# Highly efficient CRISPR-mediated homologous recombination via NHEJ deficiency rather than HDR factors overexpression in *Populus*

Ali Movahedi[1§*], Hui Wei[1§], Zhong-Hua Chen[2], Weibo Sun[1], Jiaxin Zhang[3], Dawei Li[1], Liming Yang[1*], and Qiang Zhuge[1*]

[1] College of Biology and the Environment, Co-Innovation Center for Sustainable Forestry in Southern China, Key Laboratory of Forest Genetics & Biotechnology, Ministry of Education, Nanjing Forestry University, Nanjing 210037, China
[2] School of Science, Hawkesbury Institute for the Environment, Western Sydney University, Penrith, NSW 2751, Australia
[3] School of Food Science and Pharmaceutical Engineering, Nanjing Normal University, Nanjing 210046, China

[*]Correspondence should be addressed to Qiang Zhuge, Ali Movahedi, and Liming Yang: College of Biology and the Environment, Co-Innovation Center for Sustainable Forestry in Southern China, Key Laboratory of Forest Genetics & Biotechnology, Ministry of Education, Nanjing Forestry University, Nanjing 210037, China. E-mail: qzhuge@njfu.edu.cn; ali_movahedi@njfu.edu.cn; yangliming@njfu.edu.cn; Fax: +86 25 85428701

§ These authors are equally as the first author

**Running title: Highly efficient homologous recombination via *XRCC4* deficiency in poplar**

**Abstract**

- Efficient homology-directed DNA repair (HDR) is a vital difficulty confronting researchers to replace the target genome's desired fragment. In plants, scientists have performed meticulous investigations on herbal, crops, and citrus trees using HDR effector proteins, CtIP and MRE11, to obtain double-stranded breaks (DSBs) more precisely. Although HDR efficiency in plants previously has been reported, no record has been declared about HDR efficiency in woody perennial *Populus*.
- Here, we hypothesized that inhibition of non-homologous recombination cofactors XRCC4 and enhancing the HDR pathway activities enable us to improve the HDR efficiency. In this study, the *BleoR* gene was used to integrate into the interested site and generated transformants against Zeocin antibiotics. We designed plasmids, including different fusions of HDR proteins and, together with the *XRCC4* target. Furthermore, we showed that TaqMan real-time PCR could be a powerful tool to verify HDR efficiency
- We confirmed that both applying HDR proteins and *XRCC4* deficiency simultaneously could improve HDR efficiency, which showed about fiftyfold more, and decremented polymorphisms about sixfold less than no affecting HDR and NHEJ pathways. We developed a new recombinant poplar genome to generate stable lines resistant to the Zeocin antibiotic.

**Keywords:** CRISPR; XRCC4; Homologous recombination; BleoR; *Populus trichocarpa*

## Introduction

The clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated (Cas) protein system has been widely developed for DNA sequence manipulation in recent years. The efficient molecular scissors of the Cas9 system generate double-strand breaks (DSB), which consequently initiates endogenous repair machinery typically via either the error-prone non-homologous end joining (NHEJ) pathway or the donor-dependent homology-directed repair (HDR) pathway (Symington and Gautier, 2011). NHEJ repair can randomly lead to small deletions or insertions (indels) at the DSB area, which then shifts the reading frame, interferes with the gene function, and has shown a predominant pathway and occurs among the cell cycle widely (Panier and Boulton, 2014). The HDR is rare but powerful for sequence insertion and replacement if a homologous DNA substrate is available (Puchta, 2005). Up to date, many studies have been carried out to improve the genetic modification of crops by HDR. For instance, one study has been carried out to increase *ARGOS8* expression by replacing the GOS2 promoter via HDR to drive *ARGOS8* expression (Shi et al., 2017). Another report showed an enhancement in the efficiency of the insertion of the 35S promoter in upstream of the *ANT1* gene in tomato (Cermak et al., 2015). Many publications also report the successful generation of null mutations in woody species via the NHEJ pathway since the first implementation in poplar (Bewg et al., 2018; Zhou et al., 2015). However, precise gene targeting and replacement were only reported in model plants such as *Arabidopsis* (Schiml et al., 2014) and rice (Li et al., 2016). Still, no report shows highly efficient HDR happened in woody perennials.

Many attempts have been made to improve gene targeting efficacy and accuracy (Liu et al., 2019). Lu et al. (2020) recently used chemically modified donor DNA to promote heritable sequence insertion and replacement via NHEJ and HDR repair pathways. Both-end phosphorothioate-linkage and 5′-phosphorylation modification facilitated NHEJ insertion with lengths up to 2,049bp. Furthermore, a tandem repeat-HDRs system was optimized for base substitutions and in-locus tags (Lu, Tian et al., 2020). Though another biotechnology breakthrough on prime editor presented the success bases of insertion in rice (Lin et al., 2020; Tang et al., 2020), HDR-mediated precise gene targeting is a promising method for a long time.

97 Manipulation of the DSB repair pathway to improve HDR frequency is the most common
98 method to enhance precise gene knock-in. One of the main difficulties in HDR efficiency is the
99 limited delivery of donor DNA templates (DDTs) into the cell nucleus. Previous data indicated
100 that it is necessary to increase the number of cells containing DDTs at S/G2 cell division phases
101 to increase HDR efficiency (Yang et al., 2016). Several traditional strategies were applied to
102 enrich the DDTs availability, such as offering sufficient DDTs via the particle bombardment
103 replicons (Gil-Humanes et al., 2017; Wang et al., 2017), protoplasts (Svitashev et al., 2016), the
104 geminiviral-based replication (Butler et al., 2016; Dahan-Meir et al., 2018; van Vu et al., 2020;
105 Wang et al., 2017), and RNA transcription (Li et al., 2019b), but this still is one of the significant
106 problems across woody plants genome. In this case and regarding the high and reliable stability
107 of genes transmitted by *Agrobacterium* and widely use to transduce genes into woody plant
108 cells (Movahedi et al., 2014; Movahedi et al., 2015), there few reports on enhancing the
109 Agrobacteria method delivery to increase the efficiency of transferring DDT and, consequently,
110 the recovery of DSBs as HDR in woody trees (Ali et al., 2020; An et al., 2020). Above all,
111 manipulating the effectors involved in either NHEJ or HDR pathway was more straightforward.
112 The positive effects of recombinant homologous factors and their impact on enhancing HDR
113 efficiency in mammalians have already been reported. Cas9 integrates with MRE11, CtIP, and
114 Rad51, Rad52, and promotes significant HDR efficiency in human cells and decreases NHEJ
115 significantly, with at least a 2-fold increase in HDR and a 6-fold increase in HDR/NHEJ ratio (Tran
116 et al., 2019). On the other hand, inhibition of DNA ligase IV (LIG4), Ku 70, and Ku 80, which are
117 outwardly involved only in NHEJ and known as the most critical NHEJ factors, protect DSB from
118 discrediting by forming one heterodimeric complex to bind tightly and load additional repair
119 proteins such as DNA ligase IV (Friesner and Britt, 2003; Grawunder et al., 1998a; Maruyama et
120 al., 2015; Pierce et al., 2001; Tran et al., 2019) and increase the HDR efficiency up to 19-fold
121 (Tran et al., 2019). In *Arabidopsis*, mutated *Ku70* or *Lig4* enhanced the HDR-based genome
122 targeting to 5~16-fold or 3~4-fold respectively (Qi et al., 2013).
123 XRCC4 is another critical NHEJ factor that has not yet been considered for its
124 interference effect on HDR efficiency. XRCC4 is one cofactor of LIG4 to interact with KU 70 and
125 KU 80 and ligate the DSB (Grawunder et al., 1998b). To date, there is no report on combining

126 HDR factors overexpressing, specially CtIP and MRE11, and NHEJ factors suppression, specially
127 XRCC4, to promote the HDR pathway in woody plants. Programmable endonucleases affect
128 DSBs at target positions in genomic DNA and create undesired breaks outside of on-target
129 positions and create off-target mutations. Cleavage at off-target sites direct to chromosomal
130 rearrangements, including translocations, insertions, and deletions, which happen in the
131 interruption of regular gene expression and the activation of oncogenes (Li et al., 2019a). Today,
132 scientists have realized that reducing off-target may allow efficient and accurate genome
133 editing (Wu and Yin, 2019). For this why, the effect of off-targets on efficient, precise genome
134 editing and ways to reduce their impacts has already been studied (Hajiahmadi et al., 2019; Li
135 et al., 2019a; Wu and Yin, 2019).

136 In eukaryotic kingdoms, Mitogen-activated protein kinase (MAPK or MPK) signaling is
137 broadly crossed. In plants, they may direct signal pathways to resist the impacts of drought,
138 cold, salt, scratch, and mechanical damage(Chen et al., 2020). Moreover, scientists have shown
139 that MAPKs direct cellular responses against heat shock, osmotic, and other environmental
140 stresses (Mohanta et al., 2015). Sun et al. (2015) investigated and comprised the *MAPK* genes
141 from Maize (*Zea mays*), poplar (*Populus trichocarpa*), Arabidopsis (*Arabidopsis thaliana*),
142 tomato (*Solanum lycopersicum*), and rice (*Oryza sativa*) to show the similarity of motifs,
143 genomic structures, and domains. These authors also proved that *MAPK* genes are
144 preferentially expressed in reproductive tissues and organs besides their roles against abiotic
145 stress and ABA signaling. In Arabidopsis, *MAPKK2* (*MKK2*) are transcriptional regulator genes
146 stimulated by environmental stresses such as salinity and cold to promote plant protection
147 (Teige et al., 2004). To date, and regarding the importance of the MAPK pathway, there is no
148 report on applying CRISPR genome editing through all involved genes to enhance its impact in
149 the mentioned above roles.

150 Here, we aimed to develop the CRISPR/Cas9-mediated knock-in system in poplar with a
151 haploid chromosome of 19 via NHEJ cofactor (XRCC4) suppression and HDR effectors (MRE11
152 and CtIP) overexpression. Also, we used the optimized homologous arms 400 bp to enhance the
153 HDR activities, and the concentration of DDT fragments to facilitate the DDT delivery. We
154 selected mitogen-activated protein kinases (MAPKs) as a case study. We also achieved the

155 stable transformant lines resistant against zeocin antibiotic with similar behavior to wild types
156 against salinity stresses.

157 ## Materials and Methods

158 ### Plant Transformation

159 We cultivated *P. trichocarpa* seedlings in a phytotron at 23±2°C under a 16/8 light/dark
160 time (Movahedi et al., 2015). To generate transgenic lines, we used stems from four weeks old
161 young clones and dipped them in the optimized of *Agrobacterium tumefaciens* stimulant and
162 pathogenic suspension ($OD_{600}$: 2.5, 120 min, pH ~5, Acetosyringone (As): 200 µM) (Movahedi et
163 al., 2014) for 5 min with gentle shaking and then transferred to the semi-solid woody plant
164 medium (WPM) added with 0.05 mg/L Indole-3-butyric acid (IBA), 0.006 mg/L thidiazuron (TDZ),
165 200 µM As and 0.5% (w/v) agar. Afterward, the stimulated stems were incubated in the dark at
166 23°C for two days. The assumed transformants were then co-cultivated in selection media
167 enriched with 0.1 mg/L IBA, 0.006 mg/L TDZ, 100 mg/L cefotaxime, 8 mg/L hygromycin, 50 mg/L
168 Zeocin and 0.8% (w/v) agar. Two weeks later, buds were regenerated and then sub-cultured
169 independently in media including 0.1 mg/L IBA, 0.001 mg/L TDZ, 100 mg/L cefotaxime, 8 mg/L
170 hygromycin, 50 mg/L Zeocin and 0.8% (w/v) agar. After six weeks, buds with four to six small
171 leaves were transferred to the MS media with 0.1 mg/L IBA, 200 mg/L cefotaxime, 70 mg/L
172 Zeocin, and 0.8% (w/v) agar to root. Five lines were used for each experiment independently,
173 and each line included about ten individuals.

174 ### Targets and protein detection

175 We decided to target the *MKK2* gene from *P. trichocarpa* (POPTR_0018s05420g;
176 Chromosome 18) because of its vital role in transcriptional regulation against environmental
177 stresses. Therefore, we used Uniprot (https://www.uniprot.org/) to download MKK2 protein
178 and then used the BLAST database of the National Center for Biotechnology Information (NCBI)
179 (https://blast.ncbi.nlm.nih.gov/) to download full DNA sequences and CDS. To detect targets,
180 we used Geneious Prime® 2020.1.1 to analyze *MKK2* locus and detect targets compared to the
181 whole genome of *P. trichocarpa*, which has been already downloaded from NCBI
182 (Supplementary Table 1) (Doench et al., 2014; Hsu et al., 2013). Geneious Prime® also has been
183 used to analyze the *XRCC4* (POPTR_0010s08650g, Chromosome 10) gene for knocking out. The

184 PAM motif target sequences were concerned with the exon 8 area from the *MKK2* and exon 1
185 area from the *XRCC4* genes. Furthermore, and to evaluate the effect of HDR proteins and also
186 proper function of edited *MKK2* gene in transformants, we used Uniprot to use CtIP
187 (POPTR_001G269700v3), MRE11 (POPTR_0001s41800g), BRCA1 (POPTR_0005s26150g), Rad50
188 (POPTR_0001s32760g), Rad51 (POPTR_0014s06360g), Lig4 (POPTR_0018s13870g).

189 ***MKK2* locus target oligo synthesis**

190 We designed a pair of oligos (Supplementary Table 2; *MKK2* Oligo-F and -R) flanked
191 by *Bsa*I adaptors. Synthesized oligos were then ligated into digested pRGEB31 vectors by *Bsa*I
192 restriction enzyme (Xie and Yang, 2013) to construct pgRNA (Supplementary 1a). Afterward, we
193 transferred all vectors into *E. coli* (DH5α) and propagated under normal conditions. Vectors
194 were then extracted using the plasmid midi kit (Qiagen, USA) and confirmed by sanger
195 sequencing (GenScript, Nanjing).

