MSH2 and *MSH6* in mismatch repair system account for soybean (*Glycine max* (L.) Merr.) tolerance to cadmium toxicity by determining DNA damage response

Qiang Zhao[a], Hetong Wang[c, 1], Yanli Du[a], Hilary J. Rogers[d], Zhixin Wu[a], Sen Jia[a], Xingdong Yao[a], Futi Xie[a, *], Wan Liu[b, **]

[a] Agricultural College, Shenyang Agricultural University, Shenyang 110866, PR China

[b] Key Laboratory of Pollution Ecology and Environmental Engineering, Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang 110016, PR China

[c] College of Life Science and Bioengineering, Shenyang University, Shenyang 110044, PR China

[d] Cardiff University, School of Biosciences, Cardiff CF10 3TL, UK

[1] Co-first author

* Corresponding author: xft299@syau.edu.cn (F.X.); Tel.: +86-24-8848-7135 (F.X.); **Correspondence: liuwan@iae.ac.cn (W.L.); Tel.: +86-24-8397-0315 (W.L.).

Abstract

Our aim was to investigate DNA mismatch repair (MMR) genes regulating cadmium tolerance in two soybean cultivars. Cultivars Liaodou 10 (LD10, Cd-sensitive) and Shennong 20 (SN20, Cd-tolerant) seedlings were grown hydroponically on Murashige & Skoog (MS) media containing 0-2.5 mg·L^{-1} Cd for 4 days. Cd stress induced less random amplified polymorphism DNA (RAPD) polymorphism in LD10 than in SN20 roots, causing G1/S arrest in LD10 and G2/M arrest in SN20 roots. Virus-induced gene silencing (VIGS) of *MLH1* in LD10-TRV-*MLH1* plantlets showed markedly diminished G1/S arrest, but enhanced root length/area under Cd stress. However, an increase in G1/S arrest and reduction of G2/M arrest occurred in SN20-TRV-*MSH2* and SN20-TRV-*MSH6* plantlets with decreased root length/area under Cd stress. Taken together, we conclude that low expression of *MSH2* and *MSH6*, involved in the G2/M arrest, results in Cd-induced DNA damage recognition bypassing the MMR system to activate G1/S arrest with the assistance of MLH1. This then leads to repressed root growth in LD10, explaining the inter-varietal difference in Cd tolerance in soybean.

Keywords: Cd toxicology, Cell cycle arrest, DNA damage, DNA mismatch repair, Root growth repression, Soybean (*Glycine max* (L.) Merr.)

Introduction

Cadmium (Cd) is one of the most toxic heavy metal contaminants. A large amount of Cd has been released into ecosystems mostly through anthropogenic activities, such as lead-zinc mining, nonferrous metal smelting and phosphate fertilizer utilization.[1] Due to its long half-life of 18-30 years in biota, Cd can persist in ecosystems for a long time. Since Cd is readily absorbed and accumulated in organisms, Cd bioaccumulation and biomagnification throughout the food chain induce widespread genetic toxicity or cytotoxicity in cells. [2,3] It is well known that Cd stress can directly induce a wide range of injury symptoms in plants, such as the inhibition of photosynthesis, causing oxidative stress and cell cycle modulation or apoptosis.[4] Therefore, research into the molecular mechanisms of Cd stress in plants is an important topic in environmental and agricultural science.

It has been shown that Cd stress can directly interact with the hydrogen bonds in the bases and base pairs of DNA, leading to a variety of reversible and/or irreversible DNA lesions in plants, such as base-base mismatches, insertion/deletion loops, DNA adducts, DNA chain cross linking and breaks.[5-7] Cd stress can induce the production of oxygen radicals and regulate gene expression through changes in the DNA structure or a destruction of the DNA repair system, which indirectly results in DNA damage.[8] DNA damage is sensed and repaired through a series of signal transduction pathways which are known as DNA damage response (DDR). These maintain high fidelity of genetic

information, and the main DNA damage repair mechanisms include base excision repair (BER), nucleotide excision repair (NER), mismatch repair, (MMR), non-homologous end joining (NHEJ) and homologous recombination (HR) in eukaryotes. Damaged DNA is recognized by cell cycle checkpoints, then cell cycle progression is slowed down or arrested to provide the cells with sufficient time to repair DNA damage or undergo cell death.[9] There are three checkpoints in the eukaryotic cell cycle: the G1/S phase checkpoint preventing the damaged DNA or mutant cells entering into S phase; the S phase checkpoint arresting the replication of damaged DNA; and the G2/M checkpoint arresting the cells with damaged DNA from entering mitosis. Cell cycle checkpoint control is a complex molecular mechanism involving multiple signaling pathways.[10] ATM and ATR kinases are sensors for various types of DNA damage, and activate signal transduction pathways regulating the DNA damage checkpoints.[11] For example, in animal cells, ATM/ATR activates the phosphorylation of p53-P21 proteins, which inhibit the activities of CDK2 and CDK4, and participate in the G1/S or G2/M arrest in response to DNA damage. Furthermore, in animal cells G2/M arrest induced by DNA damage is regulated by ATM/ATR-Chk2/Chk1 signaling pathways, including the protein activities of cell division cycle (Cdc25), WEE1, CDK1, BRCA1, RAD51 and Cyclin B. [12,13]

