throughput functional analysis of genes in planta. *Breakthr. Technol.* 133: 462–469.
- Deriano, L., and D. B. Roth, 2013 Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. *Annu. Rev. Genet.* 47: 433–455.
- Downs, J. a, and S. P. Jackson, 2004 A means to a DNA end: the many roles of Ku. *Nat. Rev. Mol. Cell Biol.* 5: 367–378.
- Farboud, B., and B. J. Meyer, 2015 Dramatic enhancement of genome editing by CRISPR/cas9 through improved guide RNA design. *Genetics* 199: 959–971.
- Fauser, F., S. Schiml, and H. Puchta, 2014 Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in *Arabidopsis thaliana*. *Plant J.* 348–359.
- Fell, V. L., and C. Schild-Poulter, 2014 The Ku heterodimer: Function in DNA repair and beyond. *Mutat. Res. Mutat. Res.* 763: 15–29.
- Gorbunova, V., and A. A. Levy, 1997 Non-homologous DNA end joining in plant cells is associated with deletions and filler DNA insertions. *Nucleic Acids Res.* 25: 4650–4657.

- Jia, Q., P. Bundock, P. J. J. Hooykaas, and S. de Pater, 2012 *Agrobacterium tumefaciens* T-DNA Integration and Gene Targeting in *Arabidopsis thaliana* Non-Homologous End-Joining Mutants. *J. Bot.* 2012: 1–13.
- Jia, Q., A. Den Dulk-Ras, H. Shen, P. J. J. Hooykaas, and S. de Pater, 2013 Poly(ADP-ribose)polymerases are involved in microhomology mediated back-up non-homologous end joining in *Arabidopsis thaliana*. *Plant Mol. Biol.* 82: 339–351.
- Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna *et al.*, 2012 A Programmable Dual-RNA – Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* 337: 816–822.
- Jinek, M., F. Jiang, D. W. Taylor, S. H. Sternberg, E. Kaya *et al.*, 2014 Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* 343: 1247997.
- Koole, W., R. van Schendel, A. E. Karambelas, J. T. van Heteren, K. L. Okihara *et al.*, 2014 A Polymerase Theta-dependent repair pathway suppresses extensive genomic instability at endogenous G4 DNA sites. *Nat. Commun.* 5: 3216.
- van Kregten M., S. de Pater, R. Romeijn, R. van Schendel, P. Hooykaas and M. Tijsterman, 2016 T-DNA integration in plants results from Polymerase Theta-mediated DNA repair. *Nature plants* 2:1-6.
- Langelier, M. F., A. a. Riccio, and J. M. Pascal, 2014 PARP-2 and PARP-3 are selectively activated by 5' phosphorylated DNA breaks through an allosteric regulatory mechanism shared with PARP-1. *Nucleic Acids Res.* 42: 7762–7775.
- Lloyd, A. H., D. Wang, and J. N. Timmis, 2012 Single molecule PCR reveals similar patterns of non-homologous DSB repair in tobacco and arabidopsis. *PLoS One* 7(2): e32255.
- Luijsterburg, M. S., I. de Krijger, W. W. Wiegant, R. G. Shah, G. Smeenk *et al.*, 2016 PARP1 Links CHD2-Mediated Chromatin Expansion and H3.3 Deposition to DNA Repair by Non-homologous End-Joining. *Mol. Cell* 61: 547–562.
- McVey, M., and S. E. Lee, 2008 MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet.* 24: 529–538.
- Nekrasov, V., B. Staskawicz, D. Weigel, J. D. G. Jones, and S. Kamoun, 2013 Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* 31: 691–693.
- Nishimasu, H., F. A. Ran, P. D. Hsu, S. Konermann, S. I. Shehata *et al.*, 2014 Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* 156: 935–949.
- Osakabe, K., Y. Osakabe, and S. Toki, 2010 Site-directed mutagenesis in *Arabidopsis* using custom-designed zinc finger nucleases. *Proc. Natl. Acad. Sci. USA.* 107: 12034–12039.
- de Pater, S., L. W. Neuteboom, J. E. Pinas, P. J. J. Hooykaas, and B. J. van der Zaal, 2009 ZFN-induced mutagenesis and gene-targeting in *Arabidopsis* through *Agrobacterium*-mediated floral dip transformation. *Plant Biotechnol. J.* 7: 821–835.
- de Pater, S., J. E. Pinas, P. J. J. Hooykaas, and B. J. van der Zaal, 2013 ZFN-mediated gene targeting of the *Arabidopsis* protoporphyrinogen oxidase gene through *Agrobacterium*-mediated floral dip transformation. *Plant Biotechnol. J.* 11: 510–515.
- Puchta, H., and F. Fauser, 2013a Gene targeting in plants: 25 years later. *Int. J. Dev. Biol.* 57: 629–637.
- Puchta, H., and F. Fauser, 2013b Synthetic nucleases for genome engineering in plants: prospects for a bright future. *Plant J.* 78: 727–741.
- Qi, Y., Y. Zhang, F. Zhang, J. a. Baller, S. C. Cleland *et al.*, 2013 Increasing frequencies of site-specific