196 **Construction of DDT and pDDT**

197 To produce DDT (Supplementary1b), we designed five fragments, constructed and
198 ligated them, respectively (Supplementary2a). To construct fragment one, the OsU3 promoter
199 and gRNA scaffold were isolated from pRGEB31 (Supplementary Table 2, OS1-F and -R) flanked
200 by *Hind*III and *BamH*I endonucleases. To increase the amount of DDT in the cell nucleus and
201 improve HDR efficiency, we decided to use the cleavage property of Cas9 with designing two
202 special gRNA targets 1 and -2 (No on- and -off-targets through whole poplar genome) besides
203 DDT (Zhang et al., 2017) (Supplementary 1b). Thus, we designed special gRNA oligos (Sgo1-F
204 and -R) (Supplementary 2a; Supplementary Table 2, special gRNA oligo1-F and -R) as the
205 described details (Xie and Yang, 2013) to form special gRNA target1 (Sgt1) and to ligate into the
206 fragment one. To construct fragment two, we isolated 5' homology arm (400 bp) sequences
207 from *P. trichocarpa* genomic DNA (Supplementary Table 2, 5' Ho-F-1 and -R-1). Afterward,
208 regular PCR was carried out using primers with the extensions of *BamH*I-special target1 (St1)
209 and 39 bp from complemented 5' of fragment 3 (Supplementary Table 2, 5' Ho-F-2 and -R-2)
210 (Supplementary 2a) to achieve component two.

211 To construct fragment three, we isolated the BleoR CDS from the PCR®-XL-Topo® vector
212 (Supplementary Table 2, BleoR-1092F and -2276R). Then, the overlap-PCR was performed

213 (Supplementary Table 2, BP1,2,3-F and -R) using isolated BleoR CDS as the template to add the
214 remained sequences from exon 8 and exon 9 to the 5' region of BleoR CDS and also 6XHis and
215 PolyA tail to the 3' area of BleoR CDS (Supplementary 2a). We isolated a 3' homology arm (400
216 bp) (Supplementary Table 2, 3' Ho-F-1 and -R-1) from *P. trichocarpa* genomic DNA to assemble
217 fragment four. Then, we performed PCR to extend 3' homology arm with 30 bp Poly-T and *Nco*I-
218 special target2 (St2) sequences (Supplementary Table 2, Ho-F-2 and -R-2) (Supplementary 2a).
219 Finally, we performed standard PCR to isolate the OsU3 promoter and gRNA scaffold from
220 pRGEB31 (Supplementary Table 2, Os2-F and Os2-R). Moreover, we designed special gRNA
221 oligos (Sgo2-F and -R) (Supplementary 2a; Supplementary Table 2, special gRNA oligo2-F and -R)
222 again as the described details (Xie and Yang, 2013) to form special gRNA target2 (Sgt2) and to
223 ligate into the fragment five.

224 To construct pDDT, we ligated fragments three and two using PCR (Supplementary 2b).
225 For this, we designed a 39 bp overhang on fragment two that was complementary to the end of
226 fragment three to form preliminary DDT (Supplementary 2b). In this PCR, we prepared a PCR
227 reaction with 500 ng of each component. We used everything in PCR reaction except primers
228 and then denatured fragments at 95 degrees for 5 minutes and allowed two annealing and
229 extension cycles. We allowed PCR products to anneal at 68 degrees to avoid nonspecific
230 hybridization amongst the long PCR products for 30 seconds and then extend for one minute at
231 74 degrees to have a double-stranded outcome. Then we added the primers to the distal ends
232 of fragments two and three and performed one standard PCR. We purified PCR products and
233 ligated them into the pEASY vector to sequence and confirm. Then we ligated the preliminary
234 DDT product to fragment four as described before and formed secondary DDT products
235 (Supplementary 2b). After sequencing and confirmation, we used the restriction cloning
236 technique to ligate secondary DDT products to fragments one and four (Supplementary 2b) to
237 achieve DDT products. Briefly, we incubated a reaction including 50 ng of each digested
238 fragments, 10x T4 DNA ligase buffer 0.5 ul, T4 DNA ligase (NEB) 1 ul, and H2O to 5 ul at 25
239 degrees for 4 hours and transferred into *E. Coli* DH5α competent cells for sequencing and
240 confirmation. Subsequently, we used the restriction cloning technique to merge the DDT
241 product and pRGEB31 vector and form the pDDT vector (Supplementary 2b).

**Synthesis of pgCtIP and pgMR**

To design a fused CtIP and Cas9 cassette, we isolated the CaMV35S promoter, 3xFLAG, and Cas9 CDS from pRGEB31 (Supplementary 3a) using designed primers (Supplementary Table 2). In the next step, we obtained CtIP CDS using RT-PCR from the *Populus trichocarpa* genome (Supplementary 3a; Supplementary Table 2, CtIP-F and -R). The 3'UTR and PolyA fragments were isolated from the pCAG-T3-hCAS-pA plasmid (Supplementary 3a; Supplementary Table 2, PolyA-F and -R). To complete pgCtIP, we ligated CaMV35S and 3xFLAG fragments using restriction cloning and formed backbone 1 (Supplementary 4a). The isolated Cas9 and the obtained CtIP CDS were also ligated, applying restriction cloning to form the backbone 2 (Supplementary 4a). The backbones 1 and 2 were then ligated using *Hind*III restriction cloning to form backbone 3 (Supplementary 4a). In the next step, the resulted backbone 3 was ligated to the assembled 3'UTR-PolyA using *Stu*I restriction cloning to form the CtIP cassette (Supplementary 4a; Supplementary 5a). We used *Sda*I and *Pme*I restriction enzymes to restrict the cloning of the CtIP cassette and pRGEB31 and achieve the pgCtIP plasmid (Supplementary 4a; Supplementary 5a).

To construct a fusion of MRE11 and Cas9, we isolated CaMV35 promoter, 3xFLAG, Cas9, 3'UTR, and PolyA as same the previous steps (Supplementary 3b; Supplementary Table 2). The MRE11 CDS was obtained recruiting extracted total RNA from *Populus trichocarpa* genome and RT-PCR as mentioned above (Supplementary 3b; Supplementary Table 2, MRE-F and R). To complete pgMR, we ligated the isolated CaMV35S and 3xFLAG fragments concerning *Xho*I endonuclease to form backbone 1 (Supplementary 4b). On the other hand, we constructed backbone 2 using the isolated Cas9 and 3'UTR-PolyA fragments (Supplementary 4b). The backbone 1, backbone 2, and MRE11 CDS product were then merged concerning *Not*I and *Nde*I restriction cloning to form MR cassette (Supplementary 4b; Supplementary 5b). Afterward, we used restriction cloning with *Sda*I and *Pme*I to construct pgMR plasmid (Supplementary 4b; Supplementary 5b).

**Synthesis of pgCtMR and pggCtMR**

To construct the CtMR cassette, we prepared all the required fragments, as mentioned above (Supplementary 3c). Afterward, we merged the CaMV35S and 3xFLAG components

271 using *Xho*I restriction cloning to form backbone 1 (Supplementary 6a). We then ligated
272 backbone 1 and the already obtained MRE11 CDS product (Supplementary Table 2, MRE-F and -
273 R) using *Not*I restriction cloning to form backbone 2 (Supplementary 6a). On the other hand,
274 the isolated Cas9 and the obtained RT-PCR product CtIP CDS were ligated using *BamH*I
275 restriction cloning to form backbone 3 (Supplementary 6a). We then used backbone 3 and
276 isolated 3'UTR-PolyA fragment to form backbone 4 (Supplementary 6a). Eventually, we cloned
277 backbones 2 and 4 to construct the CtMR cassette (Supplementary 6a; Supplementary 5c) and
278 thereupon implemented *Sda*I and *Pme*I restriction cloning to ligate CtMR cassettes into
279 pRGEB31, forming pgCtMR plasmid (Supplementary 6a; Supplementary 5c). To target
280 the *XRCC4* gene and *MKK2* simultaneously, we designed one cassette, including both *XRCC4,* by
281 adding one CRISPR site (Located on 5' region of target CDS) to mutate *XRCC4* (Non-off-target
282 site on whole poplar genome; Activity score: 0.415; Specificity score: 100%) (Doench et al.,
283 2014; Hsu et al., 2013), and *MKK2* gRNAs. For this purpose, we used primers (Supplementary
284 Table 2, XR-Cass1-F and -R) to isolate the OsU3 promoter and gRNA scaffold from the pRGEB31
285 vector and then used *MKK2* designed oligos (Supplementary Table 2, *MKK2* Oligo-F and -R) to
286 ligate *MKK2* target duplex (Supplementary 3d). Besides, we used primers (Supplementary Table
287 2; XR-Cass2-F and -R) to isolate the OsU3 promoter and gRNA scaffold again. In this process, we
288 applied *XRCC4* designed oligos (Supplementary Table 2; XRCC4-Oligo1 and -2) to ligate the
289 *XRCC4* target duplex (Supplementary 3d). The achieved fragments were then cloned using *Kas*I
290 restriction cloning to form *XRCC4*-Cassette (Backbone 1) (Supplementary 6b; Supplementary
291 5d). Afterward, the *XRCC4*-Cassette was cloned into pRGEB31 using *Hind*III and *Sda*I restriction
292 cloning to form backbone 2 (Supplementary 6b). Finally, we used *Sda*I and *Pme*I restriction
293 cloning to clone the CtMR cassette into the backbone 2, forming pggCtMR plasmid
294 (Supplementary 6b; Supplementary 5d). We performed PCR, cloning into pEASY T3 vector, and
295 DNA sequencing in all the above processes for confirming the right ligation.

296 **RT-PCR, DNA sequencing, Southern blotting, and Western blotting**

297 Total RNA (100 ng/ml) was extracted from young leaves of five weeks grown buds on
298 Zeocin with TRIzol. We then carried out reverse transcription using total RNA and oligo-dT
299 primers to synthesize the first cDNA strand (PrimeScript One-Step RT-PCR Kit Ver.2, Takara

300 Biotechnology, Dalian, China) according to the manufacturer's instructions. Afterward, we
301 designed two RT-PCR for both investigations of right *MKK2* transcription and right happening
302 HDR. The first RT-PCR was intended to isolate a 920 bp of *MKK2* CDS (Supplementary Table 2,
303 RT-F and R), while the forward primer was designed from 5' region of exon 9 (15 bp) and 3'
304 region of exon 8 (15 bp). This RT-PCR purpose was to show the precise attaching of exon 8 and
305 9 to direct the transcription of *MKK2* correctly. The second RT-PCR was performed to isolate a
306 413 bp of recombinant CDS (Supplementary Table 2, RT-F-107 and RT-R-519). The forward
307 primer was designed from *BleoR*, and the reverse primer was designed from exon 7 of *MKK2*.
308 The purpose of this RT-PCR was to show the explicit HDR happening through our experiments
309 via transcription of single mRNA from *MKK2* and BleoR.

310 Genomic DNA was extracted from young leaves of five weeks grown buds on Zeocin,
311 applying the DNeasy Plant Mini Kit (Qiagen, USA). The quality of the extracted genomic DNA
312 (250–350 ng/µl) was determined by a BioDrop spectrophotometer (UK). To DNA sequencing,
313 we carried out PCR using designed primers (Supplementary Table 2, MKK2-S-7F and MKK2-S-
314 1139R), Easy Taq polymerase (TransGene Biotech), and 50 ng of extracted genomic DNA as a
315 template. All desired bands were then cut off from gels, purified, and sent to the company for
316 sequencing (GeneScript, Nanjing), alignment, and analysis (Supplementary 7-11). Southern
317 blotting was performed to verify the integration of *BleoR* into the poplar genome. 500 ng of
318 genomic DNA was cleaved with *BamH*I and *Hind*III at 37 °C for 4 h. The digested DNA was then
319 used as a PCR template to label a 160 bp probe from integrated *BleoR* CDS into the genomic
320 DNA (Supplementary Table 2; S-F and -R). In this step, we used the DIG (digoxigenin) reagent,
321 according to the manufacturer's instruction (catalog number 11745832910; Roche, Basel,
322 Switzerland). PCR product was then segregated on a 0.8% agarose gel. The separated fragments
323 were shifted on a Hybond N+ nylon membrane (Amersham Biosciences BV, Eindhoven, The
324 Netherlands).

325 For extraction of proteins, 150 mg fresh leaves of five weeks grown buds were milled in
326 500 µl extraction buffer (125 mM Tris, pH 6.8, 4 M Urea, 5% β -mercaptoethanol, 4% w/v SDS).
327 The centrifuge was then performed at 13,000 rpm for 10 min, and the supernatant was
328 obtained for gel analysis. The extracted protein was then boiled in loading buffer (24% w/v

329 glycerol, 100 mM Tris, a drop amount of Bromophenol Blue, 4% v/v β -mercaptoethanol, 8%
330 w/v SDS) for 10 min. The extracted protein was analyzed by SDS-PAGE and conceived using
331 Coomassie brilliant blue R-250 staining. After that, we carried out western blotting according to
332 Sambrook et al. (1989) using a rabbit anti-His polyclonal antibody developed in our laboratory
333 as the primary and peroxidase-conjugated goat antirabbit IgG (Zhongshan Biotechnique, Beijing,
334 China) as the secondary antibody.

335 **TaqMan real-time PCR**

336 To test the effect of designed parameters in all experiments on the proper integration of
337 exogenous *BleoR* with both homology arms, we decided to run the TaqMan assay applying dye
338 labels such as FAM and VIC adopting Applied Biosystem real-time PCR (Applied Biosystems,
339 USA). We used high quality extracted genomic DNA (Refer to the southern blotting) as the
340 template for running TaqMan real-time PCR. In this assay, two fluorescent FAM and VIC will
341 attach to the 5' region of the probe, while a non-fluorescent quencher (NFQ) binds to the 3'
342 region. Thus, we designed primers to probe two 150 bp fragments FAM1 (Supplementary 2,
343 FAM1-F and -R) and FAM2 (Supplementary 2, FAM2-F and -R). These primers were designed so
344 that FAM1 was able to probe 114 bp nucleotides from the 5' homology arm and 36 bp from
345 BleoR. Besides, FAM2 could probe 105 bp nucleotides from the 3' homology arm and 45 bp
346 from the BleoR (Supplementary 12). We also designed primers (Supplementary Table2, VIC-F,
347 and -R) to probe one 106 bp fragment VIC on the *actin* gene as the reference with a stable copy
348 number (Supplementary 12). All samples were analyzed in quadruplicate.