The MMR system is a key DNA repair pathway, and is involved in a wide range of important cellular processes such as sensing and correcting DNA damage, governing cell cycle progression, confirming fidelity of DNA replication,

and maintaining genomic stability in the presence of structurally anomalous nucleotide lesions under different stresses.[14,15] In plants, the functions of the MMR system are through a complex interaction among MutS and MultL protein families.[16,17] The MMR proteins MSH1-MSH7 are from the MutS family, while MLH1, MLH3, PMS1 and PMS2 belong to the MutL family.[18,19] DNA errors involving base-base mismatches and single (1-2 bases) insertion/deletion loops are recognized by heterodimer complexes known as MutSα (MSH2–MSH6) and MutSγ (MSH2-MSH7), whereas 2-12 base insertion/deletion loops are recognized by MutSβ (MSH2–MSH3).[20,21] Also, the protein complex known as MutLs (MutLα and MutLβ) participate in MMR progression.[22] In fact, the MutLα (MLH1-PSM1) can bind to MutSα or MutSβ to deliver the DNA damage signal through PCNA and/or RFC pathway, and activate the MMR reaction to repair DNA base mismatch damage.[23,24] In mammals, DNA damage recognized by MMR proteins can activate the G2/M cell cycle checkpoint. For example, hMSH2 and hMLH1 modulate G2/M phase arrest by activating the hMSH2/hMLH1-BRCA1-ATR-CHk1 pathway in the HCC1937 human cancer cell line under 6-Mercaptopurine (6-TG) stress.[25] hMLH1 is necessary for activating the ATM-dependent DNA damage response in the HCT116 human cancer cell line under selenium stress.[26] In plants, MSH2 had been shown to play an important role in regulation of cell cycle progression in Arabidopsis seedlings after UV-B treatment,[16] and MSH2 and MSH6 mediated Cd-induced G2/M checkpoint arrest through the MutSα-ATR-WEE1 pathway in Arabidopsis

seedling roots.[27]

Previous research showed that Cd stress-induced G1/S and G2/M phase arrest was linked with DNA damage and decreased level of cyclin B1 mRNA in suspension culture soybean cells.[28] DNA damage tolerance determines whether cells maintain the complete DNA synthesis process to sustain plant growth or enter the cell death process,[29] and this may play an important role in soybean Cd stress tolerance. However, little is known about the roles of MMR proteins in Cd-induced cell cycle arrest and Cd tolerance in soybean seedlings.

In this study, two soybean cultivars, LD10 and SN20, with contrasting Cd sensitivity were used to (1) determine the levels of DNA damage in soybean seedling root tips under Cd stress by RAPD analysis; (2) measure cell cycle progression in response to Cd stress in soybean seedling roots by flow cytometry method (FCM) and qRT-PCR analyses; and (3) evaluate the potential roles of MMR genes in Cd-induced cell cycle arrest and Cd tolerance in seedling roots of soybean in which virus-induced gene silencing (VIGS) was used to silence three MMR genes: TRV-*MLH1*, TRV-*MSH2* and TRV-*MSH6*.

Materials and methods

Materials, growth and treatment conditions

Soybean (*Glycine max* L.) Merr.) seeds used in this study were harvested on Oct. 3, 2018 from the experimental station of the Soybean Institute (41°82′N, 123°57′E), College of Agriculture, Shenyang Agricultural University, Liaoning, PR China (Table S1). The pods were dried naturally and stored at 4 ℃.

Soybean seeds were sterilized using chlorine gas (made by mixing 4 mL 12 M HCl and 100 mL 5.25% hypochlorite) in a glass desiccator for 8-10 h. To investigate the effect of Cd stress on soybean seed germination, sterilized seeds were sown in a 90 mm culture dish onto gauze saturated with a Cd solution (0, 0.25, 0.5, 2.5 $mg \cdot L^{-1}$ $CdCl_2 \cdot 2H_2O$). The seed germination efficiency was measured at 28 ± 1 ℃ in darkness after 2 days.

To avoid the effects from heterogeneity in germination amongst seeds from the same batch, sterilized seeds were germinated on gauze soaked in distilled water and checked for uniformity of germination. The seeds were kept for approximately 2 days at 28 ± 1 ℃ in darkness, until the hypocotyls were 1-1.5 cm. The uniformly germinated soybean seeds were selected and transferred into Murashige & Skoog (MS, Caisson, USA) liquid medium with different Cd concentrations of 0 (control), 0.25, 0.5, 2.5 $mg \cdot L^{-1}$ in the form of $CdCl_2 \cdot 2H_2O$ of analytical grade with purity 99.5% (PR China), and incubated for 4 days at 28 ± 1 ℃ with a light regime of 16 h light / 8 h dark. The Cd solution was changed every other day. Before harvesting, the roots were rinsed three times with sterile water and scanned using a WinRHIZO Pro 2012b root scanning image analysis system (Regent Instruments, Inc., Quebec, Canada) to measure total root length/area. Root length reduction (%) = (root length of the control seedling - root length of Cd treated seedling) / root length of the control seedling × 100%. The fresh weight of soybean seedling was quickly measured, and then about 1 cm long root tips were cut and flash-frozen in liquid nitrogen prior to storage at

153 −80 ℃. Soybean seedlings were oven-dried at 105 °C for 30 min and then at
154 85 °C until a constant weight was achieved. All treatments and analyses were
155 repeated in at least three independent replicates.

156 **DNA extraction and RAPD analysis**

157 Total genomic DNA was extracted and purified using a Plant Genomic DNA
158 Isolation Kit (Tiangen, Beijing, PR China) from about 100 mg of fresh root tips
159 frozen at −80 ℃. The RAPD analysis was performed using 2 primers (primers
160 2 and 6) screened from 12 random primers as described previously (Table
161 S2).[30] Following PCR amplification, polymorphism frequency of RAPDs, was
162 assessed by polyacrylamide gel electrophoresis (PAGE) gel electrophoresis,
163 and was calculated according to Wang et al.[7] The genome template stability
164 (GTS) was calculated using the equation: GTS= (1- a/n) × 100%, where a and
165 n represent the average frequency of RAPDs polymorphism in Cd treated and
166 control roots, respectively. For all treatments, bands were considered
167 reproducible, and were used for polymorphism analysis when detected
168 simultaneously in at least two experimental replicates.

