



Universiteit
Leiden
The Netherlands

DNA repair and gene targeting in plant end-joining mutants

Jia, Q.

Citation

Jia, Q. (2011, April 21). *DNA repair and gene targeting in plant end-joining mutants*. Retrieved from <https://hdl.handle.net/1887/17582>

Version: Corrected Publisher's Version

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

Downloaded from: <https://hdl.handle.net/1887/17582>

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

Chapter 5

Characterization of the plant specific DNA ligase AtLig6

Qi Jia, B. Sylvia de Pater and Paul J.J. Hooykaas

Abstract

DNA ligases catalyze the joining of DNA ends and thus play important roles in DNA replication and DNA repair. Eukaryotes possess multiple ATP-dependent DNA ligases with distinct roles in DNA metabolism. Some of them have been well characterized, such as DNA ligase I (Lig1) and DNA ligase IV (Lig4). A novel plant-specific DNA ligase has been identified, termed DNA ligase VI (Lig6), but its function is still unclear. The expression pattern analyzed via Genevestigator showed that the expression level of *AtLig6* was lower than that of *AtLig1* and *AtLig4* and was especially induced at the stages of seed germination and flowering. Two homozygous mutants of *AtLig6* were isolated and crossed with the *Atlig4* mutant to obtain the double mutants (*Atlig4lig6-1*, *Atlig4lig6-2*). All these four homozygous ligase mutants were phenotypically indistinguishable from the wild-type under normal growth conditions. The two *Atlig6* single mutants could tolerate bleomycin treatment equally well as the wild-type. The *Atlig4lig6-1* and *Atlig4lig6-2* double mutants were hypersensitive to bleomycin, but no difference was observed from the *Atlig4* single mutant. The frequency of T-DNA integration was also not disturbed by the deficiency of *AtLig6*, and the frequency of gene targeting seems not to be increased in absence of both *AtLig6* and *AtLig4*. This indicates that other ligases function in NHEJ when *AtLig6* and *AtLig4* are inactive. One candidate was identified by *in silico* searching for homologs of ATP-dependent DNA ligases in Arabidopsis, which may represent a novel plant specific DNA ligase involved in back-up non-homologous end joining (B-NHEJ) for DSB DNA repair.

Introduction

DNA ligases seal broken DNA molecules with 3' OH and 5' PO₄ ends, which is essential for many biological processes, including DNA replication, DNA repair and DNA recombination (1). On the basis of the different cofactor preferences, the large family of DNA ligases is divided into two groups (2;3). The first group consists of the NAD⁺-dependent DNA ligases, which are utilized by most eubacteria. The second comprises the ATP-dependent DNA ligases, which are utilized mainly by eukaryotes (4). Eukaryotic organisms have evolved multiple ATP-dependent ligase isoforms, including DNA ligase I (Lig1), DNA ligase III (Lig3), DNA ligase IV (Lig4) and DNA ligase VI (Lig6). Lig1 and Lig4 are expressed and conserved in all eukaryotes, whereas Lig3 is unique to vertebrates and Lig6 is plant-specific (4;5). Each ligase has a distinct function in DNA metabolism for the maintenance of genomic integrity (6).

Lig1 plays a vital role in DNA replication by joining the Okazaki fragments and also in DNA repair pathways, such as nucleotide excision repair (NER), base excision repair (BER), single strand break (SSB) repair and probably double strand break (DSB) repair (3;7). Lig4 mediates the final ligation step in the classical non-homologous end joining (C-NHEJ)

pathway (8;9), which is the predominant mechanism for DSB repair in mammals and plants. C-NHEJ joins the DNA ends directly independent of sequence homology by utilizing Ku70/80 proteins and the Lig4/XRCC4/XLF complex. The vertebrate-specific Lig3 has two variants, Lig3 α and Lig3 β . Lig3 α is ubiquitously distributed, whereas Lig3 β has only been found in testes and may function in meiotic recombination (10). Lig3 α plays a role in BER and SSB repair and interacts with XRCC1 (11). Evidence also points to Lig3 to be involved in the back-up NHEJ (B-NHEJ), which was identified to repair DSBs in absence of the major components of C-NHEJ (12;13). The B-NHEJ pathway is prone to use micro-homology, and is therefore sometimes referred to as the micro-homology mediated end joining (MMEJ) pathway (14). In mammals both Lig3 and Lig1 are involved in MMEJ, but Lig4 is not (15).

Lig6 is a novel DNA ligase that was recently discovered in higher plants by the analysis of plant genomic databases (5;16). Due to its domain structure, it is distinct from the other DNA ligases. It displays a significant sequence similarity to Lig1 (5) and contains three conserved regions in the N-terminus, which are characteristic for Pso2/Snm1 proteins. These proteins belong to the β -CASP family and play an important role in interstrand DNA crosslink repair (5;16;17). Recently, Waterworth *et al.* (17) reported that *Arabidopsis thaliana* Lig6 is required for rapid seed germination and it is a determinant of seed longevity and quality. AtLig6 is probably involved in a rapid and strong DNA DSB response, activated in the earliest stages of seed imbibitions to repair DNA damage that accumulated over time. In order to further analyze its role in DNA repair, we isolated the *Atlig6* and *Atlig4lig6* double mutants and determined the effects of the mutations on T-DNA integration and gene targeting frequencies. Further *in silico* studies on homologs of DNA ligase in *Arabidopsis* were also done to find other putative DNA ligases.

Material and methods

Plant materials

Two T-DNA Col-0 insertion lines of *Atlig6* were obtained from the SALK T-DNA collection (*Atlig6-1*: SALK_065307, *Atlig6-2*: SALK_079499). Information about them is available at <http://signal.salk.edu/cgi-bin/tdnaexpress> (18). The homozygous mutants were isolated. They were crossed with *Atlig4* (chapter 2) to get the *Atlig4lig6-1* (*Atlig6-1* crossed with *Atlig4*) and *Atlig4lig6-2* (*Atlig6-2* crossed with *Atlig4*) double mutants.

Expression profiling with Genevestigator tools

The expression pattern of three ligase genes (*AtLig6*: At1G66730, *AtLig4*: At5G57160, *AtLig1*: At1G08130) in *Arabidopsis* were analyzed with Genevestigator analysis tools (<https://www.genevestigator.com/gv/index.jsp>) using publicly available microarray data. The data were selected from the 22k Affymetrix ATH1 Genechip arrays of high quality, totaling 5747 arrays and including datasets obtained from all available datasets online.