349 **Evaluation of HDR efficiency**

350 To evaluate the HDR efficiency, we decided to calculate and compare the ΔΔCt mean of
351 BleoR expression integrated into the poplar genome from all grown buds in five designed
352 experiments separately. In this step, we used the synthesized cDNA (Point to the RT-PCR
353 section) and designed primers (Supplementary Table 2, BleoR-52F and -151R) to carry out real-
354 time PCR. We used the Fast Start Universal SYBR Green Master (Rox; No. 04913914001: Roche,
355 USA) and performed three technical repeats for each event. Then, we used ANOVA-One way to
356 analyze the achieved mean data and compared.

**RT-qPCR**

We applied RT-qPCR using synthesized cDNA from grown buds on Zeocin (as mentioned above) as the template and designed primers (Supplementary Table2, RT-qPCR part) to investigate the expression of *BleoR* and *MKK2* genes and their impact on each other. We also explored our method's impact to develop HDR efficiency on HDR (CtIP, MRE11, BRCA1, Rad50, and Rad51) and NHEJ (Lig4, XRCC4) influential factors.

**Functional analyses and phenotypic properties**

Regarding the critical roles of MKK2 in plant protection against environmental stresses (Mohanta et al., 2015; Teige et al., 2004), and to confirm the correct HDR happened through recovered events, we decided to assess the functional analyzes of *MKK2* and phenotypic properties affected by salt stress and compare with WT. Recovered events were then planted on soil and transferred to the greenhouse. After two weeks of acclimation, we extracted total RNA from all transferred recovered events. Then, we irrigated all recovered events daily with 50 mM NaCl for one week. Some plants withered. Therefore, we extracted total RNA from survived events to perform RT-qPCR with triplicates for each event and analyzed their stem lengths (mm) and -diameter (mm) before and after salt stress.

**Statistical analysis**

All data were analyzed using ANOVA One-Way with Turkey means comparison calculated by OriginPro 2018 and Excel 2019 software (Microsoft, Redmond, WA, USA). Differences were analyzed statistically when the confidence intervals presented no overlap of the mean values with an error value of 0.05.

## Results and Discussion

Despite extensive research on using factors included in homology-directed repair pathway to enhance HDR efficiency in plants (An et al., 2020; Gil-Humanes et al., 2017; Hummel et al., 2018; Li et al., 2015; Svitashev et al., 2015), no research has been reported using this system in haploid woody species such as poplar.

**Strategies for target detection and HDR**

The importance of *the MKK2* gene role in plant protection has been shown against salt (Chen et al., 2020) and low temperature (Gao et al., 2017). In this study, the *MKK2* gene was

386 targeted for integrating *BleoR* into the poplar genome by improving HDR efficiency with Zeocin
387 resistance (Figure 1a). Guide RNA was designed near the 3' UTR to avoid rendering the gene
388 expression denoting higher specificity and less off-target (Hsu *et al.*, 2013; Doench *et al.*, 2014).
389 We targeted a detected CRISPR site located on exon 8 with the highest activity score and no off-
390 target effects on CDS throughout the whole *P. trichocarpa* genome (Figure 1a and
391 Supplementary Table 1). Afterthought and regarding (Song and Stieger, 2017), we optimized
392 the lengths of homologous arms (data not shown) to apply 400 bp upstream and downstream
393 of the PAM site as the 5' and 3' homology arms, respectively (Figure 1b). Furthermore,
394 particular sequences instead of remained nucleotides from exon 8 (Leu-Ala-Thr-Leu-Lys-Thr-Cys)
395 and exon 9 (Val-Leu-Val-Lys-Met) were added to the end of the 5' homology arm
396 (Supplementary 1b). Then, 375 bp *BleoR* CDS, 18 bp 6xHis tag, and 30 bp Poly A were designed
397 to attach the DDT sequences.

398 The DDT cassette was ligated into the pRGEB31 vector containing the Cas9 expression
399 cassette to construct the pDDT. It is worth remarking that to improve the HDR efficiency, we
400 overexpressed HDR effectors CtIP and MRE11 (Tran et al., 2019) and mutated the HDR inhibitor
401 XRCC4 (Pierce et al., 2001) simultaneously in this research. We constructed multiple fusion
402 vectors, including Cas9 expression cassette, CtIP overexpression cassette, MRE11
403 overexpression cassette, and *XRCC4* mutative cassette, improving the efficiency
404 (Supplementary 5a-d).

405 In this study, and to expand the genome editing possible for woody plants, we applied
406 Agrobacterium to deliver the optimized CRISPR–Cas9 system into poplar stems and generate
407 stable transgenic poplars. Because (Yang et al., 2016) reported that the HDR efficiency is
408 directly related to the amount of DDT at S/G2 cell division phases, we used pathogenic
409 suspension with an $OD_{600}$=2.5 (~$2\times10^9$ cell ml-1) and the ratio of 4:1 pDDT/pgRNA to increase
410 DDT fragments during S/G2 cell division (Tran et al., 2019) and to avoid off-target editing
411 caused by the extra accumulation of pgRNA (Hajiahmadi et al., 2019) (Figure 2a). Transformant
412 grown buds on zeocin were selected to analyze (Figure 2a).

413 According to figure 2b, we transferred pgRNA and pDDT into the explants via
414 experiment I (ExI). After that, we decided to design a plasmid that included a fused CtIP (Tran et

al., 2019) and Cas9 (pgCtIP) instead of pgRNA with a ratio of 4:1 pDDT/pgCtIP to promote HDR efficiency in poplars via experiment II (ExII) (Figure 2b). Only seventeen events were observed to be grown from a total of 42 regenerated buds. Also, only one recovered event was discerned after transferring on rooting media. In continuous and to investigate the effect of MRE11 (Tran et al., 2019) to improve HDR efficiency in poplars, we designed plasmid harboring a combined MRE11 and Cas9 (pgMR) instead of pgRNA with the same ratio of 4:1 pDDT/pgMR via experiment III (ExIII) (Figure 2b). In this experiment, we observed fifteen grown buds, and only one recovered edited event. Because our experiments did not show significant recovered events in overcoming NHEJ to integrate *BleoR*, we determined to design experiment IV (ExIV), including one plasmid harboring fused both MRE11 and CtIP with Cas9 (pgCtMR). Recovered events were increased insignificantly to four (Figure 2b). Therefore, we decided to target *XRCC4* as one key factor in the NHEJ pathway (Maruyama et al., 2015) besides CtIP and MRE11 overexpressing. For this purpose, we designed experiment V (ExV) using one plasmid harboring *XRCC4* gRNA and also fused both MRE11 and CtIP with Cas9 (pggCtMR). We tried to transfer this plasmid into the plant cells with the same ratio of 4:1 pDDT/pggCtMR. In this experiment, recovered events were shown increased surprisingly to twelve events from thirty-one grown buds on selection media (Figure 2b).

**Transformants verification via Western blotting, RT-PCR, and Southern blotting**

Several methods, such as western blotting, RT-PCR, and Southern blotting, were applied to verify the happened HDR in resistant transformants. Several methods, such as western blotting, RT-PCR, and Southern blotting, were applied to verify the happened HDR in resistant transformants. We decided to combine a 6XHis tag to the BleoR CDS (Figure 1b) to show the integration of *BleoR* and *MKK2* through target genomes using Western blotting. Through screening the western blotting of all transformants grown on zeocin, we detected no edited events in ExI, but one ExII event (II#29) showed a band of 54 KDa (Figure 3a), which might be the candidate for the successful integration of the BleoR (~13.7 KDa) fused by the MKK2 (~40.5 KDa) (Figure 3b). In continue and through ExIII events, we found only one event (III#6) with a band of 14 KDa (Figure 3a), and we deemed that only BleoR CDS was integrated completely and sudden mutations might knock out *MKK2* throughout exons7, 8, or 9 (Figure 3b). We then

444 screened ExIV events and found three events (IV#17, #54, and #68) with bands about 54 KDa,
445 and one event (IV#90) about 14 KDa (Figure 3a). Surprisingly, we found ten events (V#21, #25,
446 #29, #32, #39, #59, #73, #88, #91, and #94) with bands about 54 KDa within ExV, and two
447 events (V#37, and #53) with bands about 14 kDa (Figure 3a).

448 We designed two RT-PCR assays to confirm the designed editions in transformed
449 poplars. Considering the targeting of exon8 from *MKK2* (Figure 1b) in this study and then our
450 attempts to repair the damaged area together with the integration of *the BleoR* gene, we
451 decided to design the first RT-PCR experiment in such a way as to repair the infected area and
452 reassure the correct *MKK2* gene expression in transformant poplars (Figure 3c and d). As
453 expected, no bands were observed from ExI events. Three ExII events (#24, #29, and #35) and
454 four ExIII events (#10, #23, #36, and #45) showed 920 bp bands (Figure 3e) respectively. Also,
455 ExIV events (#9, #17, #39, #45, #54, #60, #68, #72, and #83) showed similar bands with WT.
456 Surprisingly, twenty events of ExV exhibited the complete *MKK2* expression with similar bands
457 with WT. Regarding the binding of MKK2 and BleoR in the target genome (Figure 1b) and verify
458 that, and also according to the design of DDT to include the required nucleotides from exon7,
459 we decided to design a second RT-PCR using pDDT as the positive control (Figure 3d). The
460 second RT-PCR also revealed no desired 413 bp band through ExI events (Figure 3f). ExII events
461 revealed only one 413 bp amplification (#29), but ExIII events revealed no desired band. We
462 then considered ExIV and observed three events (#17, #54, and #68), while ExV revealed
463 significantly increased ten events (#21, #25, #29, #32, #39, #59, #73, #88, #91, and #94) as 413
464 bp bands. All events showing bands with Western blotting and RT-PCR assays were selected for
465 further analyses. The probe of *BleoR* sequences was designed for Southern blotting (Figure 3g).
466 Several events (III#6, IV#90, V#37, and V#53), which were not amplified through the second RT-
467 PCR, were reverified in Southern blotting.

468 **Accurate investigation of edited events and HDR efficiency**

469 Cermak et al. (2015) could improve the HDR efficiency by ten-fold in tomato and
470 integrated the 35S promoter in upstream of the *ANT1* gene. Tran et al. (2019) could promote
471 the HDR efficiency in mammalians with at least a 2-fold increase in HDR and a 6-fold increase in
472 HDR/NHEJ ratio. Aslan et al. (2017) increased HDR efficiency to 25.7% in the *Xenopus*

473 *tropicalis* genome by inserting mall pieces of DNA, while Danilo et al. (2018) succeeded in
474 knocking in a DDT with 400 bp into the tomato genome with low efficiency of 1.29%. Jang et al.
475 (2018) attempted to develop HDR efficiency to 38% in mouse lines by applying multiple sgRNAs,
476 and Menchaca et al. (2020) tried to enhance the HDR efficiency to 61.5% in sheep utilizing
477 single strand oligodeoxynucleotides. Still, there are some reports on *Populus* genome editing,
478 but they are limited only to knock out genes, and mutations happened by Cas9 and Cas12a (An
479 et al., 2020; Di Fan et al., 2015; Liu et al., 2015), but no report has been carried out on
480 improving the HDR efficiency in poplar. In this study, we tried to create a recombinant genome
481 in poplars using an HDR system. We then tried to detect the actual HDR events via TaqMan
482 real-time PCR.

483 Two probes, FAM1 and FAM2, were designed for the 3' and 5'-end *BleoR* CDS area
484 (Figure 4a). The transformants were assumed to be edited via exhibiting both fluorescent
485 signals of FAM1 and FAM2 (Figure 4b). In ExI, the averages of fluorescent signals of FAM1 ΔΔCt
486 and FAM2 ΔΔCt were shown proximal to 0 (Supplementary 13a). Most signals of ExI events
487 exhibited as the mutant or WT, with a few partial FAM1 or FAM2 fluorescence (Figure 4c).
488 Sanger sequencing was conducted to confirm these results. So, no edited events were achieved
489 through ExI (Supplementary 7). Transformants of ExII and -III showed an increase in the signal
490 densities of the FAM1 and FAM2 (Figure 4d, and e). The signal average of ExII transformants
491 were about 14.5 and 13.5, while ExIII events exhibited more FAM1 ΔΔCt about 16 and a lesser
492 FAM2 ΔΔCt about 10 (Supplementary 13b, and c). The alignment of Sanger sequencing well
493 conformed to this result. In ExII, we found four fully edited events (II#7, II#19, II#53, and II#59),
494 four FAM1 partial edited events (II#13, II#21, II#35, and II#41), and four FAM2 partial edited
495 events (II#3, II#11, II#14, and II#23) (Supplementary 8; Supplementary 14a). In ExIII, we verified
496 three fully edited events (III#21, III#45, and III#61), five FAM1 partial edited events (III#10,
497 III#23, III#27, III#32, and III#53), and three FAM2 partial edited events (III#6, III17, and III#36)
498 (Supplementary 9; Supplementary 14b). In ExIV, the signal density of edited events increased
499 significantly (Figure 4f). The mean fluorescent signals of FAM1 ΔΔCt and FAM2 ΔΔCt showed an
500 increase of about 19 and 15, respectively (Supplementary 13d). In total, nine fully edited events
501 (IV#9, IV#27, IV#39, IV#45, IV#54, IV#68, IV#79, IV#83, and IV#90), seven FAM1 partial edited

events (IV#13, IV#17, IV#19, IV#46, IV#60, IV#75, and IV#85), and four FAM2 partial edited events (IV#13, IV#76, IV#80, and IV#92) (Supplementary 10; Supplementary 14c) were detected. Finally, the signal density of edited transformants in ExV was remarkably increased (Figure 4g). The mean fluorescent signals of FAM1 ΔΔCt and FAM2 ΔΔCt were increased about 21.5 and 18, respectively (Supplementary 13e). In total, 15 fully edited events (V#3, V#9, V#21, V#25, V#29, V#33, V#39, V#67, V#73, V#79, V#88, V#91, V#92, V#94, and V#101) were discovered (Supplementary 11; Supplementary 14d; Supplementary 17). Total FAM fluorescent signals (FAM1, FAM2, and FAM1&2) indicated the promotion of happened HDR through experiment events (Figure 4h). Overexpression of HDR effectors, CtIP, and MRE11 (ExIV), could dramatically increase FAM signals in comparison with the only overexpression of CtIP (ExII) or MRE11 (ExIII). Moreover, *XRCC4* deficiency, simultaneously with the overexpression of CtIP and MRE11, dramatically improved the HDR in ExV.