169 **FCM analysis of cell cycle progression in soybean seeding root tips**

170 Nuclei were extracted using chopping buffer[7] from approximately 0.1 g of
171 fresh soybean seedling root tips (about 1 cm long). The root tips were chopped
172 into 0.5 mm strips using a single-edged razorblade in a glass Petri dish
173 (diameter, 5 cm) with 2 mL ice cold chopping buffer. After 5 minutes in an ice
174 bath, the mixture was filtered through a 30 μm nylon mesh twice to remove cell

175 debris. 1 mL of mixture was transferred into a 1.5 mL centrifuge tube, and
176 incubated with 15 μg·L^{-1} RNase A in a water bath at 37 ℃ for 2 h. The mixture
177 was stained with 50 μg·L^{-1} propidium iodide (PI, Beyotime, PR China) at 4 ℃
178 for 0.5-1 h. The ploidy level of the control and Cd-treated samples was analyzed
179 using a Guava easyCyte 6-2 L flow cytometer (EMD Millipore, USA) equipped
180 with a 488 nm laser. Fluorescence intensity was analyzed in the Red-B-HLin
181 channel with more than 5000 nuclei measured for each sample, and three
182 independent replicates were performed for each sample. Gates (Figures S3
183 and S9) were determined empirically and ploidy distribution was analyzed using
184 Flowjo 7.6.1 win 64 software (BD Biosciences, San Jose, CA).

185 **RNA extraction, first-strand cDNA synthesis and qRT-PCR analysis**

186 Total RNA was isolated and purified using a Plant Total RNA Isolation Kit
187 (Qiagen, Hilden, Germany) from about 100 mg of fresh root tips frozen at −80 ℃
188 according to the manufacturer's manual. First-strand cDNA was synthesized
189 from 1 μg of total RNA using a TransScript® All-in-One First-Strand cDNA
190 Synthesis SuperMix (TransScript, Beijing, PR China) in a final volume of 20 μL,
191 and stored at -20 ℃.

192 The reaction mixture (1 μL) was used for qRT-PCR in a 20 μL reaction
193 volume using TransScript® Top Green qPCR SuperMix (TransScript, Beijing,
194 PR China). The soybean *Tubulin A*（NM_001250372）or *Actin* (NM_001289231)
195 gene was used for signal normalization. The primers used for amplifying
196 specific genes were designed using the online QuantPrime software

(http://quantprime.mpimp-golm.mpg.de/) and are listed in Table S2. The qRT-PCR products were confirmed as the correct amplification products by analysis on 2% (w/v) agarose gels and sequencing. The operational formula $2^{-\triangle\triangle Ct}$ was used to calculate relative expression levels of the selected genes between different treatments.[31] The qRT-PCR experiments and analyses were performed with three biological replications, and each biological replication was measured in three technical replications.

Construction of VIGS-induced gene silencing plasmids

TRV1 and TRV2 plasmids were used to produce amiRNAs (artificial miRNAs) for gene silencing via VIGS technology.[32] All the constructs used for VIGS-induced gene silencing were assembled into the TRV2 plasmid. Gene fragments of *PDS* (XM_028355994), *MLH1* (XM_003522549), *MSH2* (XM_003549757) and *MSH6* (XM_006604676) were amplified by PCR from cDNA of LD10 and SN20 leaves. The specific primers used in PCR were designed with BamH Ⅰ and Xho Ⅰ restriction sites in the forward and reverse primers, respectively (listed in Table S2). The sizes of PCR products were confirmed by 1.5% agarose gel electrophoresis, and then validated by sequencing. The validated fragments were inserted into the TRV2 plasmid between the BamH Ⅰ and Xho Ⅰ restriction sites to construct the -TRV-*PDS*, -TRV-*MLH1*, -TRV-*MSH2* and -TRV-*MSH6* VIGS-induced gene silencing plasmids (as shown in Figure S4).

Soybean sprout vacuum-infiltration for VIGS

A sprout vacuum-infiltration method[33] was used to develop VIGS-induced gene silencing lines of LD10 including LD10-TRV2, LD10-TRV-*PDS*, LD10-TRV-*MLH1*, LD10-TRV-*MSH2* and LD10-TRV-*MSH6*, and of SN20 including SN20-TRV2, SN20-TRV-*PDS*, SN20-TRV-*MLH1*, SN20-TRV-*MSH2* and SN20-TRV-*MSH6*. For VIGS research, plasmids of TRV1, TRV2, and TRV2 construction derivatives (TRV-*PDS*, TRV-*MLH1*, TRV-*MSH2* and TRV-*MSH6*) were transformed into competent *Agrobacterium tumefaciens* strain GV3101 cells using a freeze–thaw method.[34] A single colony for each transformation was selected and confirmed by colony PCR (primers listed in Table S2). The verified bacterial cells were inoculated into 4 mL of liquid Luria–Bertani (LB) medium (with 50 $mg \cdot L^{-1}$ kanamycin and 40 $mg \cdot L^{-1}$ gentamicin) on a rotary shaker at 180 rpm at 28 ℃ for 16 h and grown to an OD_{600} of 1.4-1.6. The *Agrobacterium* strains were flash-frozen in liquid nitrogen with glycerol at a final concentration of 30% (v/v) prior to storage at −80 ℃.

The stored *Agrobacterium* strains (20 μL) were inoculated into 3 mL of LB medium as above on a rotary shaker at 180 rpm at 28 ℃ for 24 h and grown to an OD_{600} of 1.4-1.6. Then 1 mL of the culture was inoculated into 100 mL of LB medium and incubated as above for 12-16 h to an OD_{600} of 1.0-1.2. The *Agrobacterium* cells were centrifuged at 3000 g at room temperature for 10 min, washed twice and then re-suspended using the infiltration solution (1/2 MS medium, 10 mM $MgCl_2$, 10 mM MES, 200 μM acetosyringone, pH 5.6) to a final

240 OD_{600} of 0.7-0.8, and placed at 26 ℃ in darkness for 4 h. The infiltration
241 solution of the *Agrobacterium* strain containing TRV1 was mixed with TRV2 or
242 an infiltration solution of the *Agrobacterium* carrying the constructs at a 1 to 1
243 ratio (v/v). Silwet L77 (GE, USA) was added into the infiltration solution to a final
244 concentration of 0.05 % (v/v) and mixed well immediately. About 30 seeds with
245 homogenous germination were placed in a 150 mL flask containing 100 mL of
246 the infiltration mixture solution. *Agrobacterium* was infiltrated into soybean
247 sprouts using a vacuum dryer (DZF-6050, Jinghong，Shanghai，PR China).
248 Vacuum was maintained at -25 kPa for 15 s, then decompressed to
249 atmospheric pressure rapidly. In each experiment, the operation was repeated
250 three times. The treated seeds were grown hydroponically on 250 mL MS media
251 in a 250 mL flask for 18-20 days at 26 ± 1 ℃ with a light regime of 16 h light /
252 8 h dark, and the MS media was changed every other day. The homogeneous
253 root seedlings were then transferred into the MS liquid medium with 0 or 0.5
254 $mg \cdot L^{-1}$ Cd for 4 days. The incubation conditions and the root measuring method
255 were performed as described in section 2.1.