Characterization of two *Arabidopsis* T-DNA insertion mutants of *AtLig6*

DNA was extracted from individual plants using the CTAB DNA isolation protocol (19). The T-DNA insertion site was mapped with a T-DNA Left Border (LB) specific primer LBa1 and a gene-specific primer. 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, and the PCR products were sequenced. Southern blot analysis was performed as described in chapter 2. The DIG probe was produced using the PCR DIG Labeling Mix (Roche) with specific primers SP271 and SP272 that amplified an 850-bp fragment from the T-DNA of pROK2. The RNA expression of *AtLig6* in the two mutants was also analyzed (chapter 2). Specific fragments (about 200 bp) were amplified from cDNA with pairs of primers around the T-DNA insertion sites. All sample values were normalized to the values of the house keeping gene *Roc1* (Primers Roc5.2, Roc3.3) and were presented as relative expression ratios. The value of the wild-type was set on 1. The sequences of all the primers are listed in Table 1.

Table 1. Sequences of primers used for characterization of the two *Atlig6* mutants and Q-PCR.

Name	Locus	Sequence
LBb1	T-DNA LB	5'-GCGTGGACCGCTTGCTGCAACT-3'
Sp264	<i>Atlig6-1</i>	5'-GTCAACTCTGTCAATGGTCC-3'
Sp265	<i>Atlig6-1</i>	5'-AATATCAAACACGAAGACGCAGAC-3'
Sp266	<i>Atlig6-2</i>	5'-TAAGTGCTACGGTAGTTTCTC-3'
Sp267	<i>Atlig6-2</i>	5'-CTGTTCTGTAGTAAGGCGGC-3'
Sp271	pROK2 Probe	5'-CCCGTGTCTCTCCAAATG-3'
Sp272	pROK2 probe	5'-CAGGTCCCCAGATTAGCC-3'
q5	<i>Atlig6-1</i>	5'-ATCAAGTAACTTATGGATCTGG-3'
q4	<i>Atlig6-2</i>	5'-CAAGGTTAAGCGAGATTATG-3'
q3	<i>Atlig6-2</i>	5'-GACACGGCAGACACTCTG-3'
q1	<i>Atku70</i>	5'-TCTACCACTCAGTCAACCTG-3'
q2	<i>Atku70</i>	5'-CAATAGACAAGCCATCACAG-3'
q6	<i>Atlig4</i>	5'-GACACCAACGGCACAAG-3'
q7	<i>Atlig4</i>	5'-AAGTTCAATGTATGTCAGTCCC-3'
Roc5.2	<i>Roc1</i>	5'-GAACGGAACAGGCGGTGAGTC-3'
Roc3.3	<i>Roc1</i>	5'-CCACAGGCTTCGTCCGCTTTC-3'
PPO-PA	<i>PPO</i>	5'-GTGACCGAGGCTAAGGATCGTGT-3'
PPO-1	<i>PPO</i>	5'-GCAAGGAGTTGAAACATTAG-3'
PPO-4	<i>PPO</i>	5'-CATGAAGTTGTTGACCTCAATC-3'
Sp319	<i>PPO</i>	5'-CTATCAAAGAGCACAGACAGC-3'

Assays for sensitivity to genotoxic agents

Seeds from wild-type, *Atlig6-1*, *Atlig6-2*, *Atlig4lig6-1* and *Atlig4lig6-2* were surface-

sterilized as described (20) and were germinated on solidified ½ MS medium (21) without additions or containing 0.1 µg/ml, 0.2 µg/ml, 0.3 µg/ml and 0.4 µg/ml Bleocin™ antibiotic (Calbiochem). The seedlings were scored after 3 weeks of growth.

Floral dip transformation and gene targeting

In order to test the frequency of T-DNA integration and gene targeting in the *Atlig6* mutants, floral dip transformation was performed, as described by Clough and Bent (22), with the *Agrobacterium* strain AGL1 (pSDM3900) for gene targeting using the protoporphyrinogen oxidase (PPO) system (23). Plasmid pSDM3900 is a pCambia 3200 derivative (phosphinothricin (ppt) selection marker). About 1 gram seeds were plated on solid MA medium (24) without sucrose containing 15 µg/ml ppt, 100 µg/ml timentin and 100 µg/ml nystatin to determine the transformation frequency. Phosphinothricin-resistant seedlings were scored 2 weeks after germination and the relative transformation frequency was determined compared to the wild-type. The value of the wild-type was set on 1. The rest of the seeds were all sowed on solid MA medium without sucrose containing 50 µM butafenacil, 100 µg/ml timentin (to kill *Agrobacterium* cells) and 100 µg/ml nystatin (to prevent growth of fungi) to identify gene targeting events. The butafenacil-resistant plants were analyzed with PCR to determine if they represent true gene targeting (TGT) events (Figure 1).

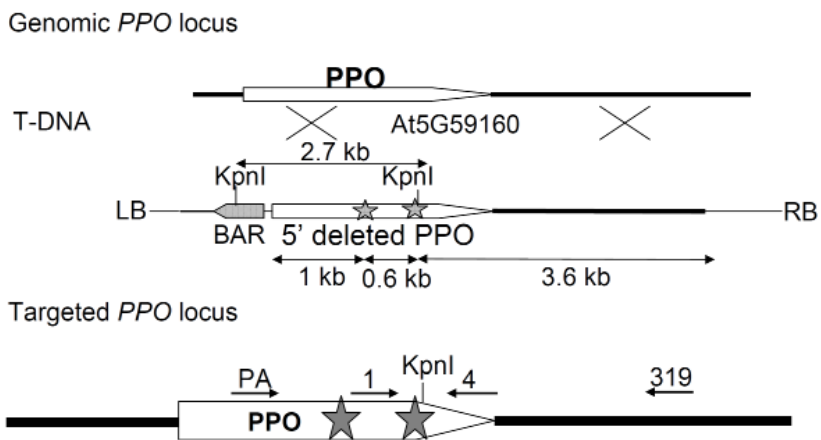


Figure 1. The design for the targeted modification of the *Arabidopsis PPO* locus. The white box marked *PPO* represents the *PPO* coding region, and the black lines represents flanking plant genomic DNA. The two mutations conferring butafenacil resistance are indicated as stars. The BAR resistance gene is linked to the truncated 5'Δ*PPO* of the T-DNA (LB for left border and RB for right border). The primers for PCR analysis of the gene targeting events are indicated by arrows.

Sequence analysis and phylogeny

The protein sequences of DNA ligases were searched using the BLAST program on the National Center for Biotechnological Information (NCBI) web page (<http://www.ncbi.nlm.nih.gov/>). AtLig1 (At1G08130), AtLig4 (At5G57160), AtLig6 (At1G66730) and DNA

ligase 3 of *Homo sapiens* (GI 73747829) were used as a query to find possible homologous proteins. Multiple sequence alignments were built using Jalview software (<http://www.jalview.org/>) with default settings. The algorithm used here is neighbour joining using % identity.