According to (Shao et al., 2017), improving the HDR factors causes to improve HDR efficiency. Tran et al. (2019) could improve the HDR efficiency up to 19-fold with overexpressing CtIP and MRE11 and inhibiting NHEJ factors (Li et al., 2018). In this study, we considered inhibiting XRCC4 as one LIG4 cofactor (Grawunder et al., 1998b) to increase HDR efficiency. The HDR efficiency investigation confirmed that the NHEJ pathway deficiency is meaningfully more efficient to HDR development than on only the overexpression of HDR factors (Figure 4i). Overexpression of CtIP, and MRE11, could dramatically increase FAM signals compared to the only overexpression of CtIP or MRE11. Furthermore, *XRCC4* deficiency, together with the overexpression of CtIP and MRE11, dramatically enhanced the HDR in ExV. According to Figure 4i, the expression of integrated exogenous *BleoR* represents HDR efficiency. We performed real-time PCR to evaluate the percentage of *BleoR* expression $\Delta\Delta C_t$ mean (Supplementary 15a) and then compare and illustrate the bar plot supported by standard distribution curves (Figure 4i; Supplementary 15b). The *BleoR* expression increased from -1.2287 in ExI to 4.40787 (5.6%) in ExII and 6.11543 (7.3%) in ExIII. The expression in ExIV and ExV raised to 19.06057 (20.26%) and 48.90032 (50.1%). We resulted that the *XRCC4* deficiency, together with the overexpression of CtIP and MRE11, was the most efficient system for HDR-based integration and improved more HDR happenings than the expression of HDR effectors CtIP and MRE11

(Figure 4i). We resulted that the *XRCC4* deficiency, together with the overexpression of CtIP and MRE11, was the most efficient system for HDR-based integration and improved more HDR happenings than the expression of HDR effectors CtIP and MRE11.

**The effect of efficient HDR on the expression of NHEJ and HDR factors**

It is required to investigate targeted *MKK2* and integrated *BleoR* gene expressions and their interdependence. The expression of each of them proves the HDR happening and shows the regular functions of these genes in the new version of the poplar genome. We analyzed RT-qPCR achieved data and used Violon plots to describe the distributed expressions of *MKK2* and *BleoR* genes through all events (Figure 5a-d) and Column plots to show their expressions for each event separately (Supplementary 16). Analysis among ExI events revealed distributed *MKK2* and *BleoR* expressions of about +1 and -1 but, the ExII events exhibited above expressions between about 100 and zero with medians about zero (Figure 5a and b). Within the ExII, we discovered three events with *MKK2* expression (#21, #29, and #35) and only one event with *BleoR* expression (#29) (Supplementary 16). We then analyzed these gene expressions from ExIII and discovered promoted distributed expressions of *MKK2* (Figure 5c). In this experiment, we found four events with *MKK2* expression (#10, #23, #36, and #45) and one event with *BleoR* expression (#6) (Supplementary 16). RT-qPCR results through ExV events revealed enhanced distributions of *BleoR* and *MKK2* expressions (Figure 5d). In this experiment, we achieved nine events with *MKK2* expression (#9, #17, #39, #45, #54, #60, #68, #79, and #83) and four events with *BleoR* expression (#17, #54, #68, and #90) (Supplementary 16). Regarding the expressions of these genes from ExV events, we observed significant promotions in *MKK2* and *BleoR* distributed expressions with a median of about 100 (Figure 5e). Also, the column bar analysis confirmed these distributions with twenty *MKK2* expressions within ExV events (#3, #9, #18, #21, #25, #29, #32, #33, #39, #59, #67, #73, #79, #82, #86, #88, #91, #92, #94, and #101) and twelve *BleoR* expressions (#21, #25, #29, #32, #37, #39, #53, #59, #73, #88, #91, and #94) (Supplementary 16).

Moreover, we decided to assess gene expressions involved in HDR and NHEJ pathways affected by our increasing HDR efficiency plans. We used a Heat-map plot to interpret the obtained data from RT-qPCR (Figure 5f). While the expressions

560 of *CtIP* (~116%), *MRE11* (~115%), *BRCA1* (~114%), *Rad50* (~113%), and *Rad51* (~116%) were
561 increased through ExI compared to WT, more expressions were observed in *Lig4* (~146%)
562 and *XRCC4* (~143%) (Figure 5f). The expression of *CtIP* was increased impressively (~166%) via
563 ExII compared with WT, while the expressions
564 of *MRE11* (~129%), *BRCA1* (~119%), *Rad50* (~120%), and *Rad51* (~121%) were increased
565 insignificantly. Through ExII, the expressions of *Lig4* (~104%) and *XRC44* (~105%) were
566 decreased compared to ExI (Figure 5f). Within ExIII, the expression of *MRE11* was increased
567 impressively (~162%), but the expressions of *CtIP* (~134%), *BRCA1* (~120%), *Rad50* (~122%),
568 and *Rad51* (~119%) were increased slightly. Within this experiment, the expressions
569 of *Lig4* (~107%) and *XRC44* (~103%) were decreased, contrasted with ExI (Figure 4f). All HDR
570 factors revealed enhanced expressions among ExIV
571 as *CtIP* (~165%), *MRE11* (~164%), *BRCA1* (~130%), *Rad50* (~128%), and *Rad51* (~129%), but the
572 expressions of *Lig4* (~101%) and *XRC44* (~99%) were decreased more compared with ExI (Figure
573 5f). *XRCC4* deficiency in ExV and enhancing *CtIP* and *MRE11* expressions caused enhancing the
574 expressions of *CtIP* (~170%) and *MRE11* (~165%) much more than WT events. Also, the
575 expression of *Lig4* (~87%) revealed more decreased than WT, and *XRCC4* was knocked out.
576 Through ExV, the other HDR factors *BRCA1* (~145%), *Rad50* (~139%), and *Rad51* (~142%) also
577 revealed more expressions (Figure 5f).

578 ***XRCC4* deficient dramatically enhanced HDR efficiency and decreased polymorphisms**

579 NHEJ is identified by introducing irregular small insertions or deletions (indels) into the
580 targeted site. Notwithstanding the mutagenicity of NHEJ for its error-prone trait, its active
581 kinetics has a role in repairing genome integrity, distinctly by crushing chromosomal
582 translocations, mostly for the bulk of repair events (Ceccaldi et al., 2016). It has been exhibited
583 a numerous percentage of about 70% of polymorphisms, especially deletions in Populus
584 genome editing by the Cas system while knocking down the PDS gene family (An et al., 2020).
585 Menchaca et al. (2020) recorded 17.8% indel mutations in sheep genome editing, with 61.5%
586 knock-in mutations happened by HDR. Also, it has been shown that silencing of NHEJ factors
587 such as *Ku70* and *Ku80* caused to reduce indels significantly from 64% to 38% and 39.4%,
588 respectively (Li et al., 2018).

In this study, and to test whether the HDR promotion affects decreasing the polymorphisms, we analyzed the variant genotypes, protein effects, and nucleotide genotyping within the 5' and 3' homologous arms and also the knocked-in fragments from recovered events. We firstly analyzed the homology arms to detect polymorphisms and detected seven polymorphism varieties, including deletions, deletion tandem repeats, insertions, insertion tandem repeats, SNP transitions (A to C or G to T and reversely), SNP transversions (Purines to pyrimidines or reversely), and substitutions (Supplementary Table 3). It has been shown that the promotion of HDR via ExI to ExV events push to migrate Insertion and Deletion (InDel) nucleotides from the 5' region of knocked-in fragments (BleoR) to the 3' region (6xHis and PolyA) (Figure 6a). It means that we achieved more *BleoR* expression via ExV events compared to the other experiments. The mean comparisons of the InDel nucleotides revealed that the promoted HDR by *XRCC4* deficiency caused to decrease in InDel nucleotides considerably via ExV events (65) compared to the ExI (454), ExII (263), ExIII (177), and ExIV events (295) (Figure 6b). We then calculated the only number of happened polymorphisms (Not nucleotides) and then compared the means obtained from the total polymorphisms issued within designed experiments. It has been observed that the HDR happening through ExI events provoked the highest polymorphisms, significantly more than ExIV and -V events (Figure 6c). We also observed more happened polymorphisms within ExII and -III than ExIV and -V (Figure 6c). We then decided to investigate all happened polymorphisms in more detail from the homology arms and knocked-in *BleoR* from the recovered poplar genome (Supplementary Table 4). We identified the highest frequency of deletions through ExI events and the least within ExV events (Figure 6d). We also observed that *XRCC4* deficiency revealed the least of SNP transition. The overexpression of *CtIP* decremented deletion tandem repeats (ExII), and the overexpression of *MRE11* decremented SNPs, SNP transitions, and SNP transversions (Figure 6d). The overexpression of *CtIP* and *MRE11* simultaneously (ExIV) decremented substitution polymorphisms (Figure 6d). Moreover, the whisker plot of total polymorphisms presented the maximum distribution of polymorphisms through ExI events and the minimum of those in ExV events (Figure 6d).

**Functional analyzes of MKK2 and phenotypic changes**

It has been shown that MAPKs genes direct cellular responses against abiotic stresses such as salinity (Mohanta et al., 2015; Sun et al., 2015). Regarding these results, we decided to evaluate *MKK2* expression from re-planted recovered events on the soil before and after salt stress to show the proper *MKK2* locus maintain by efficient HDR through exon7, 8, and 9. The functional analysis of *MKK2* in survived recovered events (II#29, IV#17, IV#54, IV#68, V#21, V#25, V#29, V#32, V#39, V#59, V#73, V#88, V#91, and V#94) from ExII, -IV and -V and comparing them with WT poplars revealed a regular expression in all events (~95%-100%) before stress and a stable overexpression induced by salt (168%-173%) after stress (Figure 7; Supplementary Table 5). Also, stem lengths and diameters of all mentioned above events before and after salt stress were measured and revealed no significant differences between survived recovered events and WT poplars (Figure 7; Supplementary Table 5).

In summary, we have proved that *XRCC4* deficiency caused to enhance the HDR efficiency meaningfully, therefore greatly expanding our capacity to improve hereditary developments in poplar. This breakthrough technology is likely to encourage biotechnological researches, breeding programs, and forest conservation of tree species.

**Supplementary information**

Supplemental information is available for this paper.

**Funding**

This project was funded by the National Key Program on Transgenic Research (2018ZX08020002), the National Natural Science Foundation of China (No. 31971682, 31570650), the Priority Academic Program Development of Jiangsu Higher Education Institutions, the Talent Funding Project of Nanjing Forestry University (No. 163108059).

**Author contribution**

AM: Conceptualization, Software, Formal analysis, Writing - Original Draft, Visualization, Project administration, and Funding acquisition; HW: Methodology, Formal analysis, Writing - Review & Editing and Data Curation; ZHC: Conceptualization, Validation, Data Curation, Writing - Review & Editing; WS, JZ, DL: Validation, Writing - Review & Editing; LY: Conceptualization,

Software, Formal analysis, Visualization, and Funding acquisition; QZ: Conceptualization, Software, Formal analysis, Visualization, Supervision, and Funding acquisition.

**Conflict of interest**

The authors declare that they have no conflict of interest.

**Acknowledgment**

We thank all researchers, especially professor Zhong-Hua Chen, to improve this research with their directions.

## References

Ali, Z., Shami, A., Sedeek, K., Kamel, R., Alhabsi, A., Tehseen, M., Hassan, N., Butt, H., Kababji, A., Hamdan, S. M. and Mahfouz, M. M. (2020) Fusion of the Cas9 endonuclease and the VirD2 relaxase facilitates homology-directed repair for precise genome engineering in rice. *Commun Biol* **3**.

An, Y., Geng, Y., Yao, J., Fu, C., Lu, M., Wang, C. and Du, J. (2020) Efficient Genome Editing in Populus Using CRISPR/Cas12a. *Front Plant Sci* **11**, 593938.

Aslan, Y., Tadjuidje, E., Zorn, A. M. and Cha, S. W. (2017) High-efficiency non-mosaic CRISPR-mediated knock-in and indel mutation in F0 Xenopus. *Development* **144**, 2852–2858.

Bewg, W. P., Ci, D. and Tsai, C.-J. (2018) Genome Editing in Trees: From Multiple Repair Pathways to Long-Term Stability. *Frontiers in Plant Science* **9**.

Butler, N. M., Baltes, N. J., Voytas, D. F. and Douches, D. S. (2016) Geminivirus-Mediated Genome Editing in Potato (Solanum tuberosum L.) Using Sequence-Specific Nucleases. *Front Plant Sci* **7**, 1045.

Ceccaldi, R., Rondinelli, B. and D'Andrea, A. D. (2016) Repair Pathway Choices and Consequences at the Double-Strand Break. *Trends in cell biology* **26**, 52–64.

Cermak, T., Baltes, N. J., Cegan, R., Zhang, Y. and Voytas, D. F. (2015) High-frequency, precise modification of the tomato genome. *Genome Biol* **16**, 232.

Chen, X. Y., Wang, P. K., Zhao, F. F., Lu, L., Long, X. F., Hao, Z. D., Shi, J. S. and Chen, J. H. (2020) The Liriodendron chinense MKK2 Gene Enhances Arabidopsis thaliana Salt Resistance. *Forests* **11**.

Dahan-Meir, T., Filler-Hayut, S., Melamed-Bessudo, C., Bocobza, S., Czosnek, H., Aharoni, A., and Levy, A. A. (2018) Efficient in planta gene targeting in tomato using geminiviral replicons and the CRISPR/Cas9 system. *Plant J* **95**, 5–16.