256 **Statistical analysis**

257 SPSS (version 23.0) was used for statistical analyses of the experimental
258 data. Results are expressed as the means ± standard deviation (SD) of three
259 independent experiments. The data were analyzed by two-way analysis of
260 variance (ANOVA) at $P < 0.05$. (Tables 1-2, Tables S1, S3-4). The differences
261 in the same cultivar among the Cd treatment and the differences between the

cultivars under the same Cd treatment were further evaluated by one-way ANOVA test at $P < 0.05$.

Results

Cd stress suppressed the root growth of soybean seedlings

To investigate the effect of Cd stress on soybean seedling growth, a total of twenty-two American and Chinese soybean cultivars were exposed to Cd (0.25-0.5 mg•L-1) stress for 4 days. There existed an obvious variation in the Cd-tolerance among the twenty-two soybean cultivars under Cd treatment (Table S1). Two soybean cultivars contrasting in Cd sensitivity, Liaodou10 (LD10) and Shennong 20 (SN20) were screen for downstream molecular studies. The result showed that, Cd treatment (0.25-0.5 $mg \cdot L^{-1}$) had no statistically significant ($P < 0.05$) effect on seed germination efficiency, fresh weight, and dry weight compared to the control in either LD10 or SN20 seedlings (Table 1). However, 2.5 $mg \cdot L^{-1}$ Cd treatment significantly reduced the germination percentage, fresh weight and dry weight of both LD10 and SN20 compared to the control. Exposure to 0.25 $mg \cdot L^{-1}$ Cd stress for 4 days significantly reduced growth of LD10 roots, resulting a reduction in root length to 29.84% of the control, and a dose-dependent decrease in root length was observed with increasing Cd concentrations. In contrast, there were statistically significant differences in root length between the control and Cd-treated SN20 seedlings only at the Cd concentrations above 0.5 $mg \cdot L^{-1}$, with a reduction in root length to 24.96% of the control at 0.5 $mg \cdot L^{-1}$ Cd and 46.16% at 2.5 $mg \cdot L^{-1}$ Cd,

284 respectively (Table 1 and Figure S1).

285 **Table 1.** Effect of Cd stress (0-2.5 mg·L^{-1}) on growth of LD10 and SN20 seedlings for 4

286 days.

Cultivar	Cd treatment (mg·L^{-1})	Germination (%)	Fresh weight (g)	Dry weight (g)	Root length (cm)	Length reduction (%)
LD10	0	72.45 ± 2.82a	3.89 ± 0.26a	0.77 ± 0.04a	5.63 ± 0.11a	0
	0.25	70.48 ± 0.95a	3.88 ± 0.51a	0.77 ± 0.01a	3.95 ± 0.15b	29.84 ± 0.16c
	0.5	70.69 ± 2.25a	3.74 ± 0.26b	0.74 ± 0.04a	2.58 ± 0.04c	54.17 ± 0.27b
	2.5	61.11 ± 1.11b	2.95 ± 0.18c	0.58 ± 0.08b	1.38 ± 0.06d	75.49 ± 0.09a
SN20	0	82.94 ± 3.10a	3.64 ± 0.18a	0.79 ± 0.05a	7.17 ± 0.15a	0
	0.25	80.60 ± 6.24a	3.65 ± 0.17a	0.78 ± 0.03a	6.92 ± 0.17a	3.49 ± 0.08c
	0.5	77.26 ± 3.80a	3.49 ± 0.09b	0.77 ± 0.01a	5.31 ± 0.11b	25.94 ± 0.15b
	2.5	65.19 ± 1.15b	3.02 ± 0.08c	0.67 ± 0.04b	3.86 ± 0.31c	46.16 ± 0.21a
Source of variation						
Cultivar (C) (df=1)		36.89 **	0.27NS	171.57**	1421.44**	75956.56**
Treatment (T) (df=3)		25.33 **	208.93**	6.39**	664.93**	122428.92**
C×T (df=3)		1.41NS	1.55NS	4.25*	23.6**	8504.53**

287 *, significant at the $P < 0.05$ level; **, significant at the $P < 0.01$ level. NS, Not significant.

288 For the same cultivar, different letters indicate statistically significantly differences ($P < 0.05$)

289 among different Cd treatment. Standard deviations were calculated with five independent

290 experiments each comprising 35 soybean seedlings.

291 **Cd stress induced DNA damage in soybean seedling roots**

292 DNA damage levels in LD10 and SN20 seedling roots grown under Cd

293 stress for 4 days was assessed using a RAPD assay on DNA extracted from

294 the control and Cd-treated (0.25-2.5 mg·L^{-1}) seedling root tips. Cd stress

295 significantly increased the frequency of RAPD polymorphism even at low

296 concentrations and substantially decreased the stability of the genome template

297 in both LD10 and SN20 roots compared with the control (Figure 1a-b, Figure

298 S2). Interestingly, higher RAPD polymorphism and more reduction in the

299 stability of genome template occurred in SN20 root tips than those in LD10

300 under Cd stress of 0.25-2.5 mg·L^{-1}.