- mutagenesis and gene targeting in Arabidopsis by manipulating DNA repair pathways. *Genome Res.* 23: 547–554.
- Roerink, S. F., R. Schendel, and M. Tijsterman, 2014 Polymerase theta-mediated end joining of replication-associated DNA breaks in *C. elegans*. *Genome Res.* 24: 954–962.
- Rulten, S. L., A. E. O. Fisher, I. Robert, M. C. Zuma, M. Rouleau *et al.*, 2011 PARP-3 and APLF function together to accelerate nonhomologous end-joining. *Mol. Cell* 41: 33–45.
- Salomon, S., and H. Puchta, 1998 Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. *EMBO J.* 17: 6086–6095.
- Shirley, B. W., S. Hanley, and H. M. Goodman, 1992 Effects of ionizing radiation on a plant genome: analysis of two Arabidopsis transparent testa mutations. *Plant Cell* 4: 333–347.
- Voytas, D. F., 2013 Plant genome engineering with sequence-specific nucleases. *Annu. Rev. Plant Biol.* 64: 327–350.
- Vu, G. T. H., H. X. Cao, K. Watanabe, G. Hensel, F. R. Blattner *et al.*, 2014 Repair of site-specific DNA double-strand breaks in barley occurs via diverse pathways primarily involving the sister chromatid. *Plant Cell* 26: 2156–2167.
- Walker, J. R., R. A. Corpina, and J. Goldberg, 2001 Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature* 412: 607–614.
- Wang, M., W. Wu, W. Wu, B. Rosidi, L. Zhang *et al.*, 2006 PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways. *Nucleic Acids Res.* 34: 6170–6182.

Summary

Double-strand breaks (DSBs) are one of the most lethal forms of DNA damage. DSBs can occur during normal cellular metabolism or can be induced by external factors, and highly threaten genomic integrity and cell survival (Deriano and Roth 2013). To prevent this, cells have evolved complex and highly conserved systems to detect these lesions, signal their presence, trigger various downstream events and finally bring about repair. Two main pathways are used for DNA DSB repair: Homologous Recombination (HR) and Non-Homologous End-Joining (NHEJ). Both of them function together to maintain genome integrity. The HR pathway precisely restores the genomic sequence of the broken DNA ends. This requires the sister chromatid as a template for accurate repair. In contrast, NHEJ promotes direct ligation of the DSB ends without the requirement for sequence homology and may result in small insertions and deletions at the break site. Both pathways are highly conserved throughout eukaryotic evolution but their relative importance may be different depending on the stage of the cell cycle or the cell type. Unicellular eukaryotes with small genomes such as yeast (*Saccharomyces cerevisiae*) mostly rely on HR to repair DSBs, whereas in higher eukaryotes, like mammals and plants with large genomes containing many repeat sequences, the NHEJ pathway is the predominant repair pathway. At least two NHEJ pathways have been identified: the classic NHEJ pathway (c-NHEJ) and the backup-NHEJ pathway (b-NHEJ) also called alternative-NHEJ (a-NHEJ) or microhomology-mediated end-joining (MMEJ).

Agrobacterium tumefaciens is widely used as a vector to produce genetically modified plants. *Agrobacterium*-mediated genetic transformation involves the transfer of T-DNA from its tumor-inducing plasmid to the host cell nucleus, where it integrates into the plant genome. However, the molecular mechanism of T-DNA integration is still unclear. T-DNAs can integrate at artificially induced DSBs, which suggests that DSB repair mechanisms are probably involved in T-DNA integration in plants (Salomon and Puchta 1998). Moreover, it was shown in our lab that *Agrobacterium* T-DNA integration in yeast (*S.cerevisiae*) depends on NHEJ proteins (van Attikum *et al.* 2001; van Attikum and Hooykaas 2003). Arabidopsis NHEJ mutants have subsequently been studied for T-DNA integration. However, the results obtained by different research groups were variable and revealed either no or limited negative effects (Friesner and Britt 2003; van Attikum *et al.* 2003; Gallego *et al.* 2003; Li *et al.* 2005). Disruption of multiple DNA repair pathways at the same time did not eliminate transformation (Jia *et al.* 2012; Mestiri *et al.* 2014; Park *et al.* 2015), suggesting that there must be other unknown proteins and pathways that mediate T-DNA integration in plants. Furthermore, a recent study showed that the Arabidopsis Polymerase θ (Pol θ) ortholog Tebichi (Teb) is essential for T-DNA integration (Kregten *et al.* 2016).

In **Chapter 1**, I review the current knowledge of the DNA damage response, the DSB repair pathways and their regulation, how these repair pathways affect DNA repair and *Agrobacterium*-mediated T-DNA integration and the artificial nuclease techniques for inducing DSBs. Compared to mammals, in which the NHEJ pathways have been well defined, there is still much to learn about NHEJ repair pathways in plants.