Danilo, B., Perrot, L., Botton, E., Nogue, F. and Mazier, M. (2018) The DFR locus: A smart landing pad for targeted transgene insertion in tomato. *Plos One* **13**.

Di Fan, Liu, T., Li, C., Jiao, B., Li, S., Hou, Y. and Luo, K. (2015) Efficient CRISPR/Cas9-mediated Targeted Mutagenesis in Populus in the First Generation. *Sci Rep-Uk* **5**, 12217.

Doench, J. G., Hartenian, E., Graham, D. B., Tothova, Z., Hegde, M., Smith, I., Sullender, M., Ebert, B. L., Xavier, R. J. and Root, D. E. (2014) Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat Biotechnol* **32**, 1262–1267.

Friesner, J. and Britt, A. B. (2003) Ku80- and DNA ligase IV-deficient plants are sensitive to ionizing radiation and defective in T-DNA integration. *Plant J* **34**, 427–440.

Gao, J., Zhang, S., He, W. D., Shao, X. H., Li, C. Y., Wei, Y. R., Deng, G. M., Kuang, R. B., Hu, C. H., Yi, G. J. and Yang, Q. S. (2017) Comparative Phosphoproteomics Reveals an Important Role of MKK2 in Banana (Musa spp.) Cold Signal Network. *Sci Rep-Uk* **7**.

Gil-Humanes, J., Wang, Y., Liang, Z., Shan, Q., Ozuna, C. V., Sanchez-Leon, S., Baltes, N. J., Starker, C., Barro, F., Gao, C. and Voytas, D. F. (2017) High-efficiency gene targeting in hexaploid wheat using DNA replicons and CRISPR/Cas9. *Plant J* **89**, 1251–1262.

Grawunder, U., Zimmer, D., Fugmann, S., Schwarz, K. and Lieber, M. R. (1998a) DNA ligase IV is essential for V(D)J recombination and DNA double-strand break repair in human precursor lymphocytes. *Mol Cell* **2**, 477–484.

Grawunder, U., Zimmer, D., Kulesza, P., and Lieber, M. R. (1998b) Requirement for an interaction of XRCC4 with DNA ligase IV for wild-type V(D)J recombination and DNA double-strand break repair in vivo. *J Biol Chem* **273**, 24708–24714.

Hajiahmadi, Z., Movahedi, A., Wei, H., Li, D. W., Orooji, Y., Ruan, H. H. and Zhuge, Q. (2019) Strategies to Increase On-Target and Reduce Off-Target Effects of the CRISPR/Cas9 System in Plants. *Int J Mol Sci* **20**.

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., Li, Y., Fine, E. J., Wu, X., Shalem, O., Cradick, T. J., Marraffini, L. A., Bao, G. and Zhang, F. (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* **31**, 827–832.

Hummel, A. W., Chauhan, R. D., Cermak, T., Mutka, A. M., Vijayaraghavan, A., Boyher, A., Starker, C. G., Bart, R., Voytas, D. F. and Taylor, N. J. (2018) Allele exchange at the EPSPS locus confers glyphosate tolerance in cassava. *Plant Biotechnol J* **16**, 1275–1282.

Jang, D. E., Lee, J. Y., Lee, J. H., Koo, O. J., Bae, H. S., Jung, M. H., Bae, J. H., Hwang, W. S., Chang, Y. J., Lee, Y. H., Lee, H. W. and Yeom, S. C. (2018) Multiple sgRNAs with overlapping sequences enhance CRISPR/Cas9-mediated knock-in efficiency. *Exp Mol Med* **50**.

Li, G. L., Liu, D. W., Zhang, X. W., Quan, R., Zhong, C. L., Mo, J. X., Huang, Y. Q., Wang, H. Q., Ruan, X. F., Xu, Z., Zheng, E. Q., Gu, T., Hong, L. J., Li, Z. C., Wu, Z. F., and Yang, H. Q. (2018) Suppressing Ku70/Ku80 expression elevates homology-directed repair efficiency in primary fibroblasts. *Int J Biochem Cell B* **99**, 154–160.

Li, J., Hong, S., Chen, W., Zuo, E. and Yang, H. (2019a) Advances in detecting and reducing off-target effects generated by CRISPR-mediated genome editing. *J Genet Genomics.*

Li, J., Meng, X., Zong, Y., Chen, K., Zhang, H., Liu, J. and Gao, C. (2016) Gene replacements and insertions in rice by intron targeting using CRISPR-Cas9. *Nat Plants* **2**, 16139.

Li, S., Li, J., He, Y., Xu, M., Zhang, J., Du, W., Zhao, Y. and Xia, L. (2019b) Precise gene replacement in rice by RNA transcript-templated homologous recombination. *Nat Biotechnol* **37**, 445–450.

Li, Z., Liu, Z. B., Xing, A., Moon, B. P., Koellhoffer, J. P., Huang, L., Ward, R. T., Clifton, E., Falco, S. C., and Cigan, A. M. (2015) Cas9-Guide RNA Directed Genome Editing in Soybean. *Plant Physiol* **169**, 960–970.

Lin, Q., Zong, Y., Xue, C., Wang, S., Jin, S., Zhu, Z., Wang, Y., Anzalone, A. V., Raguram, A., Doman, J. L., Liu, D. R. and Gao, C. (2020) Prime genome editing in rice and wheat. *Nat Biotechnol* **38**, 582–585.

Liu, M., Rehman, S., Tang, X., Gu, K., Fan, Q., Chen, D. and Ma, W. (2019) Methodologies for Improving HDR Efficiency. *Frontiers in Genetics* **9**.

Liu, T.-t., Di Fan, Ran, L.-y., Jiang, Y.-z., Liu, R. and Luo, K.-m. (2015) Highly efficient CRISPR/Cas9-mediated targeted mutagenesis of multiple genes in Populus. *Yi Chuan = Hereditas* **37**, 1044–1052.

Lu, Y., Tian, Y., Shen, R., Yao, Q., Wang, M., Chen, M., Dong, J., Zhang, T., Li, F., Lei, M. and Zhu, J. K. (2020) Targeted, efficient sequence insertion and replacement in rice. *Nat Biotechnol* **38**, 1402–1407.

Maruyama, T., Dougan, S. K., Truttmann, M. C., Bilate, A. M., Ingram, J. R. and Ploegh, H. L. (2015) Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of non-homologous end joining. *Nat Biotechnol* **33**, 538–542.

Menchaca, A., dos Santos-Neto, P. C., Souza-Neves, M., Cuadro, F., Mulet, A. P., Tesson, L., Chenouard, V., Guiffes, A., Heslan, J. M., Gantier, M., Anegon, I. and Crispo, M. (2020) Otoferlin gene editing in sheep via CRISPR-assisted ssODN-mediated Homology Directed Repair. *Sci Rep-Uk* **10**.

Mohanta, T. K., Arora, P. K., Mohanta, N., Parida, P. and Bae, H. (2015) Identification of new members of the MAPK gene family in plants shows diverse conserved domains and novel activation loop variants. *BMC genomics* **16**, 58.

Movahedi, A., Zhang, J. X., Amirian, R. and Zhuge, Q. (2014) An Efficient Agrobacterium-mediated Transformation System for Poplar. *Int J Mol Sci* **15**, 10780–10793.

Movahedi, A., Zhang, J. X., Gao, P. H., Yang, Y., Wang, L. K., Yin, T. M., Kadkhodaei, S., Ebrahimi, M. and Qiang, Z. G. (2015) Expression of the chickpea CarNAC3 gene enhances salinity and drought tolerance in transgenic poplars. *Plant Cell Tiss Org* **120**, 141–154.

Panier, S. and Boulton, S. J. (2014) Double-strand break repair: 53BP1 comes into focus. *Nat Rev Mol Cell Biol* **15**, 7–18.

Pierce, A. J., Hu, P., Han, M., Ellis, N. and Jasin, M. (2001) Ku DNA end-binding protein modulates homologous repair of double-strand breaks in mammalian cells. *Genes Dev* **15**, 3237–3242.

Puchta, H. (2005) The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. *J Exp Bot* **56**, 1–14.

Qi, Y., Zhang, Y., Zhang, F., Baller, J. A., Cleland, S. C., Ryu, Y., Starker, C. G. and Voytas, D. F. (2013) Increasing frequencies of site-specific mutagenesis and gene targeting in Arabidopsis by manipulating DNA repair pathways. *Genome Res* **23**, 547–554.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). *Molecular cloning: a laboratory manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Schiml, S., Fauser, F. and Puchta, H. (2014) The CRISPR/Cas system can be used as nuclease for in planta gene targeting and as paired nickases for directed mutagenesis in Arabidopsis resulting in heritable progeny. *Plant J* **80**, 1139–1150.

Shao, S. M., Ren, C. H., Liu, Z. T., Bai, Y. C., Chen, Z. L., Wei, Z. H., Wang, X., Zhang, Z. Y. and Xu, K. (2017) Enhancing CRISPR/Cas9-mediated homology-directed repair in mammalian cells by expressing Saccharomyces cerevisiae Rad52. *Int J Biochem Cell B* **92**, 43–52.

Shi, J., Gao, H., Wang, H., Lafitte, H. R., Archibald, R. L., Yang, M., Hakimi, S. M., Mo, H. and Habben, J. E. (2017) ARGOS8 variants generated by CRISPR-Cas9 improve maize grain yield under field drought stress conditions. *Plant Biotechnol J* **15**, 207–216.

Song, F. and Stieger, K. (2017) Optimizing the DNA Donor Template for Homology-Directed Repair of Double-Strand Breaks. *Mol Ther Nucleic Acids* **7**, 53–60.

Sun, W., Chen, H., Wang, J., Sun, H. W., Yang, S. K., Sang, Y. L., Lu, X. B. and Xu, X. H. (2015) Expression analysis of genes encoding mitogen-activated protein kinases in maize provides a key link between abiotic stress signaling and plant reproduction. *Funct Integr Genomic* **15**, 107–120.

Svitashev, S., Schwartz, C., Lenderts, B., Young, J. K. and Cigan, A. M. (2016) Genome editing in maize directed by CRISPR-Cas9 ribonucleoprotein complexes. *Nat Commun* **7**.

Svitashev, S., Young, J. K., Schwartz, C., Gao, H., Falco, S. C. and Cigan, A. M. (2015) Targeted Mutagenesis, Precise Gene Editing, and Site-Specific Gene Insertion in Maize Using Cas9 and Guide RNA. *Plant Physiol* **169**, 931–945.

Symington, L. S., and Gautier, J. (2011) Double-strand break end resection and repair pathway choice. *Annu Rev Genet* **45**, 247–271.

Tang, X., Sretenovic, S., Ren, Q., Jia, X., Li, M., Fan, T., Yin, D., Xiang, S., Guo, Y., Liu, L., Zheng, X., Qi, Y. and Zhang, Y. (2020) Plant Prime Editors Enable Precise Gene Editing in Rice Cells. *Mol Plant* **13**, 667–670.

Teige, M., Scheikl, E., Eulgem, T., Dóczi, R., Ichimura, K., Shinozaki, K., Dangl, J. L., and Hirt, H. (2004) The MKK2 pathway mediates cold and salt stress signaling in Arabidopsis. *Mol Cell* **15**, 141–152.

Tran, N. T., Bashir, S., Li, X., Rossius, J., Chu, V. T., Rajewsky, K. and Kuhn, R. (2019) Enhancement of Precise Gene Editing by the Association of Cas9 With Homologous Recombination Factors. *Front Genet* **10**.

Van Vu, T., Sivankalyani, V., Kim, E.-J., Doan, D. T. H., Tran, M. T., Kim, J., Sung, Y. W., Park, M., Kang, Y. J. and Kim, J.-Y. (2020) Highly efficient homology-directed repair using CRISPR/Cpf1-geminiviral replicon in tomato. *Plant Biotechnol J*.

Wang, M. G., Lu, Y. M., Botella, J. R., Mao, Y. F., Hua, K. and Zhu, J. K. (2017) Gene Targeting by Homology-Directed Repair in Rice Using a Geminivirus-Based CRISPR/Cas9 System. *Mol Plant* **10**, 1007–1010.

Wu, J., and Yin, H. (2019) Engineering guide RNA to reduce the off-target effects of CRISPR. *J Genet Genomics*.

Xie, K. B., and Yang, Y. N. (2013) RNA-Guided Genome Editing in Plants Using a CRISPR-Cas System. *Mol Plant* **6**, 1975–1983.

Yang, D., Scavuzzo, M. A., Chmielowiec, J., Sharp, R., Bajic, A. and Borowiak, M. (2016) Enrichment of G2/M cell cycle phase in human pluripotent stem cells enhances HDR-mediated gene repair with customizable endonucleases. *Sci Rep* **6**, 21264.

Zhang, J. P., Li, X. L., Li, G. H., Chen, W., Arakaki, C., Botimer, G. D., Baylink, D., Zhang, L., Wen, W., Fu, Y. W., Xu, J., Chun, N., Yuan, W., Cheng, T. and Zhang, X. B. (2017) Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. *Genome Biol* **18**, 35.

Zhou, X., Jacobs, T. B., Xue, L. ‐ J., Harding, S. A. and Tsai, C. ‐ J. (2015) Exploiting SNP s for biallelic CRISPR mutations in the outcrossing woody perennial *Populus* reveals 4 ‐ coumarate: CoA ligase specificity and redundancy. *New Phytologist* **208**, 298–301.

## Figure legends

**Figure 1:** Schematics of *MKK2* locus before and after edition and integrating exogenous *BleoR* CDS into the poplar genome. **(a)** Schematic of this research purpose to integrate exogenous BleoR into the poplar genome. Dash line reveals the target site. **(b)** Protospacer Adjacent Motif (PAM) was detected at the end of exon 8 to lead Cas9. 400 bp sequences from both sides of the CRISPR target were selected for HDR in this study. The 5' homology arm included part sequences of the intron between exon 6 and -7, exon 7, intron sequences

822 between exon 7 and -8, and a part of exon 8. The 3' homology arm included intron sequences
823 between exon 8 and -9 and 3' UTR of the *MKK2* locus up to 400 bp. Designed DDT included
824 remained sequences of exon 8, exon 9, BleoR CDS, 6xHis, and PolyA sequences flanked by the
825 3'- and 5' homology arms. We added two special targets besides DDT. The DDT was then ligated
826 into the pRGEB31 vector to form pDDT.