Results

Isolation and characterization of the two *Atlig6* mutants

In yeast and fungi, deletion of Lig4 leads to a disruption of NHEJ and an almost complete loss of T-DNA integration (25). In plants this is not the case (8). In the *Atlig4* mutant T-DNA is integrated with equal efficiency as in the wild-type suggesting that another ligase is responsible for T-DNA integration. Recently the plant specific ligase Lig6 was discovered (5). In order to investigate AtLig6 functions in T-DNA integration, we ordered seeds from T-DNA insertion mutants from the Salk collection and propagated these in order to obtain homozygous *Atlig6-1* and *Atlig6-2* mutants. The homozygotes were identified by PCR analysis. When two gene-specific primers flanking the insertion site were used, PCR products were amplified for wild-type and heterozygotes. No PCR products were obtained for homozygous mutants by using these two gene-specific primers, because the PCR products in the mutants would be >10 kb in size and will not be amplified with the PCR condition used here. When a T-DNA-specific primer from left border (LB) or right border (RB) was used in combination with one gene-specific primer, PCR products for the T-DNA insertion mutants were amplified, whereas no PCR products were obtained for the wild-type. We identified homozygous mutants harboring a T-DNA insertion in the *AtLig6* gene in the offspring of the heterozygous plants obtained from the Salk collection. The insertion point was mapped by sequencing of the PCR products generated using one of the T-DNA specific primers in combination with one of the gene-specific primers. For the *Atlig6-1* and *Atlig6-2* mutants, there were PCR products produced with LBa1 and both gene-specific primers (*Atlig6-1*: Sp264 and Sp265, *Atlig6-2*: Sp266 and Sp267), indicating that at least 2 T-DNA copies were inserted as an inverted repeat in the *Atlig6* locus. The combination of the primers is shown in the Figure 2. The genomic DNA was digested by HindIII for Southern blotting (Figure 2). With T-DNAs inserted in the loci identified by PCR, the amplified bands will be detected on the blot with the following sizes: for *Atlig6-1*: 2677 bp and 3584 bp and for *Atlig6-2*: 3742 bp and 4058 bp. If T-DNAs are inserted in other loci, additional bands probably with different sizes will be detected. If there are 2 or more T-DNAs inserted in one locus as direct repeat, an additional band of 4317 bp, representing a complete T-DNA, will be detected. If there are 2 or more T-DNAs inserted in one locus as inverted repeat, an additional band of 3634 bp will be detected with LBs in tail to tail orientation. Sequencing results and Southern blot analysis indicated that the T-DNAs were all inserted at the position as reported by the Salk database. A detailed characterization of the T-DNA insertions is shown in Figure 2. There were additional bands for both *Atlig6*

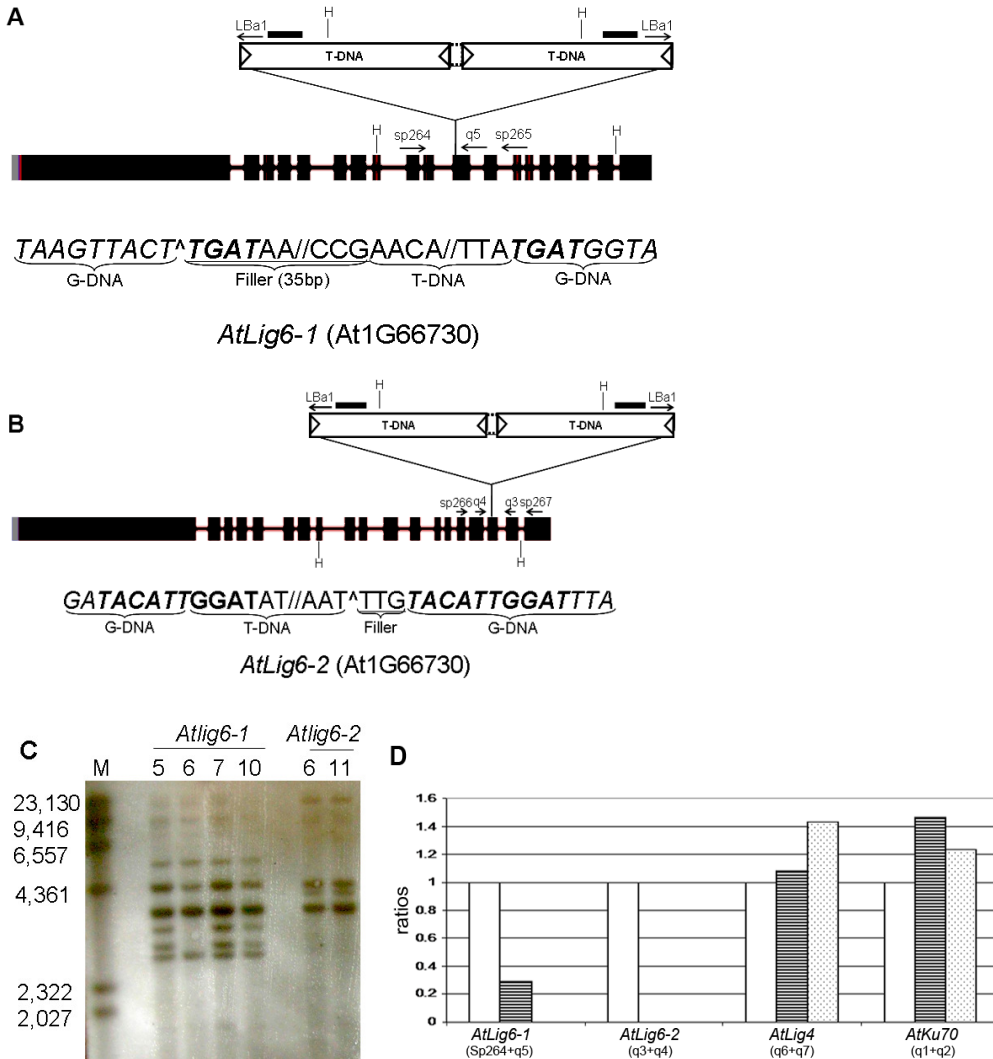


Figure 2. Molecular analysis of the T-DNA insertion in the *AtLig6* locus.

Genomic organization of the *AtLig6* locus is indicated with the positions of the inserted T-DNAs in *Atlig6-1* (A) and *Atlig6-2* (B). Exons are shown as black boxes. 3' and 5' UTRs are shown as gray boxes. Introns are shown as lines. The primers used for genotyping and Q-RT-PCR analysis are indicated. The probes (—) and the restriction enzyme digestion sites used for Southern blot analysis are also indicated. Genomic DNA sequences (g-DNA) flanking the T-DNA insertion are shown in italic. (C) Southern blot analysis of the T-DNA insertion. The genomic DNA was digested by HindIII. M: λ HindIII Marker, H: HindIII. (D) RNA expression of the *AtLig6*, *AtLig4* and *AtKu70* genes were determined by Q-RT-PCR in wild-type, *Atlig6-1* and *Atlig6-2* plants. Expression of *AtLig6* was analyzed with two different sets of primers. All the sample values were normalized to *Roc* values. The values of the wild-type were set on 1.