Chapter 2 describes the involvement of *Parp3* and *Xrcc1* in DNA repair and the effect of a combination of deficiencies of different NHEJ factors on T-DNA integration in Arabidopsis. Homozygous *parp3* and *xrcc1* mutants were isolated and characterized, and

parp1parp3, *ku80xrcc1* double mutants and the *parp1parp2parp3* triple mutant were obtained by crossing. The results from DNA damaging treatments showed that Parp3 and Xrcc1 are involved in DNA repair. We further examined transient and stable root transformation frequencies of these mutants together with the *ku70*, *ku80*, *parp1*, *parp2*, *parp1parp2*, *lig4*, *lig6* and *lig4lig6* mutants, which were characterized by our lab previously, after co-cultivation with *Agrobacterium*. The aim of this study was to investigate NHEJ pathways and analyze whether *Agrobacterium*-mediated transformation efficiency was affected by the absence of NHEJ factors in Arabidopsis. Deficiency in either the c-NHEJ or b-NHEJ pathway, did not lead to a significant decrease in root transformation. However, the *ku80xrcc1* and *ku80p1p2* mutants showed a significant decrease in stable root transformation efficiency. However, no significant differences were observed in transient transformation. These results indicate that the known NHEJ repair pathways are required for optimal T-DNA integration, but that there must still be other important factors and/or pathways involved in T-DNA integration in Arabidopsis.

Chapter 3 describes experiments done to reveal a putative role of Lig1a in DNA repair. The *lig1a* mutant was isolated and characterized, and the *lig4lig1a* double mutant and the *lig4lig6lig1a* triple mutant were obtained by crossing. Genotoxic tests showed that *lig1a* must have a role in repair as *lig4lig6lig1a* was more sensitive to DNA damage than *lig4lig6*. Also the comet assay revealed more DNA damage in the triple mutant. These results suggested that Lig1a probably plays a role in DNA repair when Lig4 and Lig6 are both mutated.

Chapter 4 describes an important function of Mre11 for b-NHEJ pathways involved in T-DNA integration. Our results from yeast-two-hybrid analysis showed that the interaction domain necessary for the formation of a Mre11-Rad50 complex is still present in the Mre11-2 protein, but absent in Mre11-4. This probably explains the phenotypic differences between the two mutants. In order to study whether Mre11 also functions in b-NHEJ in plants, the *ku80mre11-2* double mutant was obtained by crossing. We found that the *ku80mre11-2* double mutant was more sensitive to DNA damage and exhibited more cell death in the roots than each of the single mutants. Furthermore, the root transformation assays showed that the transformation frequencies of the single mutants were not significantly different from the wild type. However, the *ku80mre11-2* double mutant is fully resistant to *Agrobacterium* mediated T-DNA integration. These results suggested that Ku80 and Mre11 are involved in two different pathways of T-DNA integration, each of which is not essential, but together are responsible for all T-DNA integrations in plants.

Chapter 5 focuses on how exactly the DNA ends join in the Arabidopsis NHEJ mutants *in vivo*, which were deficient in c-NHEJ, b-NHEJ or both. TALENs and CRISPR/Cas9 were used for DSB-mediated targeted mutagenesis in the *cruciferin 3* (*CRU3*) and *protoporphyrinogen oxidase* (*PPO*) genes. Both nucleases were expressed and successfully induced DSBs in *ku80*, *parp1parp2* and *ku80parp1parp2* mutants. The results from footprint analyses showed that larger deletions predominated after DSB repair in *ku80* and *ku80parp1parp2* mutants. Furthermore, templated insertions were observed at the repair junctions more frequently in *ku80* and *ku80p1p2* mutants than in wild type and *parp1parp2* mutants, although such insertions were found in all four genotypes. These results indicate a shift to a more error-prone back up repair mechanism of DSB repair in the absence of Ku80

and that other Parp-independent back up pathways exist probably involving Pol θ mediated end-joining (TMEJ) responsible for the template insertions.

In short, the studies described in this thesis showed that back-up error-prone NHEJ repair pathways, together with classical NHEJ, are involved in *Agrobacterium*-mediated T-DNA integration. Mre11 could be a key player in this process, and together with Ku80 they are responsible for all T-DNA integration in plants. A model summarizing the main results described in this thesis is shown in **Figure 1**.