827 **Figure 2:** The transformation strategy and designing of experiments. **(a)** pDDT and
828 pgRNA were mixed 4:1 and introduced to the Agrobacterium tumefaciens to form inoculator
829 suspension. We condensed the inoculator up to $OD_{600}$=2.5 and then dipped all cut off stems.
830 The putatively edited events were regenerated on Zeocin. We allowed putative edited events
831 to bud. The grown buds were then transferred on selective rooting media and allowed to be
832 recovered. Recovered events were then planted on soil and following two weeks of acclimation
833 introduced by salt stress. **(b)** Designed experiments for this study including (I) No HDR factors,
834 (II) overloaded CtIP, (III) overloaded MRE11, (IV) Overloaded CtIP+MRE11, and (V) Overloaded
835 CtIP+MRE11 with *XRCC4* deficiency.

836 **Figure 3:** Western blotting to reveal the fused 6xHis tag with BleoR integration into the
837 poplar genome. **(a)** Different experiments exhibited different quantities of 6xHis tag fusion. **(b)**
838 Schematic of fusion 6xHis tag with edited poplar genome triggered by different experiments.
839 Shape 1 reveals successful fusion of *BleoR* and *MKK2* with about 54 kDa. Shape 2 reveals an
840 unsuccessful combination of mentioned proteins with about 14 kDa. **(c)** Schematic of right HDR
841 happening caused to attach exon 8 and 9 in the edited genome. **(d)** Schematic of proper
842 integration in edited genome caused to connect the BleoR to the C-terminal of MKK2. **(e)** RT-
843 PCR exhibited the HDR in exon 8 and 9, revealing a 920 bp of transcribed *MKK2* RNA in
844 triggered events from ExII to ExV. The β-actin was used as the control in all RT-PCR assays; WT
845 was positive. **(f)** RT-PCR revealed that BleoR CDS was adequately inserted in the target region
846 with amplifying 413 bp of transcribed RNA in the recovered events. The *β-actin* was used as the
847 control in all RT-PCR assays; BleoR protein extracted from pDDT plasmid was used as the
848 positive control. WT was used as the negative control. **(g)** Schematic of probing BleoR in edited
849 events and WT as the control using Southern blotting. **(h)** Southern blot proved that *BleoR* CDS

was integrated into the precise recombinant genome. Digested pDDT plasmid was used as the positive control.

**Figure 4:** The 2D kernel density plot of TaqMan real-time PCR fluorescent intensities and HDR efficiency percentage. **(a)** The TaqMan real-time PCR assay designing to detect HDR happened, and evaluation included FAM1 and FAM2 DNA binding probes. **(b)** Strategy to classify edited events. **(c)** Experiment I revealed no density for the edited events. **(d)** The density plot of FAM1 and -2 intensities resulted from experiment II revealed an expansion in edited events against partial, mutant, and wild-types. **(e)** The density plot of FAM1 and -2 ΔΔCt resulted from experiment III revealed an increased intensity of partial FAM1 events. **(f)** Experiment IV revealed a remarkable increase of edited events signals in confronting with three earlier experiments. **(g)** The Density plot of experiment V revealed a significant increase of FAM1 and -2 intensities in edited events compared to the earlier experiments and a significant decrease in intensities in WT and mutated events. All samples were analyzed in quadruplicate. **(h)** Diamond box and whisker plot revealed the identification of all FAM signals visualized in the experiments and showed more signals remarkably measured in ExV than ExI, II, and-III; Error bars represent SE; Asterisks represent p-value as *≤0.05, **≤0.01, and ***≤0.001. **(i)** The bar plot represents the HDR efficiency in different experiments; The overlap data are shown as bin bars, and the standard distribution curves are added. HDR efficiency plot revealed that *XRCC4* deficiency (ExV) led to HDR happening significantly more than the fusion of CtIP (ExII), MRE11 (ExIII), and CtIP+MRE11 (ExV). Also, ExIV meaningfully revealed more HDR happening than ExII and -III.; Error bars represent SE; Asterisks represent p-value as **≤0.01, ***≤0.001, and ****≤0.0001; Triplicate technical repeats were considered for each sample.

**Figure 5:** Violin plots reveal the *BleoR* and *MKK2* expression and the success happening HDR via different experiments. **(a-e)** The differences between *BleoR* and *MKK2* expression. Three technical repeats were used for each event in this assay; Dash lines present quartiles; Solid lines present median. **(f)** Heat-map to show the effect of efficient HDR on the expression of NHEJ and HDR factors. Overexpression *CtIP* and/or *MRE11* caused to enhance the expression of *BRCA1*, *Rad50*, and *Rad51* and to demote the expression of *Lig4* and *XRCC4*. The highest expression of the HDR factors visualized in ExV means that *XRCC4* deficiency decreased the

expression of NHEJ factor *Lig4* and intensified HDR efficiency. Triplicate technical repeats were considered for each sample.

**Figure 6:** Polymorphisms analysis. **(a)** Analyses of distributed indel nucleotides happened on 5' and 3' homology arms and knocked in fragments throughout experiment events. **(b)** Diamond box and whiskers for the mean comparisons of happened indel nucleotides through experiment events. The exact numbers of indels are presented via the punching column bars on the top-right corner. **(c)** Identification of the happened polymorphisms in homology arms through the experiments. Box and Whisker plot revealed that most polymorphisms happened in homology arms by ExI, and it was significantly more than those in ExV and –IV; Asterisks represent p-value as *≤0.05; Error bars represent SE. **(d)** Stacked column plot of total polymorphisms happened in DDT integration into the poplar genome. Insertions and deletions were occurred much more than the other types. SNP and substitutions were occurred less than the other types. Whisker and standard distribution curves exposed that the total polymorphisms caused by *XRCC4* deficiency were less than the other experiments.

**Figure 7:** Radar diagrams of *MKK2* expressions, stem lengths, and -diameters from WT and survived recovered events after NaCl treatment. No significant differences in *MKK2* expressions and phenotypic changes before and after salt stress between WT and survived recovered events confirmed the proper HDR.

## Supplementary data

**Supplementary 1**: Schematic of pgRNA, DDT, and pDDT. **(a)** pgRNA included the *MKK2* target seed and Cas9. **(b)** pDDT included DDT ligated into pRGEB31 by restriction enzyme cloning method.

**Supplementary 2**: Schematic construction of DDT and pDDT fragments, primers, and oligos.

**Supplementary 3:** Schematic construction of CtIP, MRE11, CtIP+MRE11, and XRCC4 cassette primers and oligos.

**Supplementary 4:** Schematic construction of cassettes (CtIP and MR) and vectors (pgCtIP and pgMR) and their primers.

907 **Supplementary 5 :** Schematics of constructed cassettes and plasmids. **(a)** pgCtIP plasmid
908 including CtIP cassette. **(b)** pgMR plasmid including MR cassette. **(c)** pgCtMR plasmid including
909 CtMR cassette. **(d)** pggCtMR plasmid including XRCC4 cassette.

910 **Supplementary 6:** Schematic construction of cassettes (CtMR, XRCC4) and vectors
911 (pgCtMR and pggCtMR) and their primers.

912 **Supplementary 7:** Alignment of events involved in experiment I.

913 **Supplementary 8:** Alignment of events involved in experiment II.

914 **Supplementary 9:** Alignment of events involved in experiment III.

915 **Supplementary 10:** Alignment of events involved in experiment IV.

916 **Supplementary 11:** Alignment of events involved in experiment V.

917 **Supplementary 12:** Schematic of TaqMan real-time PCR FAM and VIC target assays in
918 this study. Yellow rectangles exhibited CDS.

919 **Supplementary 13:** Box-and-whisker (Min-Max) plots of one-dimensional FAM delta-
920 delta Ct signals in designed experiments. All signals were calculated as quadruplicates.

921 **Supplementary 14:** Schematics of sequence analyzing of triggered events from different
922 experiments. **(a)** Sequence analysis of triggered events included in EXII reveals one recovered
923 event. **(b)** Sequence analysis of triggered events included in EXIII reveals one recovered event.
924 **(c)** Sequence analysis of triggered events included in EXIV reveals four recovered events. **(d)**
925 Sequence analysis of triggered events included in EXV reveals 12 recovered events.

926 **Supplementary 15:** The raw data of real-time PCR evaluates the percentage of delta-
927 delta Ct mean from *BleoR* in all experiments. **(a)** Delta-delta Ct mean of *BleoR* expression from
928 grown buds. Each sample was investigated with three technical repeats. **(b)** Descriptive statistic
929 table of raw data calculated by ANOVA-One way.

930 **Supplementary 16:** Column plots of the expression of integrated *BleoR* and new
931 recombinant *MKK2* genes via different designed experiments. Three technical repeats were
932 used for each event in this assay; Error bars represent SD; WT and pDDT were used to
933 control *MKK2* and *BleoR* expression, respectively.

934 **Supplementary 17:** Chromatogram alignments of events included in experiment V.

## Supplementary Table

**Supplementary Table 1:** CRISPR sites located on 3' region of *MKK2*. The yellow highlight reveals the selected CRISPR target in this study.

**Supplementary Table 2:** Oligos and primers used in this study.

**Supplementary Table 3:** All polymorphisms detected in homology arms happened by HDR through experiments.

**Supplementary Table 4:** Variant nucleotides happened from experiments.

**Supplementary Table 5**: The raw data of RT-qPCR of *MKK2* expressions and phenotypic analyzes in WT and survived recovered events before and after salt stress. All events were analyzed in triplicates for RT-qPCR.

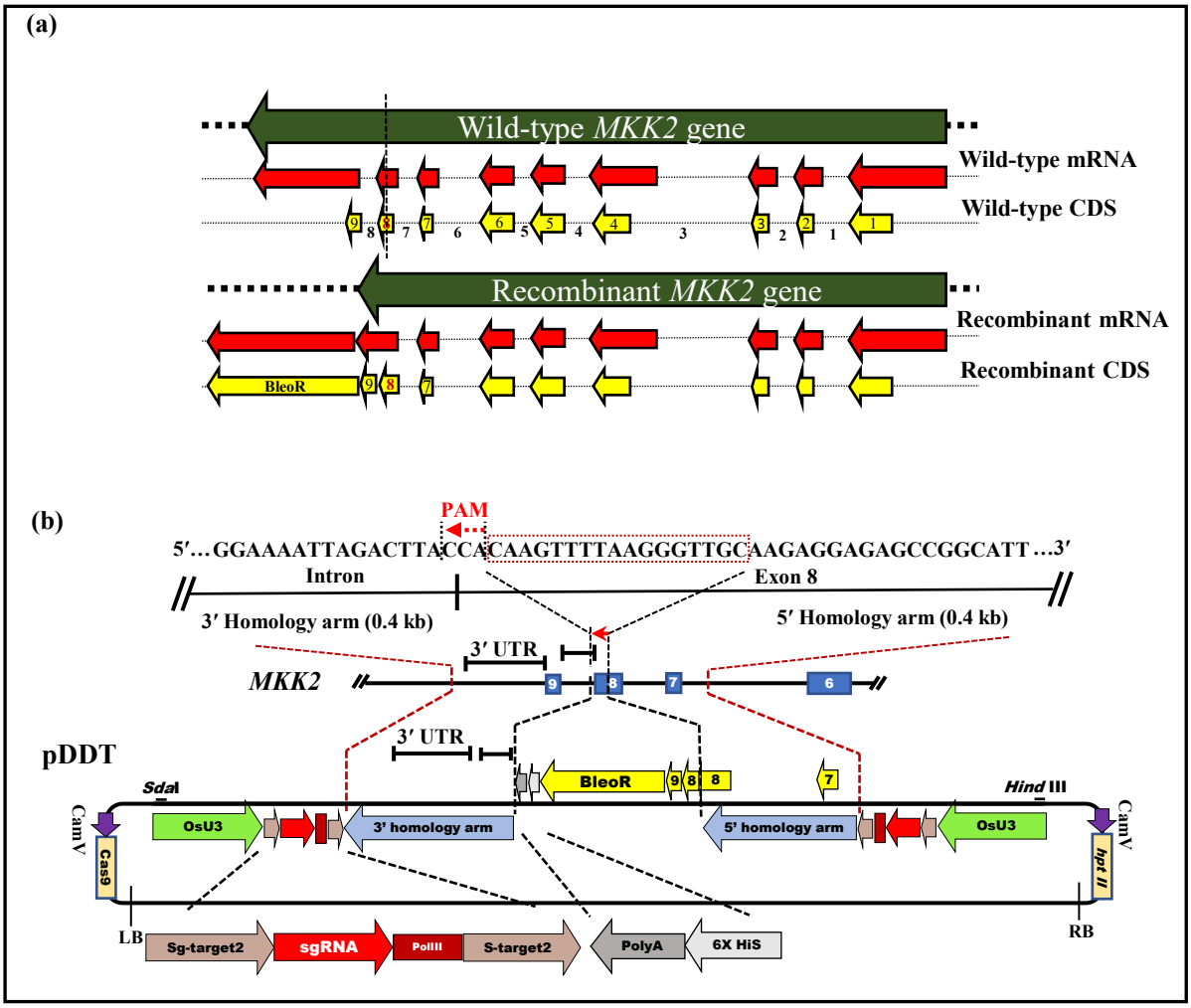
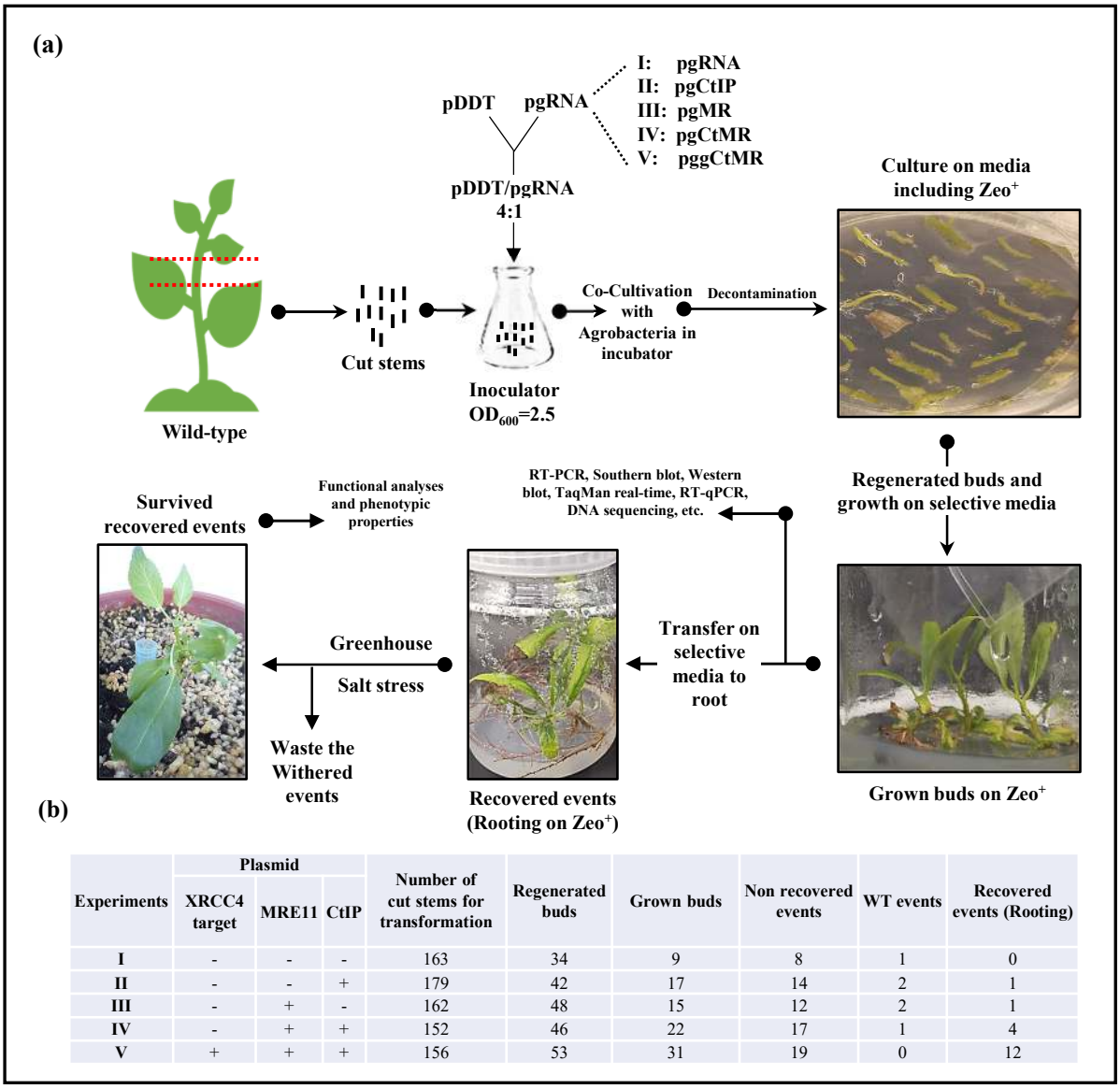