(□ wild-type; ▨ *Atlig6-1*; ▩ *Atlig6-2*)

mutants on the Southern blot, indicating that more than 2 T-DNAs were inserted in the genome. The T-DNA of *AtLig6-1* was integrated in exon 11 and had 35 bp filler DNA. The

two bands had segregated in the individual plant 6 of *Atlig6-1*, indicating that this plant lost two additional T-DNAs inserted in other loci, and therefore it was chosen for further research. The T-DNA of *AtLig6-2* was integrated in exon 17 and had 3 bp filler DNA. An additional band of around 4300 bp was shown for both *Atlig6* mutants, suggesting that they contained additional T-DNA copies in direct repeat.

Q-RT-PCR analysis was performed for the *Atlig6-1* and *Atlig6-2* T-DNA insertion lines using primers flanking the respective insertion sites (*Atlig6-1*: Sp264+q5, *Atlig6-2*: q3+q4). This resulted in a product with the two pairs of primers for *AtLig6* in the wild-type, but no correct products in the *Atlig6-1* and *Atlig6-2* mutants with the respective pairs of primers flanking the insertion (Figure 2). In the *Atlig6-1* mutant, a small amount of PCR product was seen, but this was shown to be of the wrong size through agarose gel electrophoresis (data not shown), suggesting that it was a non-specific PCR product. This indicated that the plants are homozygous mutants indeed. The expression levels of the *AtKu70* and *AtLig4* gene were also checked in the wild-type and the two *Atlig6* mutants as a reference. The *Atlig6* mutants had similar expression levels for the *AtKu70* and *AtLig4* genes as the wild-type.

The two *Atlig6* mutants were crossed with the *Atlig4* mutant (chapter 2) and homozygous *Atlig4lig6-1* and *Atlig4lig6-2* mutants were obtained. No obvious differences in growth were observed in these mutants compared with the wild-type under normal growth conditions.

Expression profiling of *AtLig6*

We compared the expression of the three ligase genes of *A. thaliana*, i.e. *AtLig1*, *AtLig4* and *AtLig6*, using the genevestigator analysis tool (Figure 3). The expression levels of the three ligase genes were quite low in all organs, except for sperm cells where the expression levels of *AtLig1* and *AtLig6* were extremely high, indicating that *AtLig6* may play a role in meiosis as does *AtLig1*. The data also revealed that the developmental expression pattern of the three ligases was broadly similar to each other and that the expression level of *AtLig6* was rather low compared with *AtLig4* and *AtLig1* during the whole plant life cycle. The highest expression of *AtLig6* occurred during seed germination and flower development. In order to find out whether the expression level of *AtLig6* responds to certain stimuli, the genevestigator data relating to the abiotic stimuli were analyzed. This showed that the expression of all the three ligase genes hardly changed upon genotoxic stresses, but it changed about 2-fold for some other abiotic stresses. For *AtLig6*, its expression was induced by high light or heat.

Sensitivity to genotoxic agents

Since *AtLig4* (chapter 2) and *AtLig1* (7) are involved in DNA repair, we hypothesized that the plant specific *AtLig6* would also be involved in this process. Therefore, we determined whether the *Atlig6* mutants were more sensitive to DNA damaging agents than the wild-type. The wild-type and the *Atlig6-1*, *Atlig6-2*, *Atlig4lig6-1* and *Atlig4lig6-2* mutants were treated with the DNA-damaging agent bleomycin. After 3 weeks, the *Atlig4*, *Atlig4lig6-1* and *Atlig4lig6-2* mutants turned out to be hypersensitive to bleomycin compared with the

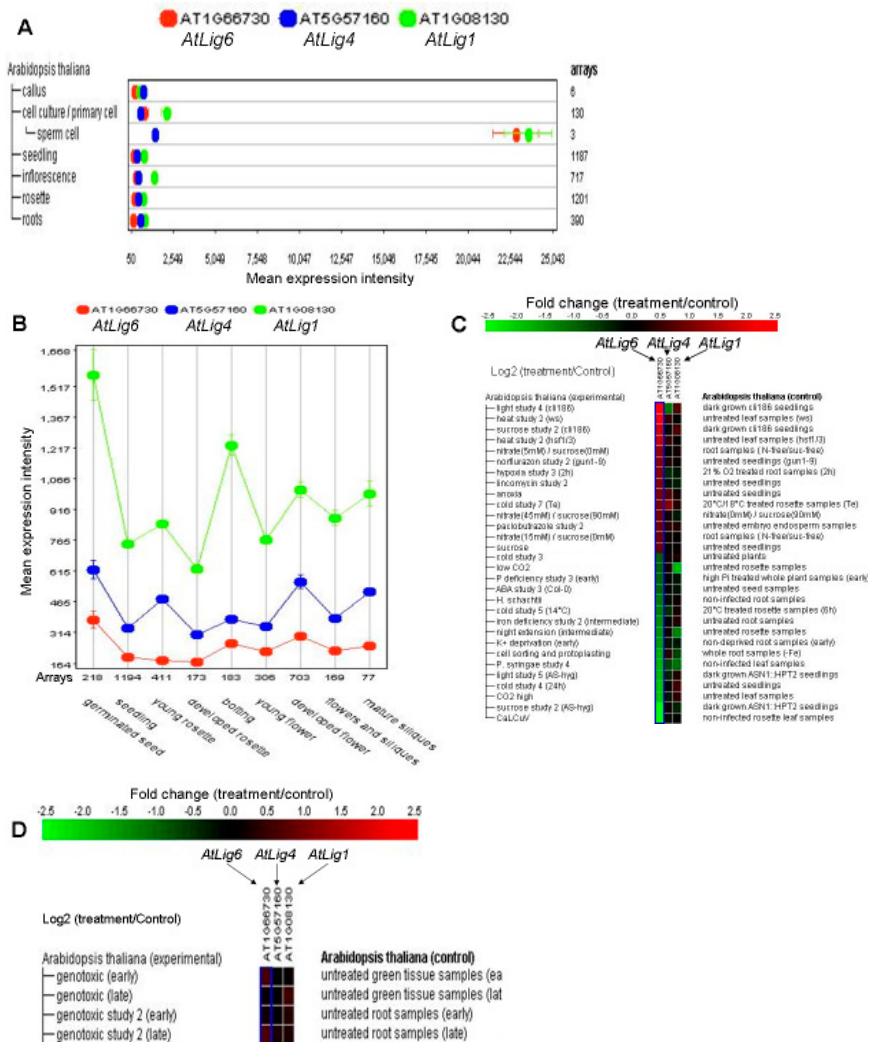


Figure 3. Expression pattern of the three ligase genes *AtLig6*, *AtLig4* and *AtLig1* in Arabidopsis. Data were retrieved using Genevestigator analysis tools of anatomy (A), development (B) and stimulus (C, D). Scatterplot outputs of tissue-specific (A) and developmental (B) expression patterns are shown. *AtLig6* is indicated as red dot, *AtLig4* as blue dot, and *AtLig1* as green dot. The numbers of arrays are also shown. Heat maps (C, D) show the expression levels of *AtLig6*, *AtLig4* and *AtLig1* in response to various stimuli. Columns represent probe sets, and rows represent stimuli. The expression was normalized for colour from red through black to green. Red, black and green indicate relatively higher, the same and lower expression levels, respectively. The stimuli which could change the expression of *AtLig6* are shown in (C). The expression levels in response to genotoxic treatments for the three ligase genes are shown in (D).