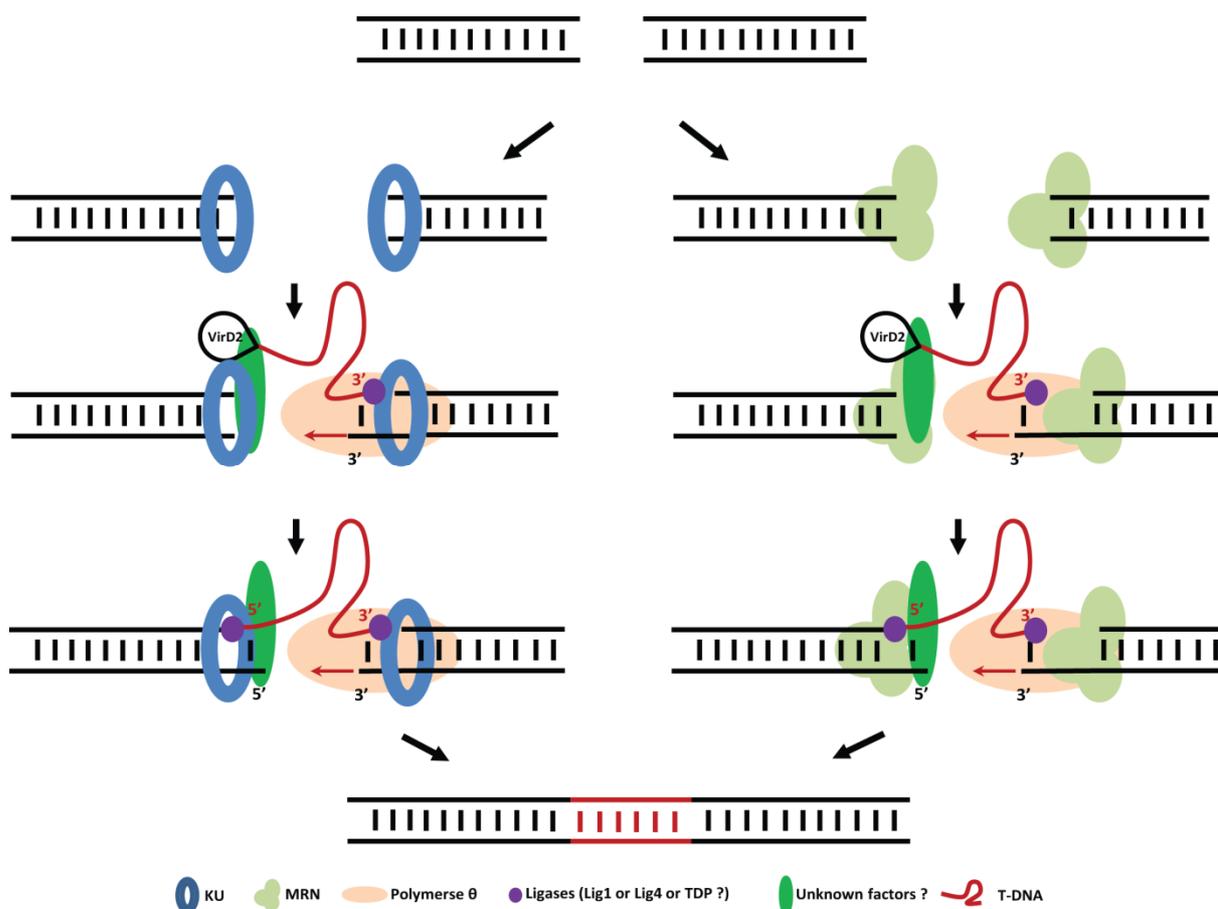


Figure 1. Model for *Agrobacterium*-mediated T-DNA integration in Arabidopsis. DSBs are recognized by Ku heterodimers or the MRN complex. Once Ku or MRN binds to the break ends, other factors including Polymerase θ (Pol θ), ligases, resection enzymes and unknown proteins are recruited to the break sites. One of these essential factors, Pol θ is responsible for the attachment of the single-stranded T-DNA left border (LB) to the plant genome by using a few bases of homology to prime DNA synthesis from the 3' end. Similar to the LB, the activities of Ku, MRN and unknown proteins may be involved in the ligation of the T-DNA right border (RB) to the other end of the break.