301 The transcriptional regulation of DNA damage repair genes *BRCA1*, *RAD51*,
302 *MRE11* and *KU70* was studied by qRT-PCR analysis in LD10 and SN20
303 seedling roots under Cd stress. qRT-PCR analyses results indicate that the
304 expression levels of these DNA damage repair genes, involved in HR (*RAD51*
305 and *BRCA1*) and NHEJ (*MRE11* and *KU70*) were significantly up-regulated by
306 0.25 mg·L^{-1} Cd stress in both LD10 and SN20 seedling roots compared with
307 the control (Figure 1c-f). Moreover, SN20 showed higher expression level of
308 these DNA damage repair genes than LD10 did under 0.25 mg·L^{-1} Cd stress.
309 However, 0.5 mg·L^{-1} Cd stress significantly down-regulated the expression
310 levels of *BRCA1* and *KU70* genes in LD10 seedling roots, and of *BRCA1* and
311 *RAD51* genes in SN20 seedling roots. A higher concentration (2.5 mg·L^{-1}) of
312 Cd treatment down-regulated the expression levels of all of these genes in both
313 LD10 and SN20 seedling roots (Figure 1c-f). Taken together, the results indicate
314 that Cd stress can induce higher expression of DNA damage regulatory genes
315 in SN20 than in LD10 seedling roots.

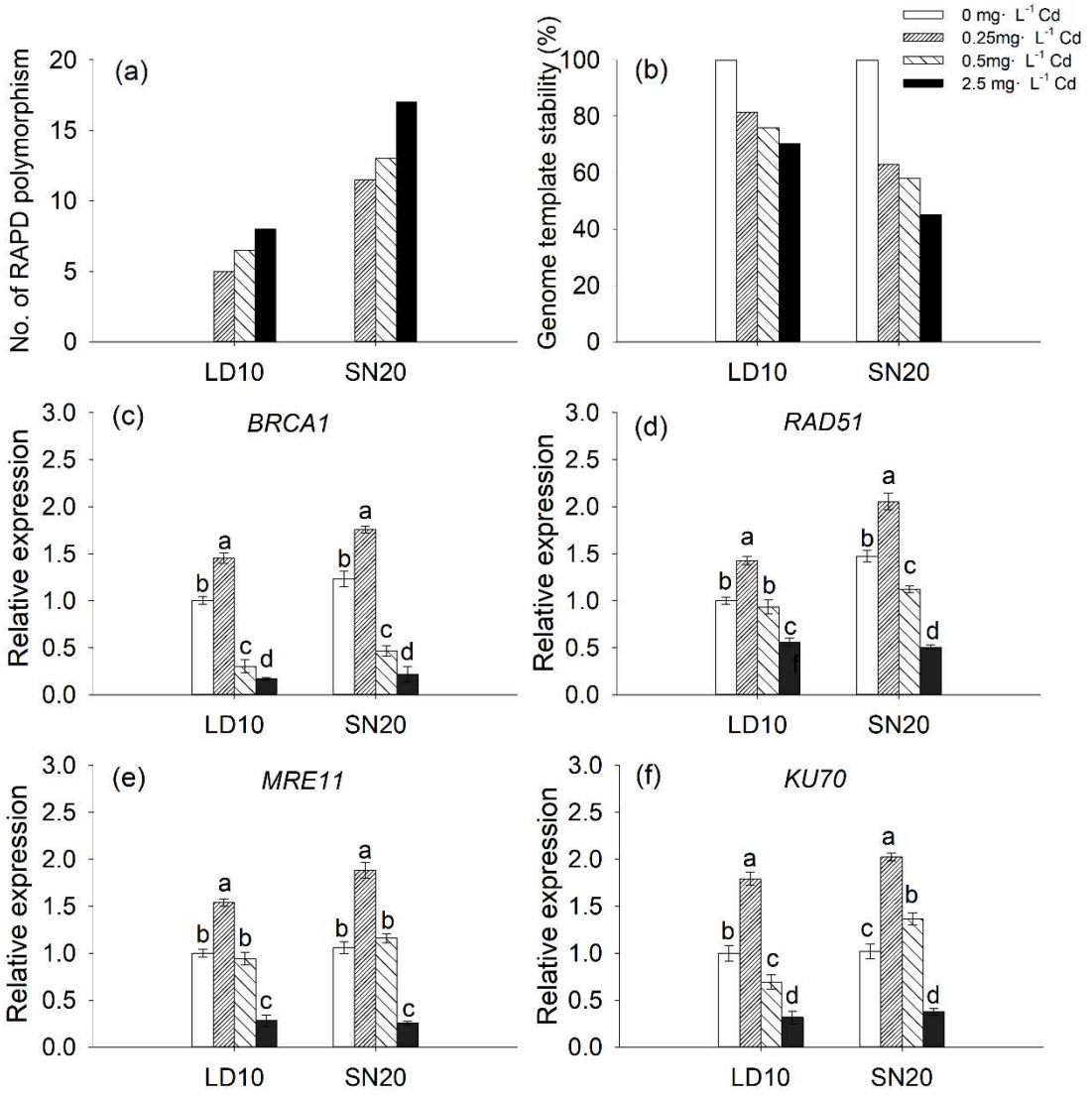


Figure 1. Cd stress induced DNA damage in LD10 and SN20 seedling root tips when grown under 0-2.5 mg·L^{-1} Cd stress for 4 days. (a) RAPD polymorphism variation; (b) The GTS; (c-f) The relative expression level of DNA damage repair genes. Gene expression levels of the LD10 under control conditions were set to 1 as the normalization in qRT-PCR analysis. Standard deviations were calculated with three independent experiments. Different letters indicate statistically significantly differences ($P < 0.05$).

Cd stress induced the cell cycle progression arrest in soybean root tips.

To evaluate cell cycle progression in root tips of LD10 and SN20 seedlings

grown under Cd stress (0-2.5 mg·L^{-1}) for 4 days, the ploidy was determined using FCM analysis. Cd treatment significantly increased the proportion of 2C nuclear content (G0/G1 phase) cells (by 15.4-50.5%) in root tips of LD10 seedlings compared to the control (32.46%), while the proportion of 4C nuclear content cells decreased significantly by 7.9-24.9% (Figure 2a and Figure S3). In contrast, the proportion of 2C nuclear content cells in SN20 seedling root tips under Cd stress decreased by 10.9-15.3% when compared with the control (53.33 %), but the proportion of 4C nuclear content cells increased by 12.1-18.5% (Figure 2b). The FCM results indicate that Cd stress could induce a G1/S phase arrest in root tips of Cd-sensitive soybean genotype LD10, and G2/M phase arrest in root tips of Cd-tolerant soybean genotype SN20, respectively.