**Figure 1:** Schematics of *MKK2* locus before and after edition and integration of exogenous *BleoR* CDS into the poplar genome **(a)** Schematic of the purpose of this research to integrate exogenous BleoR into the poplar genome. Dash line reveals the target site. **(b)** Protospacer Adjacent Motif (PAM) was detected at the end of exon 8 to lead Cas9. 400 bp sequences from both sides of the CRISPR target were selected for HDR in this study. The 5' homology arm included part sequences of the intron between exon 6 and -7, exon 7, intron sequences between exon 7 and -8, and a part of exon 8. The 3' homology arm included intron sequences between exon 8 and -9 and 3' UTR of the *MKK2* locus up to 400 bp. Designed DDT included remained sequences of exon 8, exon 9, BleoR CDS, 6xHis, and PolyA sequences flanked by the 3'- and 5' homology arms. We added two special targets besides DDT. The DDT was then ligated into the pRGEB31 vector to form pDDT.




| Experiments | Plasmid | | | Number of cut stems for transformation | Regenerated buds | Grown buds | Non recovered events | WT events | Recovered events (Rooting) |
|---|---|---|---|---|---|---|---|---|---|
| | XRCC4 target | MRE11 | CtIP | | | | | | |
| I | - | - | - | 163 | 34 | 9 | 8 | 1 | 0 |
| II | - | - | + | 179 | 42 | 17 | 14 | 2 | 1 |
| III | - | + | - | 162 | 48 | 15 | 12 | 2 | 1 |
| IV | - | + | + | 152 | 46 | 22 | 17 | 1 | 4 |
| V | + | + | + | 156 | 53 | 31 | 19 | 0 | 12 |

**Figure 2:** The transformation strategy and designing of experiments. **(a)** pDDT and pgRNA were mixed 4:1 and introduced to the Agrobacterium tumefaciens to form inoculator suspension. We condensed inoculator up to OD600=2.5 and then dipped all cut off stems. The putatively edited events were regenerated on Zeocin. We allowed putative edited events to bud. The grown buds were then transferred on selective rooting media and allowed to be recovered. Recovered events were then planted on soil and following two weeks of acclimation introduced by salt stress **(b)** Designed experiments for this study including (I) No HDR factors, (II) overloaded CtIP, (III) overloaded MRE11, (IV) Overloaded CtIP+MRE11, and (V) Overloaded CtIP+MRE11 with *XRCC4* deficiency.

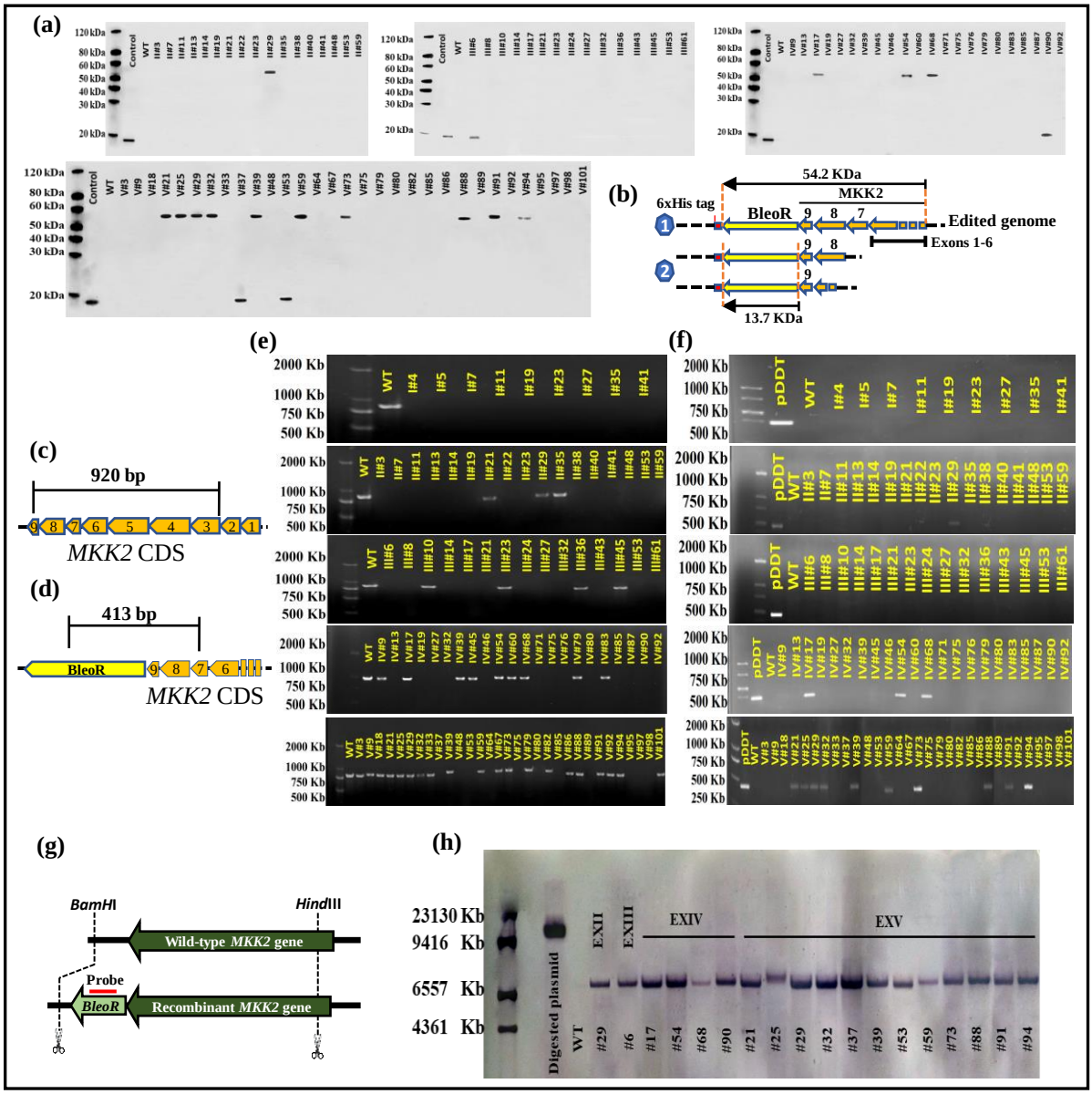



**Figure 3:** Western blotting to reveal the fused 6xHis tag with BleoR integration into the poplar genome. **(a)** Different experiments exhibited different quantities of 6xHis tag fusion. **(b)** Schematic of fusion 6xHis tag with edited poplar genome triggered by different experiments. Shape 1 reveals successful fusion of BleoR and MKK2 with about 54 kDa. Shape 2 reveals an unsuccessful combination of mentioned proteins with about 14 kDa. **(c)** Schematic of right HDR happening caused to attach exon 8 and 9 in the edited genome. **(d)** Schematic of proper integration in edited genome caused to connect the BleoR to the C-terminal of MKK2. **(e)** RT-PCR exhibited the HDR in exon 8 and 9, revealing a 920 bp of transcribed MKK2 RNA in triggered events from ExII to ExV. The β-actin was used as the control in all RT-PCR assays; WT was positive. **(f)** RT-PCR revealed that BleoR CDS was adequately inserted in the target region with amplifying 413 bp of transcribed RNA in the recovered events. The β-actin was used as the control in all RT-PCR assays; BleoR protein extracted from pDDT plasmid was used as the positive control. WT was used as the negative control. **(g)** Schematic of probing BleoR in edited events and WT as the control using Southern blotting. **(h)** Southern blot proved that BleoR CDS was integrated into the precise recombinant genome. Digested pDDT plasmid was used as the positive control.