References

- van Attikum, H., P. Bundock, and P. J. J. Hooykaas, 2001 Non-homologous end-joining proteins are required for *Agrobacterium* T-DNA integration. *EMBO J.* 20: 6550–6558.
- van Attikum, H., P. Bundock, R. M. Overmeer, L. Y. Lee, S. B. Gelvin *et al.*, 2003 The *Arabidopsis* AtLIG4 gene is required for the repair of DNA damage, but not for the integration of *Agrobacterium* T-DNA. *Nucleic Acids Res.* 31: 4247–4255.
- van Attikum, H., and P. J. J. Hooykaas, 2003 Genetic requirements for the targeted integration of *Agrobacterium* T-DNA in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 31: 826–832.
- Deriano, L., and D. B. Roth, 2013 Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. *Annu. Rev. Genet.* 47: 433–455.
- Friesner, J., and A. B. Britt, 2003 Ku80- and DNA ligase IV-deficient plants are sensitive to ionizing radiation and defective in T-DNA integration. *Plant J.* 34: 427–440.
- Gallego, M. E., J.Y. Bleuyard, S. Daoudal-Cotterell, N. Jallut, and C. I. White, 2003 Ku80 plays a role in non-homologous recombination but is not required for T-DNA integration in *Arabidopsis*. *Plant J.* 35: 557–565.
- Jia, Q., P. Bundock, P. J. J. Hooykaas, and S. de Pater, 2012 *Agrobacterium tumefaciens* T-DNA Integration and Gene Targeting in *Arabidopsis thaliana* Non-Homologous End-Joining Mutants. *J. Bot.* 2012: 1–13.
- Kregten, M. Van, S. De Pater, R. Romeijn, R. Van Schendel, P. J. J. Hooykaas *et al.*, 2016 T-DNA integration in plants results from Polymerase Theta-mediated DNA repair. *Nat. Plants* 2: 1–6.
- Li, J., M. Vaidya, C. White, A. Vainstein, V. Citovsky *et al.*, 2005 Involvement of KU80 in T-DNA integration in plant cells. *Proc. Natl. Acad. Sci. USA.* 102: 19231–19236.
- Mestiri, I., F. Norre, M. E. Gallego, and C. I. White, 2014 Multiple host-cell recombination pathways act in *Agrobacterium*-mediated transformation of plant cells. *Plant J.* 77: 511–520.
- Park, S.Y., Z. Vaghchhipawala, B. Vasudevan, L.Y. Lee, Y. Shen *et al.*, 2015 *Agrobacterium* T-DNA integration into the plant genome can occur without the activity of key non-homologous end-joining proteins. *Plant J.* 81: 934–946.
- Salomon, S., and H. Puchta, 1998 Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. *EMBO J.* 17: 6086–6095.

Samenvatting

Dubbelstrengs breuken (DSB-en) zijn één van de meest dodelijke vormen van DNA schade. DSB-en kunnen optreden tijdens normaal cellulair metabolisme of kunnen worden veroorzaakt door externe factoren, en bedreigen genomische integriteit en celoverleving (Deriano en Roth 2013). Om dit te voorkomen, hebben cellen complexe en geconserveerde systemen ontwikkeld om DSB-en te detecteren, hun aanwezigheid te signaleren en verschillende stroomafwaartse reacties plaats te laten vinden die uiteindelijk leiden tot herstel. Twee belangrijke trajecten worden gebruikt voor DNA-DSB reparatie: Homologe Recombinatie (HR) en Niet-Homologe End-Joining (NHEJ). Beiden werken samen om de integriteit van het genoom te behouden. Het homologe recombinatie traject herstelt nauwkeurig de DNA volgorde van het gebroken DNA met behulp van de zuster chromatide als sjabloon voor nauwkeurige reparatie. Daarentegen bevordert NHEJ directe ligatie van de DSB uiteinden zonder gebruik te maken van een sjabloon, wat kan leiden tot kleine inserties en deleties op de plaats van herstel. Beide routes zijn sterk geconserveerd tijdens de evolutie van eukaryoten maar hun relatieve belang kan verschillen afhankelijk van het stadium van de celcyclus of celtype. Eencellige eukaryoten met kleine genomen zoals gist (*Saccharomyces cerevisiae*) gebruiken voornamelijk homologe recombinatie om DSB-en te repareren, terwijl in hogere eukaryoten, zoals zoogdieren en planten met grote genomen met veel herhalings-sequenties, de NHEJ route de meest gebruikte reparatie route is. Ten minste twee NHEJ routes zijn geïdentificeerd: de klassieke NHEJ route (c-NHEJ) en de back-up-NHEJ route (b-NHEJ) ook wel alternatieve-NHEJ (a-NHEJ) genoemd of microhomology-gemedieerde end-joining (MMEJ).

Agrobacterium tumefaciens wordt tegenwoordig veel gebruikt als vector voor het genetisch modificeren van planten. Tijdens *Agrobacterium*-gemedieerde genetische transformatie, wordt het T-DNA afkomstig van het tumor-inducerende plasmide naar de kern van de gastheercel getransporteerd, waar het integreert in het plantengenoom. Het moleculaire mechanisme van T-DNA integratie is nog onduidelijk. T-DNA strengen zijn in staat om te integreren in kunstmatig gemaakte DSB-en, wat suggereert dat DSB herstelmechanismen waarschijnlijk betrokken zijn bij T-DNA-integratie in planten (Salomon en Puchta 1998). Bovendien werd aangetoond in ons lab dat *Agrobacterium* T-DNA integratie in gist (*S. cerevisiae*) gemedieerd wordt door c-NHEJ (van Attikum et al 2001; Van Attikum en Hooykaas 2003). *Arabidopsis* NHEJ mutanten zijn vervolgens onderzocht op T-DNA-integratie. Echter, de resultaten verkregen door verschillende onderzoeksgroepen waren variabel en lieten ofwel een duidelijk negatief effect zien (Li et al 2005) of geen of maar een zeer beperkt negatief effect van een NHEJ mutatie op T-DNA integratie (Friesner en Britt 2003; Van Attikum et al 2003; Gallego et al 2003). Ook verstoring van meerdere DNA herstelmechanismen tegelijkertijd elimineerden niet *Agrobacterium*-gemedieerde transformatie (Jia et al 2012; Mestiri et al 2014; Park et al 2015). Dit suggereerde dat er andere onbekende herstelroutes moeten zijn voor T-DNA integratie in planten. Inderdaad toonde een recente studie in ons lab aan dat de *Arabidopsis* ortholoog van Polymerase θ (Tebichi; Teb) essentieel is voor T-DNA-integratie (Kregten et al. 2016).