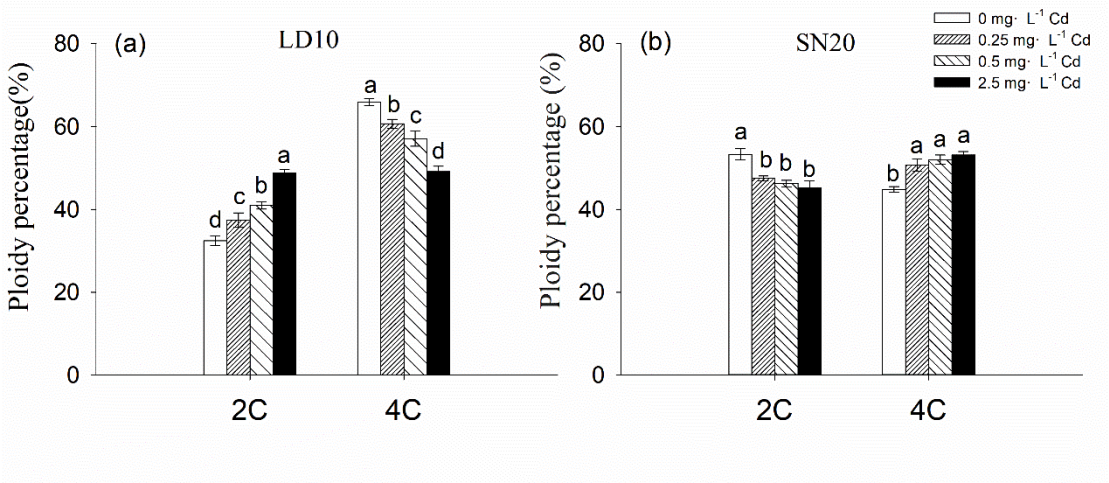


Figure 2. FCM analysis of the nuclear DNA contents of soybean genotypes LD10 (a) and SN20 (b) seedling roots under Cd stress for 4 days. The percentage of 2C and 4C nuclear content cells in the total cell population was calculated. Standard deviations were calculated with three independent experiments. Different letters indicate statistically significant differences ($P < 0.05$).

To assess the cell cycle progression in root tips of soybean seedlings after Cd stress for 4 days, the expression levels of *PCNA1, E2Fa* and *HISTONE H4* (G1/S phase transition regulation/marker genes), *ATM* and *ATR* (DNA damage response genes), *CYCB1;1, CDKA;1* and *WEE1* (G2/M phase transition regulation/marker genes) were measured in the LD10 and SN20 seedling root tips by qRT-PCR. Expression levels of *PCNA1*, *E2Fa*, *HISTONE H4, CYCB1;1* and *CDKA;1* were significantly down-regulated with a dose-dependent response related to the concentration of Cd treatment in both LD10 and SN20 seedling root tips (Figure 3). In contrast, in root tips of both LD10 and SN20 seedlings, the expression level of *WEE1*, *ATM* and *ATR* genes was up-regulated by 1.2 to 1.7- fold at 0.25 $mg \cdot L^{-1}$ Cd treatment, but a dose-dependent decrease was observed in the expression of *WEE1* with Cd concentrations ≥ 0.5 $mg \cdot L^{-1}$ and a significant suppression in expression of *ATM* and *ATR* genes only at 2.5 $mg \cdot L^{-1}$ Cd. Interestingly, SN20 root tips showed higher expression level *of ATR* and lower level *of ATM* than those of LD10 under 0.25 $mg \cdot L^{-1}$ Cd stress. Taken together, these data indicate that Cd stress had striking effects on the expression of cell cycle marker genes in LD10 and SN20 seedling root tips, and provides evidence towards the hypothesis that Cd stress induces G1/S phase arrest in LD10 and G2/M phase arrest in SN20 seedling root tips.

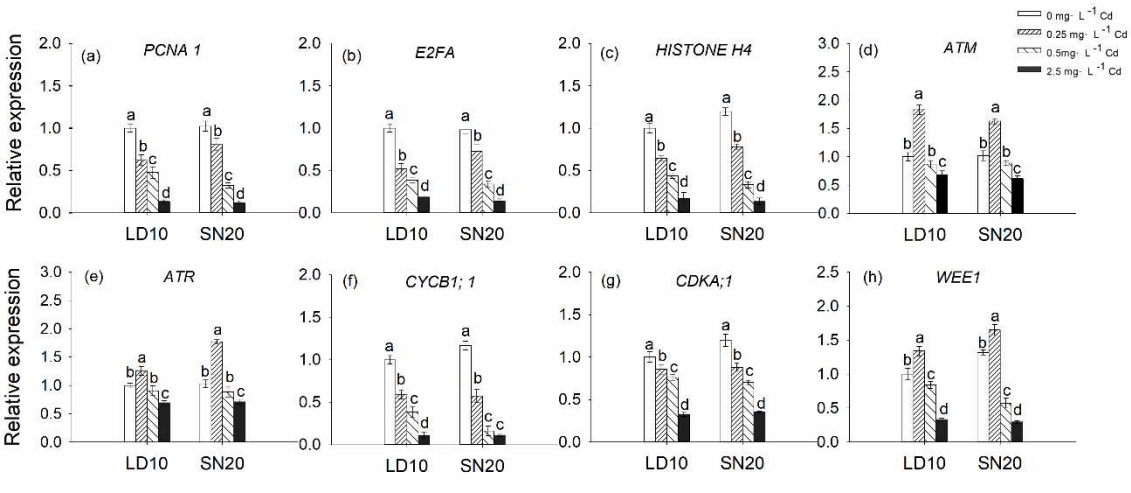
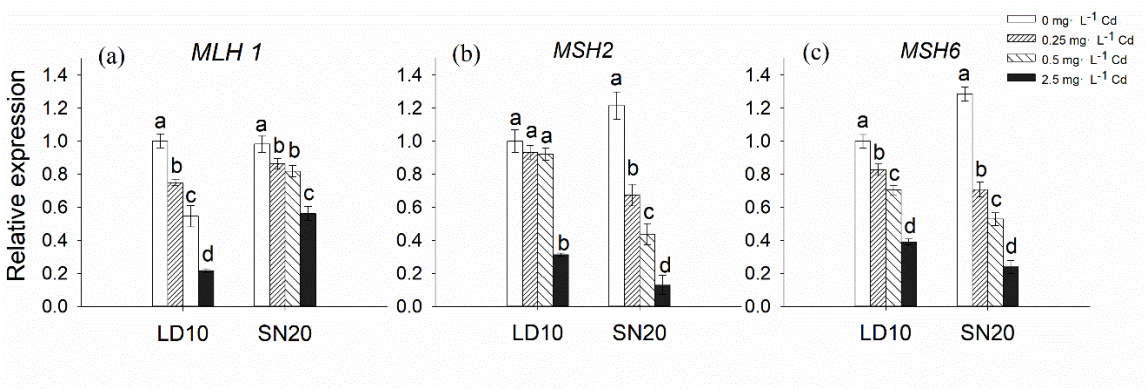