wild-type, whereas the two *AtLig6* single mutants seemed to tolerate the stress of bleomycin as well as the wild-type (Figure 4). As described in chapter 2, *AtLig4* is the main DNA ligase involved in the NHEJ pathway. Therefore, the mutants, in which *AtLig4* is deficient, were expected to be sensitive to DNA damaging agents. The double mutants were not more

sensitive to bleomycin than the *Atlig4* single mutant, suggesting that if any, *AtLig6* has only a minor role in the process of DSB DNA repair.

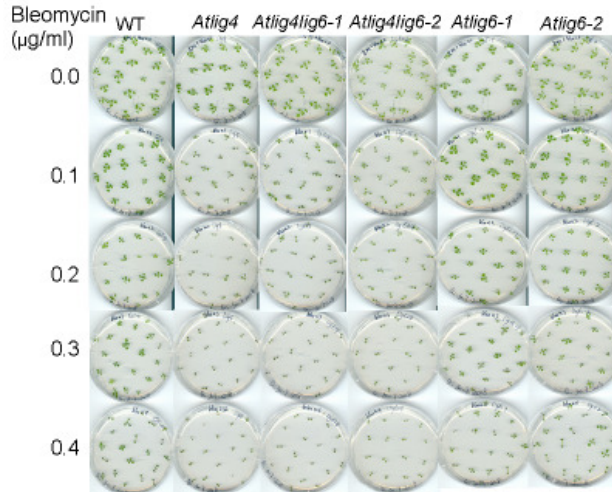


Figure 4. Response of *Atlig6* mutants to the DNA-damaging agent bleomycin. The wild-type, the *Atlig6-1*, *Atlig6-2*, *Atlig4lig6-1* and *Atlig4lig6-2* mutants were treated with different concentrations of bleomycin on the solid ½ MS media and were scored 2 weeks after germination.

T-DNA integration and gene targeting in the *Atlig6* mutants

Our previous study showed that host proteins involved in DNA repair were involved in *Agrobacterium* T-DNA integration via floral dip transformation, and thus *AtLig6* could also be involved in that process. Our group has found that *Lig4* is essential for T-DNA integration in yeast and fungi, but not in plants. It might be that this was due to redundancy of *Lig4* with the plant-specific *Lig6*. Therefore, we determined T-DNA transformation in the *Atlig6* and *Atlig4Atlig6* double mutants, and also established whether the gene targeting efficiency was increased in these mutants. The floral dip transformation experiments were done at least 3 times independently for each mutant. The transformation frequencies of the

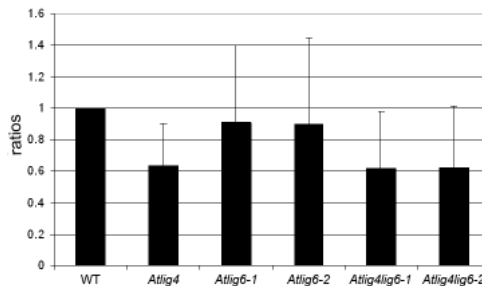


Figure 5. Transformation frequencies using the floral dip assay. One gram of seeds from the wild-type and the NHEJ mutants obtained after floral dip transformations were selected on ppt. The number of ppt resistant seedlings was scored 2 weeks after germination. The transformation frequency is presented as the percentage of ppt resistant seedlings compared with the wild-type, and the value of the wild-type is set on 1.

Atlig6-1 and *Atlig6-2* mutants were at the same level as that of the wild-type (Figure 5). The transformation frequency of the *Atlig4* mutant was reduced mildly compared with the wild-type (chapter 2), and the transformation frequencies of the *Atlig4lig6-1* and *Atlig4lig6-2* mutants were similar to that of the *Atlig4* mutant (Figure 5). This indicated that AtLig6, like AtLig4 (chapter 2), was not required for efficient *Agrobacterium* T-DNA integration in plant germline cells.

Subsequently, the frequency of gene targeting was determined with the same mutants. About 1 butafenacil-resistant plant in 1000 transformants of the wild-type (chapter 2) were obtained. One butafenacil-resistant plant in 744 transformants of the *Atlig4lig6-1* mutant, and none butafenacil-resistant plants were found in 1000 or more transformants for the *Atlig4*, *Atlig6-1*, *Atlig6-2* and *Atlig4lig6-2* mutants (Table 2). The butafenacil-resistant plants

Table 2. The numbers of different events found in gene targeting experiments.

Plant lines	Transformants tested	Butafenacil resistant	ppt resistant	Transformation frequency
WT	2600	2	1	1
<i>Atlig4</i>	1537	0	-	0.64
<i>Atlig6-1</i>	1820	0	-	0.91
<i>Atlig6-2</i>	1998	0	-	0.90
<i>Atlig4lig6-1</i>	744	1	1	0.62
<i>Atlig4lig6-2</i>	996	0	-	0.63

The transformation frequency was shown as the ratios compared with the wild-type and the value of the wild-type was set on 1.

were analyzed by PCR to determine whether they were indeed the result of gene targeting (GT) events. The PCR products with the primers PPO-PA and PPO-4 were sensitive to KpnI, indicating that the 5' end of *PPO* is replaced via HR (Figure 6A). This was confirmed by KpnI digestion for the PCR products of nested PCR reactions with the primer PPO-1 and PPO-4 (Figure 6B). In order to test whether the 3' end of *PPO* is also replaced via HR, PCR products with the primers PPO-PA and Sp319 were cleaved by KpnI. The butafenacil-resistant plants of the wild-type represented true gene targeting (TGT) plants (chapter 2). However, the PCR products from the butafenacil resistant *Atlig4lig6-1* plant were resistant to KpnI, indicating it represented an ectopic gene targeting (EGT) event. The number of transformants tested of the double mutants was too low to make a solid conclusion, but it seemed that the absence of both AtLig4 and AtLig6 did not increase the gene targeting efficiency.