In **hoofdstuk 1** wordt een overzicht gepresenteerd van onze huidige kennis van detectie, herstel en regulering van DNA-schade, hoe deze herstelmechanismen *Agrobacterium*-gemedieerde T-DNA-integratie beïnvloeden en hoe artificiële nucleasen zoals

CRISPR/Cas gebruikt kunnen worden voor het induceren van DSB-en. Hoewel het proces van NHEJ in zoogdieren en gist goed is gedefinieerd, is er nog veel te leren over NHEJ herstelmechanismen in planten.

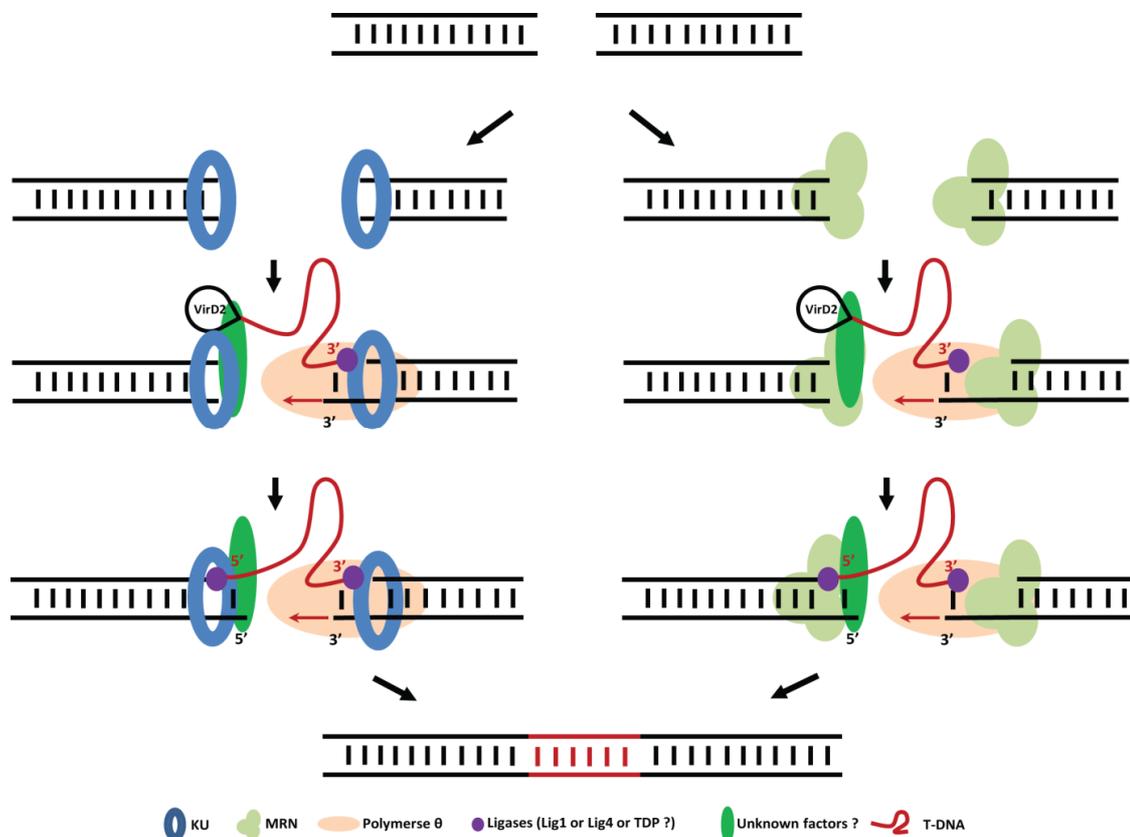
Hoofdstuk 2 beschrijft de betrokkenheid van Parp3 en XRCC1 bij DNA herstel en het effect van een combinatie van mutaties van verschillende NHEJ factoren op T-DNA integratie in Arabidopsis. Homozygote *parp3* en *xrcc1* mutanten werden geïsoleerd en gekarakteriseerd, en *parp1parp3*, *ku80xrcc1* dubbele mutanten en de *parp1parp2parp3* drievoudige mutant werden verkregen door kruisen. De resultaten van DNA-beschadigende behandelingen liet zien dat Parp3 en XRCC1 betrokken zijn bij DNA-herstel. We hebben gekeken naar transiënte en stabiele worteltransformatie in deze mutanten samen met de *ku70*, *ku80*, *parp1*, *parp2*, *parp1parp2*, *lig4*, *lig6* en *lig4lig6* mutanten, die eerder werden gekarakteriseerd door ons lab. Het doel van deze studie was om de NHEJ herstel routes te onderzoeken en te analyseren of de efficiëntie van Agrobacterium-gemedieerde transformatie beïnvloed wordt door de afwezigheid van NHEJ factoren in Arabidopsis.