Figure 3. Relative gene expression levels in root tips of LD10 and SN20 seedling exposed to 0-2.5 mg·L^{-1} Cd for 4 days. In a-h, G1/S phase transition regulation / marker genes *PCNA1*, *E2FA*, *HISTONE H4*; DNA damage response genes *ATM*, *ATR*; G2/M phase transition regulation / marker genes *CYCB1;1*, *CDKA;1*, *WEE1*. Gene expression levels of the LD10 seedling root tips under normal condition were set to 1 as the normalization for qRT-PCR analysis. Standard deviations were calculated with three independent experiments. Different letters indicate statistically significant differences ($P < 0.05$).

Cd stress regulated MMR transcripts in soybean seedling roots

To investigate the role of the MMR system in the Cd-induced DNA damage response pathway in soybean, the transcriptional regulation of MMR genes was determined by qRT-PCR analysis. As shown in Figure 4, exposure to Cd stress (0.25-0.5 mg·L^{-1}) for 4 days significantly decreased the expression level of *MLH1* and *MSH6* in both SN20 and LD10, but *MSH2* was only down regulated in SN20 root tips compared with the control at these Cd concentrations. However, expression of all three genes was significantly down-regulated by 2.5 mg·L^{-1} Cd treatment in both LD10 and SN20. Surprisingly, the expression levels

379 of *MSH2* and *MSH6* genes were significantly ($P < 0.05$) higher in SN20 than
380 those in LD10 root tips under the control conditions. Inversely, LD10 had a
381 higher *MSH6* expression level than SN20 did when exposed to Cd stress (0.25-
382 2.5 mg·L^{-1}) for 4 days. The above results reveal a significant difference in basal
383 expression of *MSH2* and *MSH6* between LD10 and SN20 cultivars under
384 normal conditions, while showing Cd hypersensitivity of *MLH1* in LD10 and
385 *MSH2* and *MSH6* in SN20.



386

387 **Figure 4.** Relative gene expression levels of *MLH1* (a), *MSH2* (b), and *MSH6* (c) genes
388 in root tips of LD10 and SN20 seedling exposed to 0-2.5 mg·L^{-1} Cd for 4 days. Gene
389 expression levels of the LD10 seedling roots under normal condition were set to 1 as the
390 normalization in the qRT-PCR analysis. Standard deviations were calculated with three
391 independent experiments. Different letters indicate statistically significant differences ($P <$
392 0.05).

393 **Soybean MMR-silenced plants developed by VIGS**

394 *MLH1*, *MSH2* and *MSH6* genes silencing seedlings of LD10 and SN20 were
395 developed using a TRV-based VIGS system to further understand the role of
396 MMR genes in soybean Cd tolerance. The soybean *PDS* gene, encoding a key

enzyme in the carotenoid synthesis pathway, was used as a reporter gene for testing the TRV-based gene silencing efficiency in soybean plantlets. As shown in Figure 5a-d, the newly formed leaves of the TRV-*PDS* plantlets of LD10 and SN20 showed very obvious photo-bleaching compared with the uninfected plants at 20 days after *Agrobacterium* infection (Figure S5). Likewise, 85.33-88.05% TRV-*PDS* infected plants showed a photo-bleaching phenotype (Figure S5).

To confirm suppression of the *MLH1*, *MSH2* and *MSH6* genes using the TRV-based VIGS system, transcript levels were measured by qRT-PCR using gene-specific primers (Table S2). As shown in Figure 5a-c, the expression levels of *MLH1*, *MSH2* and *MSH6* genes in their corresponding TRV-based MMR gene silencing plantlets were significantly reduced ($P < 0.05$) when compared with the TRV2 infected or the uninfected plants. In contrast, the transcript level of housekeeping genes (*Tubulin A* and *Actin* gene) or cell cycle-regulation genes including *PCNA1, E2Fa*, *HISTONE H4*, *CYCB1;1, CDKA;1* and *WEE1* was not significantly different between the TRV-based gene silencing plants and TRV2 or the uninfected plant root tips under normal culture conditions (Figure S6). This demonstrates that there was no general effect on mRNA stability in the *MMR*-silenced soybean root tips, suggesting that the effects were transcript specific. In addition, the LD10/SN20-TRV-based gene silencing plants showed no visible root phenotype differences under normal growth conditions compared to the corresponding TRV2 and the LD10/SN20

uninfected plants (Figure S7, Table S4).