Discussion

Under both normal growth conditions and under genotoxic stress, no specific phenotype

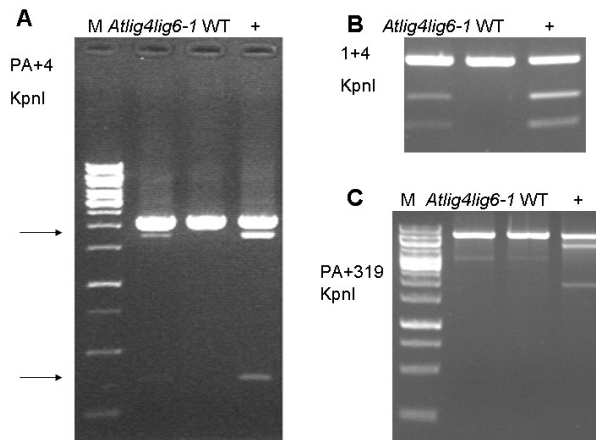


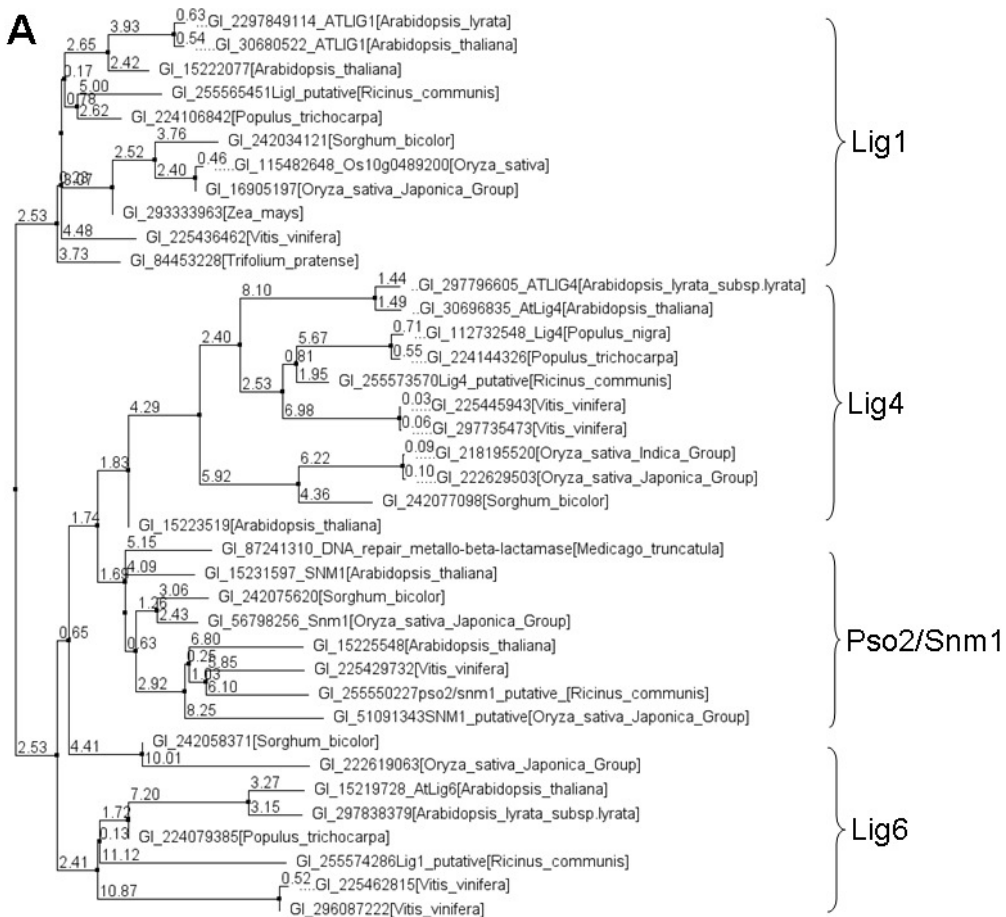
Figure 6. Gene targeting analysis for the butafanecil-resistant plant of *Atlig4lig6-1* mutant. PCR products were amplified with primers on *PPO* gene or non-coding regions (A: PPO-PA+4, B: PPO-1+PPO-4, C: PPO-PA+Sp319) followed by digesting with KpnI. The two lanes for *Atlig4lig6-1* in (A) are PCR products from the same butafanecil-resistant *Atlig4lig6-1* plant. WT: wild-type. +: positive control for true gene targeting event.

was observed in the *Atlig6* mutants. Since all DNA ligases are closely related, it could be that their function is redundant. The expression level of *AtLig6* is much lower than that of *AtLig1* and *AtLig4*, and therefore the function of *AtLig6* may be overshadowed when *AtLig1* and *AtLig4* are present. The *AtLig6* gene is most highly expressed during seed germination and flower development. Waterworth *et al.* (17) reported that *AtLig6* plays a role in seed germination and is a determinant of seed quality and longevity, probably by functioning in DNA repair at the earliest stages of seed germination to repair the DNA damage that had accumulated during seed storage. This suggests that *AtLig6* primarily functions only during certain specific stages of plant development. During the remaining parts of the life cycle, the other two DNA ligases seem to play the major role in DNA repair. The expression analysis also revealed that the expression of *AtLig6* was influenced by light and was induced by heat. Light and heat stresses can also induce DNA damage and *AtLig6* could specifically be recruited for DNA repair under such aversive growth conditions.

The absence of *AtLig6* did not disturb T-DNA integration neither in the background of the wild-type nor in that of the *Atlig4* mutant, indicating that *AtLig6*, like *AtLig4* (chapter 2), had no or only a minor role in T-DNA integration in germline cells. In mammalian cells, a low level of *Lig4* and *Lig3* is sufficient for efficient NHEJ (26). This could also be the case in plant cells, and all the three known ligases may collaboratively be involved in T-DNA integration. When both *AtLig4* and *AtLig6* were absent, T-DNAs were still integrated efficiently in the plant genome, suggesting that there must still be another ligase to take over that function. The frequency of gene targeting was not increased in the double *Atlig4lig6-1* mutant.