Deficiëntie in een van de NHEJ routes (c-NHEJ of b-NHEJ), leidde niet tot een significante afname van worteltransformatie. Mutaties in beide NHEJ herstelroutes (*ku80xrcc1* en *ku80p1p2* mutanten) leidde wel tot een significante afname in de efficiency van stabiele worteltransformatie. Er werden echter geen significante verschillen waargenomen bij transiënte transformatie. Deze resultaten geven aan dat de bekende NHEJ herstelmechanismen nodig zijn voor optimale T-DNA integratie, maar dat er nog andere belangrijke factoren of routes betrokken zijn bij T-DNA integratie in Arabidopsis.

Hoofdstuk 3 beschrijft een rol van Lig1a bij DNA-herstel. De homozygote *lig1a* mutant werd geïsoleerd en gekarakteriseerd, en de *lig4lig1a* dubbele mutant en de *lig4lig6lig1a* drievoudige mutant werden verkregen door kruisen. Genotoxische testen toonden aan dat de *lig1a* mutatie zorgde voor extra gevoeligheid in de *lig4lig6* achtergrond (*lig4lig6lig1a* in vergelijking met *lig4lig6*). De resultaten uit de komeet-test lieten een effect in de drievoudige mutant zien ten opzichte van de *lig4lig1a* dubbel mutant. Deze resultaten geven aan dat Lig1a waarschijnlijk een rol speelt bij DNA-herstel wanneer Lig4 en Lig6 beide defect zijn.

Hoofdstuk 4 analyseert het MRE11 eiwit, dat belangrijk is onder andere voor b-NHEJ in zoogdieren. De resultaten van de 'yeast-two-hybrid' analyse toonden aan dat het interactie domein noodzakelijk voor de vorming van een MRE11-RAD50 complex nog aanwezig is in het Mre11-2 eiwit, maar afwezig in Mre11-4. Dit verklaart de fenotypische verschillen tussen de twee mutanten. Om te onderzoeken of MRE11 ook functioneert in b-NHEJ in planten, werd de dubbele mutant *ku80mre11-2* verkregen door het kruisen van de enkele mutanten. We vonden dat de *ku80mre11-2* dubbele mutant gevoeliger was voor DNA beschadiging dan beide enkele mutanten en dat er meer celdood in wortels van de dubbele mutant plaats vond dan in wortels van de enkele mutanten. Opmerkelijk was dat de *ku80mre11-2* dubbele mutant volledig resistent was geworden voor Agrobacterium-gemedieerde T-DNA-integratie, terwijl de enkele mutanten elk wel goed transformeerbaar waren door Agrobacterium. Deze resultaten suggereerden dat Ku80 en MRE11 betrokken zijn in twee verschillende trajecten van T-DNA-integratie, die elk niet essentieel zijn, maar samen verantwoordelijk voor alle T-DNA-integratie in planten.

Hoofdstuk 5 richt zich op de vraag hoe precies DNA uiteinden *in vivo* aan elkaar worden gezet in Arabidopsis NHEJ mutanten, die in c-NHEJ, b-NHEJ of in beide routes deficiënt waren. TALENs en CRISPR / Cas9 werden gebruikt voor DSB geïnduceerde gerichte mutagenese van de Arabidopsis cruciferin 3 (CRU3) en protoporphyrinogeenoxidase (PPO) genen. Beide nucleasen werden tot expressie gebracht en induceerden DSB-en in *ku80*, *parp1parp2* en *ku80parp1parp2* mutanten. De analyses van de DNA volgorden van herstelde DNA breuken toonden aan dat grotere deleties overheersten na DSB reparatie in *ku80* en *ku80parp1parp2* mutanten. Verder werden vaker ‘templated’ inserties waargenomen bij de reparatie in *ku80* en *ku80p1p2* mutanten dan in wildtype en *parp1parp2* mutanten, hoewel dergelijke ‘templated’ inserties werden gevonden in alle vier de genotypes. Deze resultaten geven aan dat er een verschuiving plaats vindt naar een foutgevoelige back up herstelroute van DSB-en in afwezigheid van Ku80 en dat er andere PARP-onafhankelijke back up wegen bestaan die waarschijnlijk verantwoordelijk zijn voor de ‘templated’ inserties in Arabidopsis.