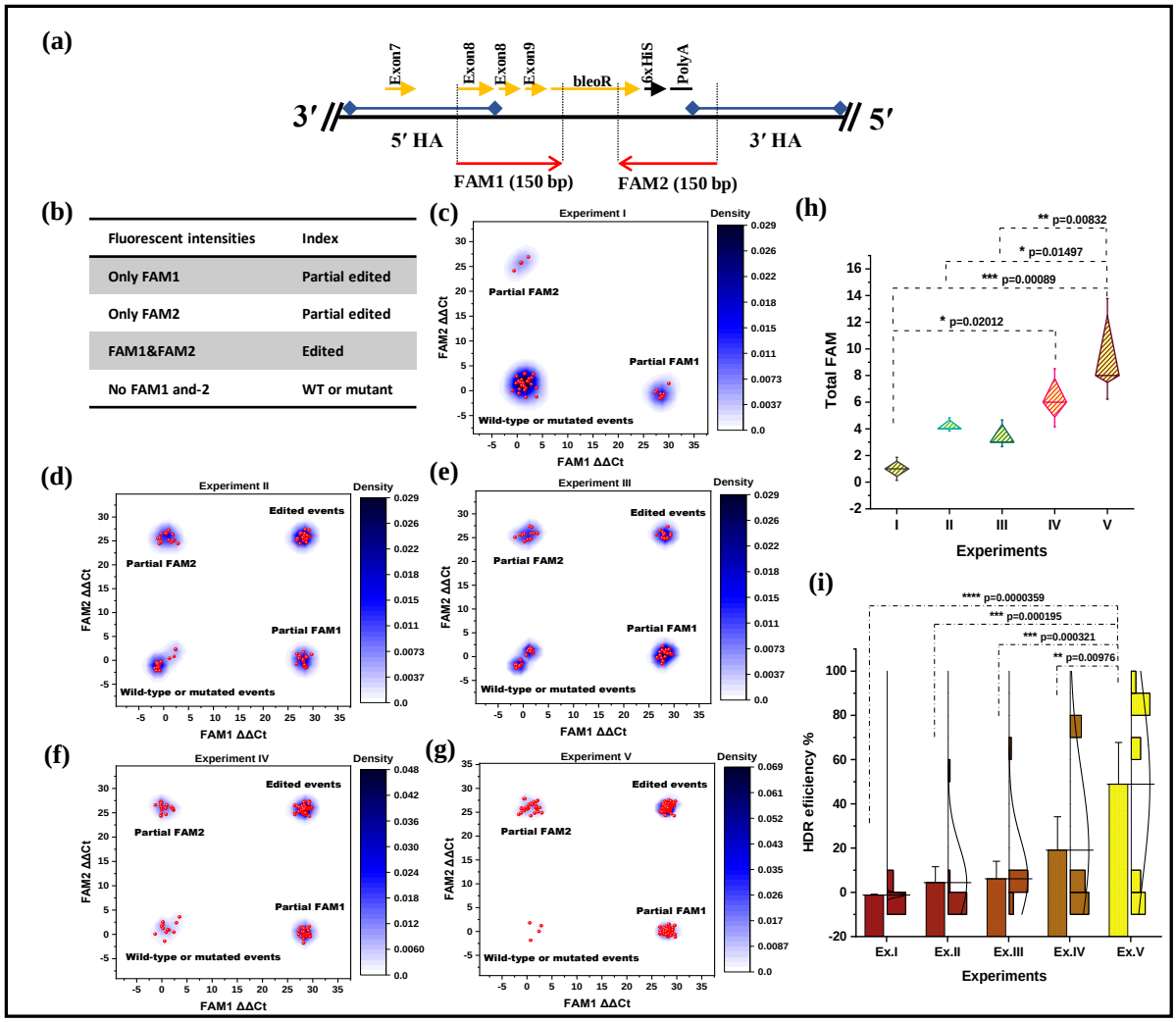



**Figure 4:** The 2D kernel density plot of TaqMan real-time PCR fluorescent intensities and HDR efficiency percentage. **(a)** The TaqMan real-time PCR assay designing to detect HDR happened, and evaluation included FAM1 and FAM2 DNA binding probes. **(b)** Strategy to classify edited events. **(c)** Experiment I revealed no density for the edited events. **(d)** The density plot of FAM1 and -2 intensities resulted from experiment II revealed an expansion in edited events against partial, mutant, and wild-types. **(e)** The density plot of FAM1 and -2 ΔΔCt resulted from experiment III revealed an increased intensity of partial FAM1 events. **(f)** Experiment IV revealed a remarkable increase of edited events signals in confronting with three earlier experiments. **(g)** The Density plot of experiment V revealed a significant increase of FAM1 and -2 intensities in edited events compared to the earlier experiments and a significant decrease in intensities in WT and mutated events. All samples were analyzed in quadruplicate. **(h)** Diamond box and whisker plot revealed the identification of all FAM signals visualized in the experiments and showed more signals remarkably measured in ExV than ExI, II, and-III; Error bars represent SE; Asterisks represent p-value as *≤0.05, **≤0.01, and ***≤0.001. **(i)** The bar plot represents the HDR efficiency in different experiments; The overlap data are shown as bin bars, and the standard distribution curves are added. HDR efficiency plot revealed that *XRCC4* deficiency (ExV) led to HDR happening significantly more than the fusion of CtIP (ExII), MRE11 (ExIII), and CtIP+MRE11 (ExV). Also, ExIV meaningfully revealed more HDR happening than ExII and -III.; Error bars represent SE; Asterisks represent p-value as **≤0.01, ***≤0.001, and ****≤0.0001; Triplicate technical repeats were considered for each sample.

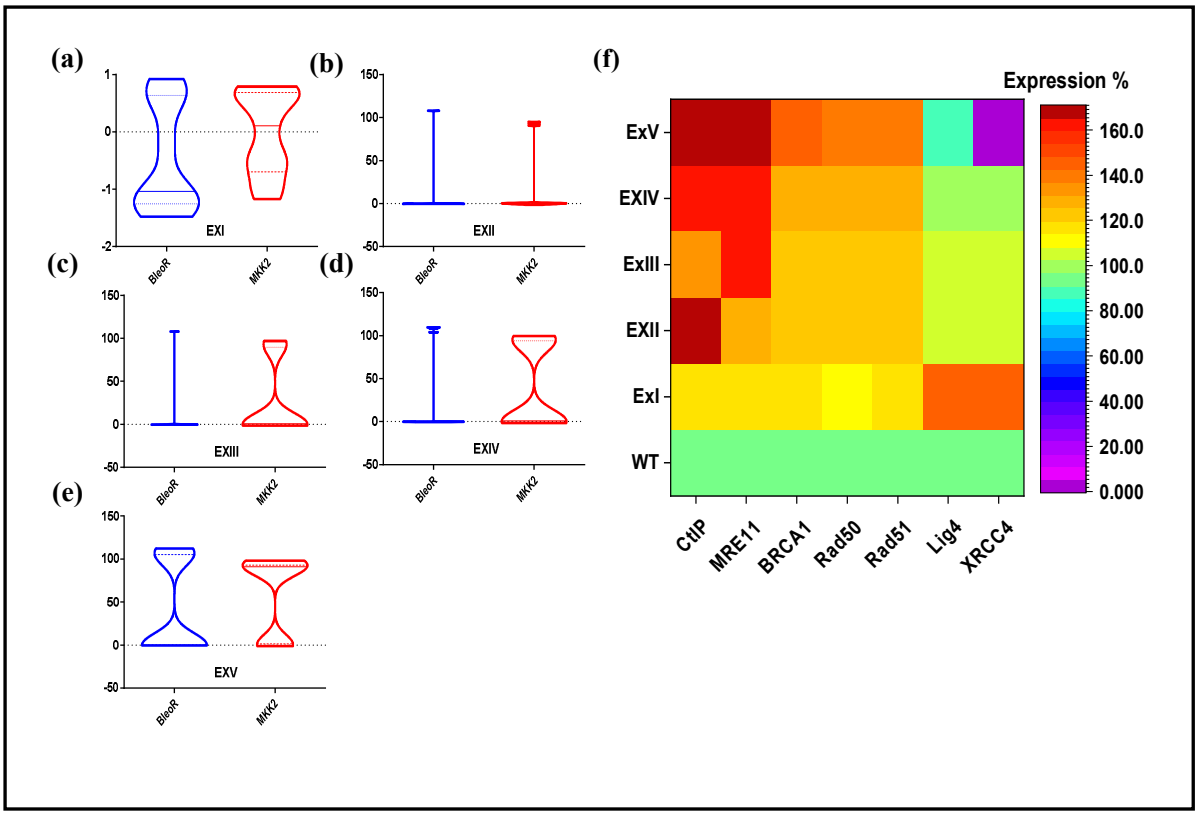



**Figure 5:** Violin plots reveal the *BleoR* and *MKK2* expression and the success happening HDR via different experiments. **(a-e)** The differences between *BleoR* and *MKK2* expression. Three technical repeats were used for each event in this assay; Dash lines present quartiles; Solid lines present median. **(f)** Heat-map to show the effect of efficient HDR on the expression of NHEJ and HDR factors. Overexpression *CtIP* and/or *MRE11* caused to enhance the expression of *BRCA1*, *Rad50*, and *Rad51* and to demote the expression of *Lig4* and *XRCC4*. The highest expression of the HDR factors visualized in ExV means that *XRCC4* deficiency decreased the expression of NHEJ factor *Lig4* and intensified HDR efficiency. Triplicate technical repeats were considered for each sample.

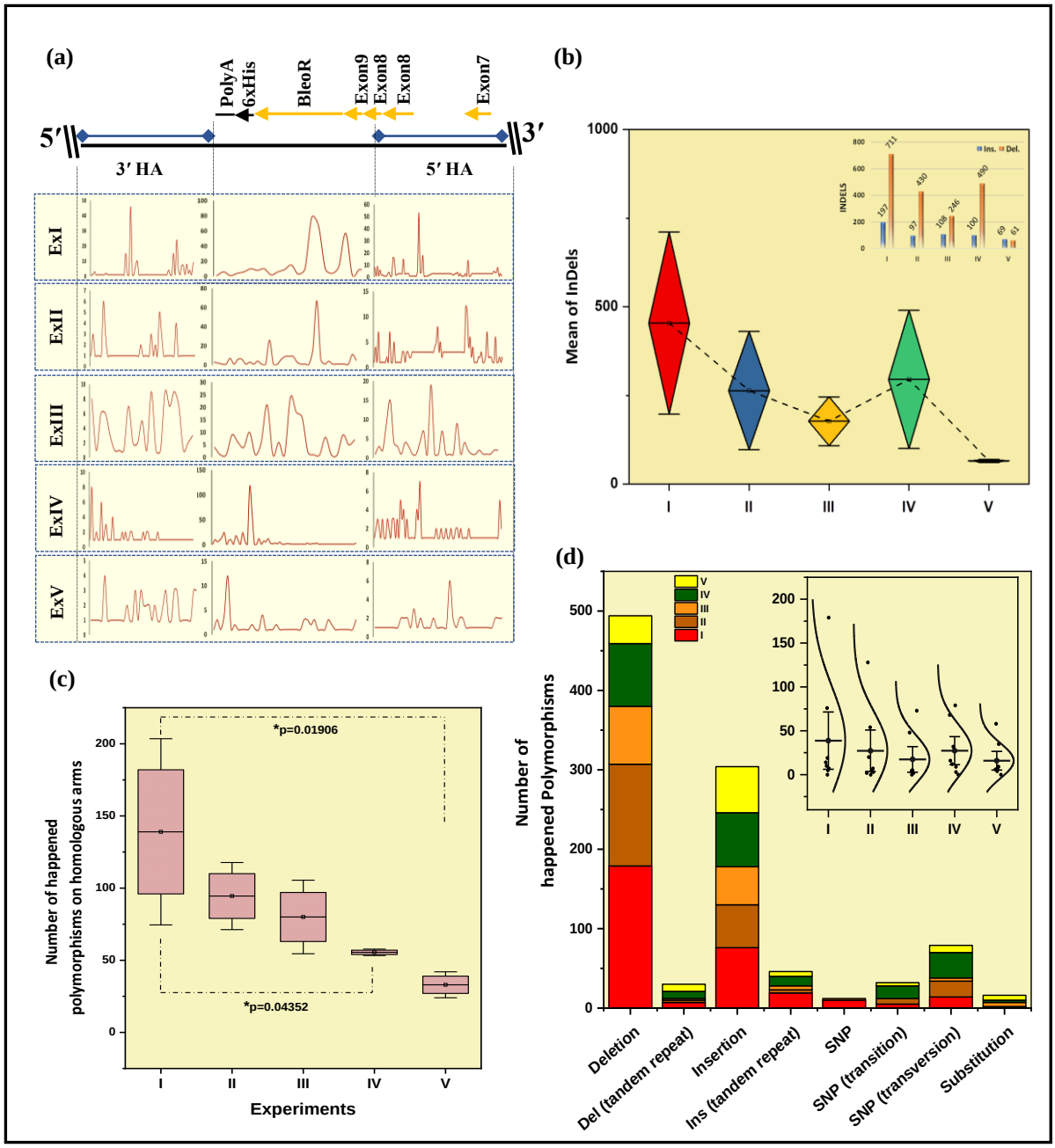



**Figure 6:** Polymorphisms analysis. **(a)** Analyses of distributed indel nucleotides happened on 5' and 3' homology arms and knocked in fragments throughout experiment events. **(b)** Diamond box and whiskers for the mean comparisons of happened indel nucleotides through experiment events. The exact numbers of indels are presented via the punching column bars on the top-right corner. **(c)** Identification of the happened polymorphisms in homology arms through the experiments. Box and Whisker plot revealed that most polymorphisms happened in homology arms by ExI, and it was significantly more than those in ExV and –IV; Asterisks represent p-value as *≤0.05; Error bars represent SE. **(d)** Stacked column plot of total polymorphisms happened in DDT integration into the poplar genome. Insertions and deletions were occurred much more than the other types. SNP and substitutions were occurred less than the other types. Whisker and standard distribution curves exposed that the total polymorphisms caused by XRCC4 deficiency were less than the other experiments.

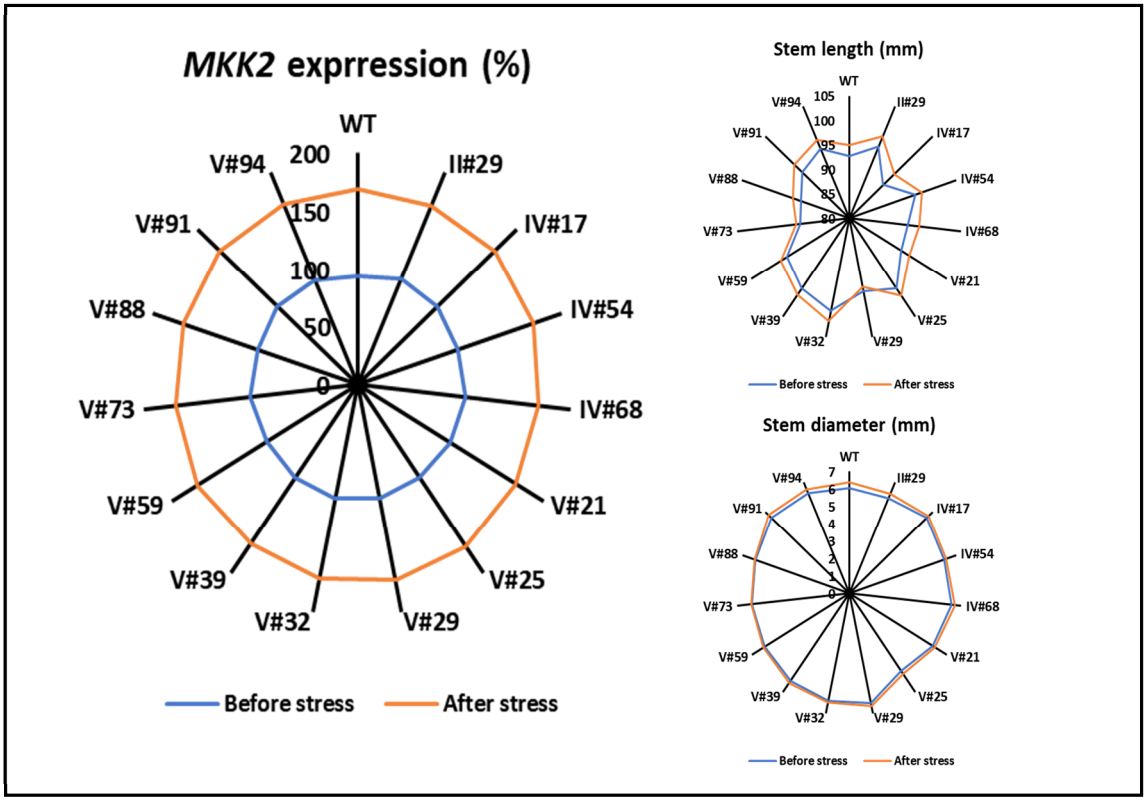



**Figure 7**: Radar diagrams of *MKK2* expressions, stem lengths, and -diameters from WT and survived recovered events after NaCl treatment. No significant differences in *MKK2* expressions and phenotypic changes before and after salt stress between WT and survived recovered events confirmed the proper HDR.