A candidate may be *AtLig1*, which is mainly involved in the ligation of Okazaki fragments during DNA replication, but may in addition also act in DNA repair (6). It is

also possible that another, so far unknown DNA ligase takes over from AtLig4 and AtLig6 in their absence. In order to find indications about the presence of other putative ligases in the *A. thaliana* genome, we retrieved sequences that shared homology and determined their relationship to the DNA ligases 1, 4 and 6 by construction of a phylogenetic tree using the neighbour joining algorithm (Figure 7A). The adenylation domain and the oligonucleotide/ oligosaccharide binding-fold (OBF) domain comprise a common catalytic core unit for the ATP-dependent DNA ligase family. The conserved domains of DNA ligases from plants, mammals and yeast are shown in Figure 8. Different DNA ligases contain other specific domains. The Lig4 ligases of higher eukaryotes contain a breast cancer suppressor protein, carboxy-terminal (BRCT) domain, while the Lig3 ligases contain a poly(ADP-ribose) polymerase and DNA ligase Zn-finger (ZF-Parp) region. The Lig6 ligases contain a lactamase (Lac) domain and a DNA repair metallo-beta-lactamase (DRMBL) domain, like the proteins of the Pso/Snm1 family. The phylogenetic analysis resulted in a tree composed of four major clades (the Lig1, Lig4, Pso/Snm1 and Lig6 clades). Two putative DNA ligase-like proteins (GI_15222077 and GI_15223519) were found in Arabidopsis. GI_15222077 is



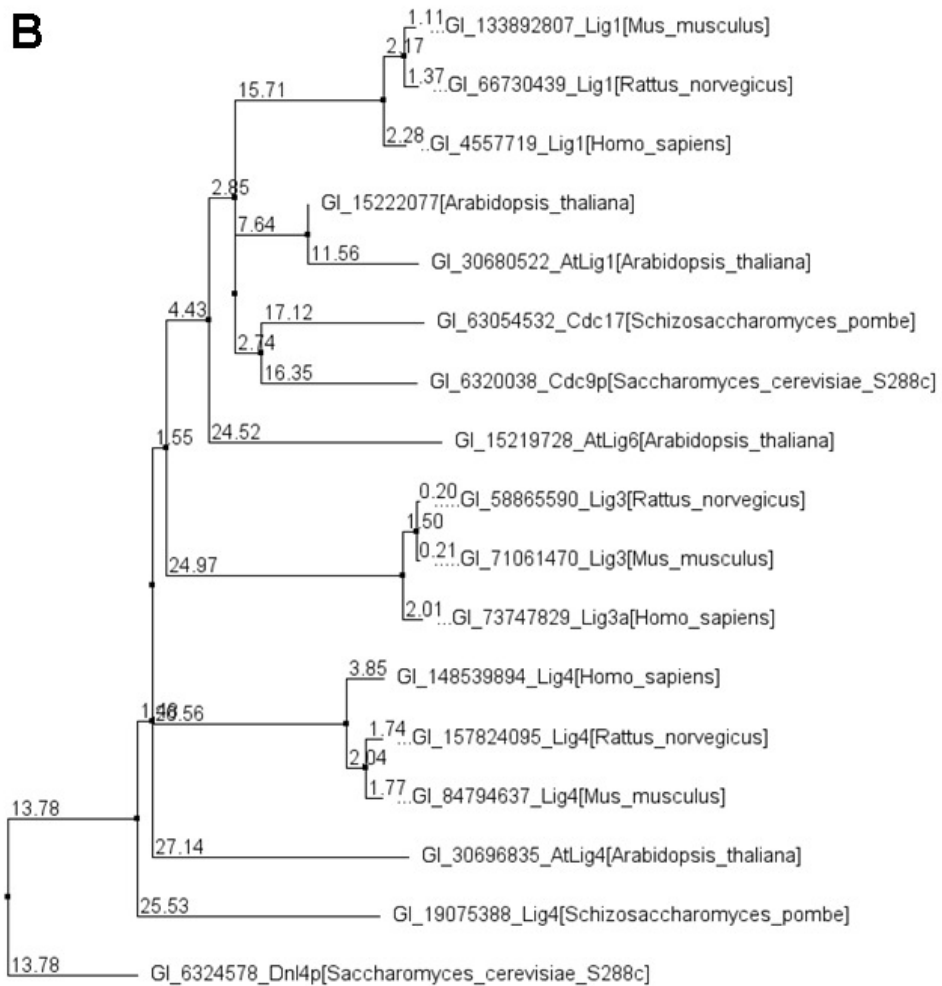
B

Figure 7. Phylogenetic analysis of DNA ligases.

The trees were calculated using Jalview software and the distances were also shown. The algorithm is the neighbour joining using % identity. (A) Tree of the plant DNA ligases. (B) Tree of the conserved adenylation and OBF domains of the DNA ligases in plants, mammals and yeast.

in the clade of Lig1 and close to AtLig1. GI_15223519 is in the clade of Pso/Snm1 (16), suggesting it could be involved in DNA crosslink repair. The GI_15222077 protein only contains the conserved domains of the adenylation domain and the OBF domain, while the GI_15223519 protein only contains the conserved lactamase and DRMBL domains; therefore, it is unlikely that this protein is a genuine ligase protein. Another phylogenetic tree was built with the adenylation domain and the OBF domain of DNA ligases from plants, mammals and yeast (Figure 7B). The Lig1 and Lig4 genes are present in lower eukaryotes indicating they are the oldest DNA ligases in nature. Until now Lig3 has been identified only in animals, and Lig6 has been found only in plants. The phylogenetic tree suggests that Lig3 and Lig6 were derived from Lig1 later in evolution in animals and plants, respectively.

Since the homozygous *Atlig1* mutant is lethal, the GI_15222077 protein is not redundant with AtLig1. The expression level of the GI_15222077 protein is even lower than AtLig6 under normal conditions according to the data of the Genevestigator data base (data not shown). But it is highly increased under hypoxia, drought, UV and light stress, and therefore may be a very good candidate for having a role in B-NHEJ.