Figuur 1. Model voor Agrobacterium-gemedieerde T-DNA-integratie in Arabidopsis. Uiteinden van DSB-en worden gebonden door Ku heterodimeren en/of het MRN complex. Zodra Ku of MRN binden aan de uiteinden, worden andere factoren, zoals Polymerase θ (Pol θ), ligasen, resectie enzymen en onbekende eiwitten aangetrokken. Eén van deze essentiële factoren is Pol θ , verantwoordelijk voor de baseparing van de enkelstrengs T-DNA linker border (LB) aan het plantengenoom door een of enkele complementaire nucleotiden gevolgd door DNA synthese vanaf het 3' uiteinde. Het kan ook zijn dat Ku en MRN uitsluitend betrokken zijn bij het integratieproces van de T-DNA rechter border (RB).

Kortom, de in dit proefschrift beschreven studies laten zien dat foutgevoelige back-up NHEJ herstel wegen, samen met de klassieke NHEJ weg, betrokken zijn bij *Agrobacterium*-gemedieerde T-DNA integratie. MRE11 is een belangrijke speler in dit proces, en samen met Ku80 verantwoordelijk voor alle T-DNA-integratie in planten. Een model met de belangrijkste resultaten van dit proefschrift, wordt getoond in **Figuur 1**.

References

- van Attikum, H., P. Bundock, and P. J. J. Hooykaas, 2001 Non-homologous end-joining proteins are required for *Agrobacterium* T-DNA integration. *EMBO J.* 20: 6550–6558.
- van Attikum, H., P. Bundock, R. M. Overmeer, L. Y. Lee, S. B. Gelvin *et al.*, 2003 The Arabidopsis AtLIG4 gene is required for the repair of DNA damage, but not for the integration of *Agrobacterium* T-DNA. *Nucleic Acids Res.* 31: 4247–4255.
- van Attikum, H., and P. J. J. Hooykaas, 2003 Genetic requirements for the targeted integration of *Agrobacterium* T-DNA in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 31: 826–832.
- Deriano, L., and D. B. Roth, 2013 Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. *Annu. Rev. Genet.* 47: 433–455.
- Friesner, J., and A. B. Britt, 2003 Ku80- and DNA ligase IV-deficient plants are sensitive to ionizing radiation and defective in T-DNA integration. *Plant J.* 34: 427–440.
- Gallego, M. E., J.Y. Bleuyard, S. Daoudal-Cotterell, N. Jallut, and C. I. White, 2003 Ku80 plays a role in non-homologous recombination but is not required for T-DNA integration in Arabidopsis. *Plant J.* 35: 557–565.
- Jia, Q., P. Bundock, P. J. J. Hooykaas, and S. de Pater, 2012 *Agrobacterium tumefaciens* T-DNA Integration and Gene Targeting in Arabidopsis thaliana Non-Homologous End-Joining Mutants. *J. Bot.* 2012: 1–13.
- Kregten, M. Van, S. De Pater, R. Romeijn, R. Van Schendel, P. J. J. Hooykaas *et al.*, 2016 T-DNA integration in plants results from Polymerase Theta-mediated DNA repair. *Nat. Plants* 2: 1–6.
- Li, J., M. Vaidya, C. White, A. Vainstein, V. Citovsky *et al.*, 2005 Involvement of KU80 in T-DNA integration in plant cells. *Proc. Natl. Acad. Sci. USA.* 102: 19231–19236.
- Mestiri, I., F. Norre, M. E. Gallego, and C. I. White, 2014 Multiple host-cell recombination pathways act in *Agrobacterium*-mediated transformation of plant cells. *Plant J.* 77: 511–520.
- Park, S.Y., Z. Vaghchhipawala, B. Vasudevan, L.Y. Lee, Y. Shen *et al.*, 2015 *Agrobacterium* T-DNA integration into the plant genome can occur without the activity of key non-homologous end-joining proteins. *Plant J.* 81: 934–946.
- Salomon, S., and H. Puchta, 1998 Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. *EMBO J.* 17: 6086–6095.

Curriculum vitae

Hexi Shen was born on the 3rd of August 1984 in Dongshan, Fujian province, China. He grew up in Dongshan and finished his high school at Dongshan No.1 middle school. In 2003, he started his bachelor studies at University of Science and Engineering in Chongqing and obtained the Bachelor degree in Bioengineering in 2007. In 2008, he started his master studies in Cell Biology at Southwest University in Chongqing under the supervision of Prof. Dr. Guolu Liang. From 2009 to 2011, he worked during his master internship on a project on the genetic improvement of *Caragana intermedia*, at the Research Institute of Forestry, the Chinese Academy of Forestry, under supervision of Dr. Liwang Qi. After he obtained his Master degree, he started his PhD research which was supported by a grant from the China Scholarship Council, in the laboratory of Prof. Dr. Paul Hooykaas, under supervision of Dr. Sylvia de Pater at the Institute of Biology, Leiden University, the Netherlands. His work about the DNA double-strand breaks repair and *Agrobacterium*-mediated T-DNA integration in *Arabidopsis* is described in this thesis.