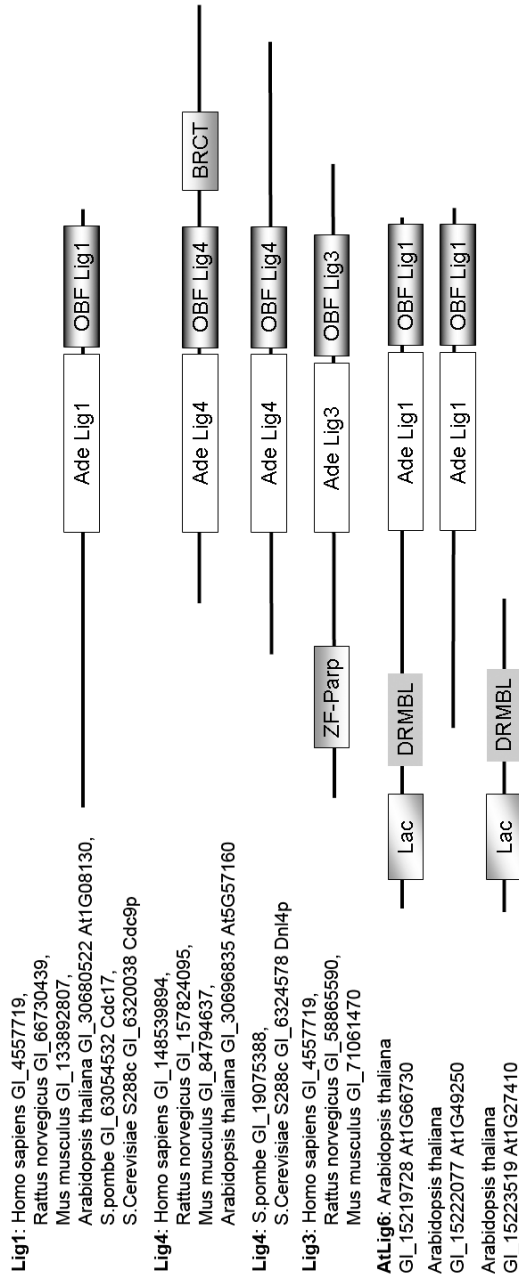


Figure 8. Conserved domains of the ATP-dependent DNA ligases. The conserved domains of the DNA ligases are shown as colored rectangles and the remaining parts are shown as black lines. Ade represents the adenylation domain. OBF represents the oligonucleotide/oligosaccharide binding-fold domain. BRCT represents the breast cancer suppressor protein, carboxy-terminal domain. ZF-Parp represents the poly(ADP-ribose) polymerase and DNA ligase Zn-finger domain. Lac represents the Lactamase domain. DRMBL represents the DNA repair metallo-beta-lactamase domain.

Acknowledgements

Amke den Dulk-Ras for the technical assistance with gene targeting project and Tiia Husso for the technical assistance with the characterization of mutants. This work was supported by the Chinese Scholarship Council (CSC) (QJ).

Reference List

1. Timson,D.J., Singleton,M.R. and Wigley,D.B. (2000) DNA ligases in the repair and replication of DNA. *Mutat. Res.*, **460**, 301-318.
2. Pascal,J.M. (2008) DNA and RNA ligases: structural variations and shared mechanisms. *Curr. Opin. Struct. Biol.*, **18**, 96-105.
3. Shuman,S. (2009) DNA ligases: progress and prospects. *J. Biol. Chem.*, **284**, 17365-17369.
4. Martin,I.V. and MacNeill,S.A. (2002) ATP-dependent DNA ligases. *Genome Biol.*, **3**, REVIEWS3005.
5. Bonatto,D., Brendel,M. and Henriques,J.A.P. (2005) A new group of plant-specific ATP-dependent DNA ligases identified by protein phylogeny, hydrophobic cluster analysis and 3-dimensional modelling. *Functional Plant Biology*, **32**, 161-174.
6. Ellenberger,T. and Tomkinson,A.E. (2008) Eukaryotic DNA ligases: structural and functional insights. *Annu. Rev. Biochem.*, **77**, 313-338.
7. Waterworth,W.M., Kozak,J., Provost,C.M., Bray,C.M., Angelis,K.J. and West,C.E. (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.
8. van Attikum,H., Bundock,P., Overmeer,R.M., Lee,L.Y., Gelvin,S.B. and Hooykaas,P.J. (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.
9. Critchlow,S.E., Bowater,R.P. and Jackson,S.P. (1997) Mammalian DNA double-strand break repair protein XRCC4 interacts with DNA ligase IV. *Curr. Biol.*, **7**, 588-598.
10. Mackey,Z.B., Ramos,W., Levin,D.S., Walter,C.A., McCarrey,J.R. and Tomkinson,A.E. (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.
11. Cappelli,E., Taylor,R., Cevasco,M., Abbondandolo,A., Caldecott,K. and Frosina,G. (1997) Involvement of XRCC1 and DNA ligase III gene products in DNA base excision repair. *J. Biol. Chem.*, **272**, 23970-23975.
12. Wang,H., Rosidi,B., Perrault,R., Wang,M., Zhang,L., Windhofer,F. and Iliakis,G. (2005) DNA ligase III as a candidate component of backup pathways of nonhomologous end joining. *Cancer Res.*, **65**, 4020-4030.
13. Audebert,M., Salles,B. and Calsou,P. (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.
14. McVey,M. and Lee,S.E. (2008) MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet.*, **24**, 529-538.
15. Liang,L., Deng,L., Nguyen,S.C., Zhao,X., Maulion,C.D., Shao,C. and Tischfield,J.A. (2008) Human DNA ligases I and III, but not ligase IV, are required for microhomology-mediated end joining of DNA double-strand breaks. *Nucleic Acids Res.*, **36**, 3297-3310.
16. Bonatto,D., Revers,L.F., Brendel,M. and Henriques,J.A. (2005) The eukaryotic Pso2/Snm1/Artemis proteins and their function as genomic and cellular caretakers. *Braz. J. Med. Biol. Res.*, **38**, 321-334.
17. Waterworth,W.M., Masnavi,G., Bhardwaj,R.M., Jiang,Q., Bray,C.M. and West,C.E.

- (2010) A plant DNA ligase is an important determinant of seed longevity. *Plant J.*, **63**, 848-860.
18. Alonso,J.M., Stepanova,A.N., Leisse,T.J., Kim,C.J., Chen,H., Shinn,P., Stevenson,D.K., Zimmerman,J., Barajas,P., Cheuk,R. *et al.* (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science*, **301**, 653-657.
 19. de Pater,S., Caspers,M., Kottenhagen,M., Meima,H., ter Stege,R. and de Vetten,N. (2006) Manipulation of starch granule size distribution in potato tubers by modulation of plastid division. *Plant Biotechnol. J.*, **4**, 123-134.
 20. Weijers,D., Franke-van,D.M., Vencken,R.J., Quint,A., Hooykaas,P. and Offringa,R. (2001) An *Arabidopsis* Minute-like phenotype caused by a semi-dominant mutation in a RIBOSOMAL PROTEIN S5 gene. *Development*, **128**, 4289-4299.
 21. Murashige,T. and Skoog,F. (1962) A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiologia Plantarum*, **15**, 473-497.
 22. Clough,S.J. and Bent,A.F. (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.*, **16**, 735-743.
 23. Hanin,M., Volrath,S., Bogucki,A., Briker,M., Ward,E. and Paszkowski,J. (2001) Gene targeting in *Arabidopsis*. *Plant J.*, **28**, 671-677.
 24. Masson,J. and Paszkowski,J. (1992) The Culture Response of *Arabidopsis-Thaliana* Protoplasts Is Determined by the Growth-Conditions of Donor Plants. *Plant Journal*, **2**, 829-833.
 25. van Attikum H., Bundock,P. and Hooykaas,P.J. (2001) Non-homologous end-joining proteins are required for *Agrobacterium* T-DNA integration. *EMBO J.*, **20**, 6550-6558.
 26. Windhofer,F., Wu,W. and Iliakis,G. (2007) Low levels of DNA ligases III and IV sufficient for effective NHEJ. *J. Cell Physiol*, **213**, 475-483.