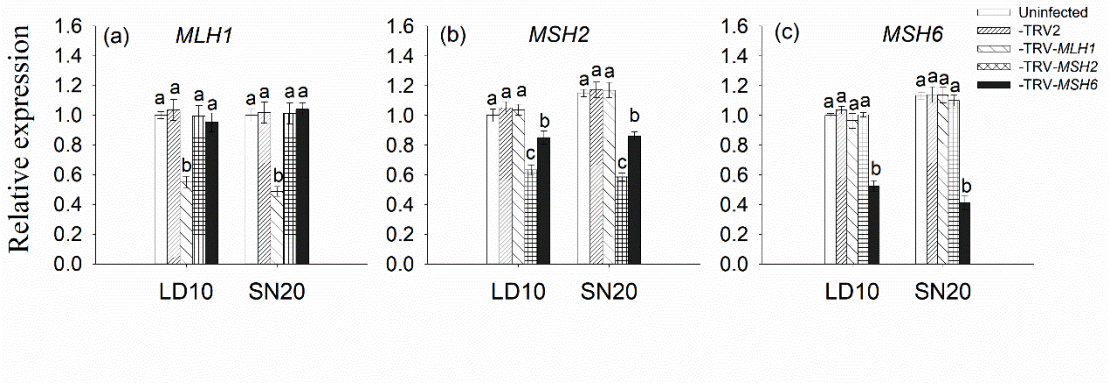


Figure 5. VIGS induced soybean MMR gene silencing. Relative gene expression levels of *MLH1* (a), *MSH2* (b), and *MSH6* (c) genes in root tips of TRV-based MMR gene silencing soybean plantlet under normal culture conditions. Gene expression levels of the uninfected LD10 seedling roots were set to 1 as the normalization in the qRT-PCR analysis. Standard deviations were calculated with three independent experiments. Different letters indicate statistically significant differences ($P < 0.05$).

Soybean MMR system was involved in the Cd-induced root growth inhibition

To investigate the effect of the MMR system on the tolerance of soybean roots to Cd toxicity, uninfected LD10 and SN20 plantlets and their VIGS-induced gene silencing lines (including TRV2, TRV-*MLH1*, TRV-*MSH2* and TRV-*MSH6* plantlets) were exposed to 0.5 mg·L^{-1} Cd stress for 4 days. Cd stress significantly inhibited total root length and total root area, but not root diameter in both uninfected LD10 and SN20 seedlings compared with the corresponding control (Table 2, Figure S7). Unexpectedly, total root length and total root area were significantly higher in LD10-TRV-*MLH1* than those in LD10, LD10-TRV, LD10-TRV-*MSH2* and LD10-TRV-*MSH6* lines with similar phenotypes under

438 0.5 mg·L^{-1} Cd stress for 4 days. However, total root length and total root area
439 in SN20-TRV-*MSH2* and SN20-TRV-*MSH6* seedlings were significantly
440 reduced compared with the uninfected SN20, SN20-TRV2 and SN20-TRV-
441 *MLH1* seedlings under 0.5 mg·L^{-1} Cd stress. Taken together, the results indicate
442 that VIGS-induced *MLH1* gene silencing increased Cd toxicity resistance in the
443 LD10 line, while *MSH2* and *MSH6* gene silencing decreased Cd toxicity
444 resistance in the SN20 line.

445 **Table 2.** Effect of Cd stress on root growth of TRV-based MMR gene silencing soybean
446 plantlets for 4 days.

Soybean lines	Cd treatment (mg·L^{-1})	Total root length (cm)	Total root area (cm^2)	Root diameter (mm)
LD10	0	132.64 ± 9.32 b	24.39 ± 2.89 b	0.61 ± 0.04 a
LD10	0.5	68.84 ± 2.85 f	10.19 ± 1.43 f	0.52 ± 0.05 a
LD10-TRV2	0.5	68.61 ± 3.05 f	10.22 ± 1.05 f	0.56 ± 0.04 a
LD10-TRV-*MLH1*	0.5	83.11 ± 3.75 e	13.52 ± 0.69 e	0.55 ± 0.03 a
LD10-TRV-*MSH2*	0.5	66.91 ± 1.84 f	11.02 ± 0.82 f	0.56 ± 0.03 a
LD10-TRV-*MSH6*	0.5	64.35 ± 3.26 f	10.22 ± 1.05 f	0.56 ± 0.09 a
SN20	0	147.9 ± 2.17 a	31.09 ± 0.98 a	0.53 ± 0.04 a
SN20	0.5	111.09 ± 2.83 c	20.62 ± 1.23 c	0.58 ± 0.05 a
SN20-TRV2	0.5	112.37 ± 5.33 c	20.42 ± 1.30 c	0.57 ± 0.03 a
SN20-TRV-*MLH1*	0.5	106.58 ± 5.39 c	19.38 ± 0.75 c	0.55 ± 0.04 a
SN20-TRV-*MSH2*	0.5	94.66 ± 4.91 d	16.11 ± 0.97 d	0.53 ± 0.08 a
SN20-TRV-*MSH6*	0.5	90.13 ± 2.66 d	16.18 ± 0.99 d	0.52 ± 0.06 a
Source of variation				
Cultivar (C) (df=9)		20.14**	23.56**	0.98NS
Treatment (T) (df=1)		632.34**	535.58**	0.02NS
C×T (df=9)		4.4**	2.05NS	1.18NS

447 Standard deviations were calculated with three independent experiments. For each
448 experiment, at least 10 soybean seedling plants were used for each treatment. Different
449 letters indicate statistically significant differences ($P < 0.05$).

450 **Contribution of the soybean MMR system in the Cd-induced DNA damage**

451 To assess the role of soybean MMR proteins in the Cd-induced DNA

damage signaling pathway, DNA damage levels of LD10 and SN20 roots exposed to 0.5 mg·L^{-1} Cd for 4 days were analyzed by a RAPD assay and compared to the VIGS silencing lines. As shown in Figure 6 and Figure S8, Cd treatment significantly increased the frequencies of RAPD polymorphism in LD10 and SN20 seedling roots compared with the untreated control plantlets. LD10-TRV-*MLH1* plantlet roots showed significantly less RAPD polymorphism compared to all the other LD10 lines. In contrast, SN20-TRV-*MSH2* and SN20-TRV-*MSH6* showed more polymorphic RAPD bands compared to the other SN20 lines. Taken together, these results suggest that in LD10, MLH1 may not determine DNA stability, while in SN20, MSH2 and MSH6 may promote DNA stability in soybean roots under Cd stress.

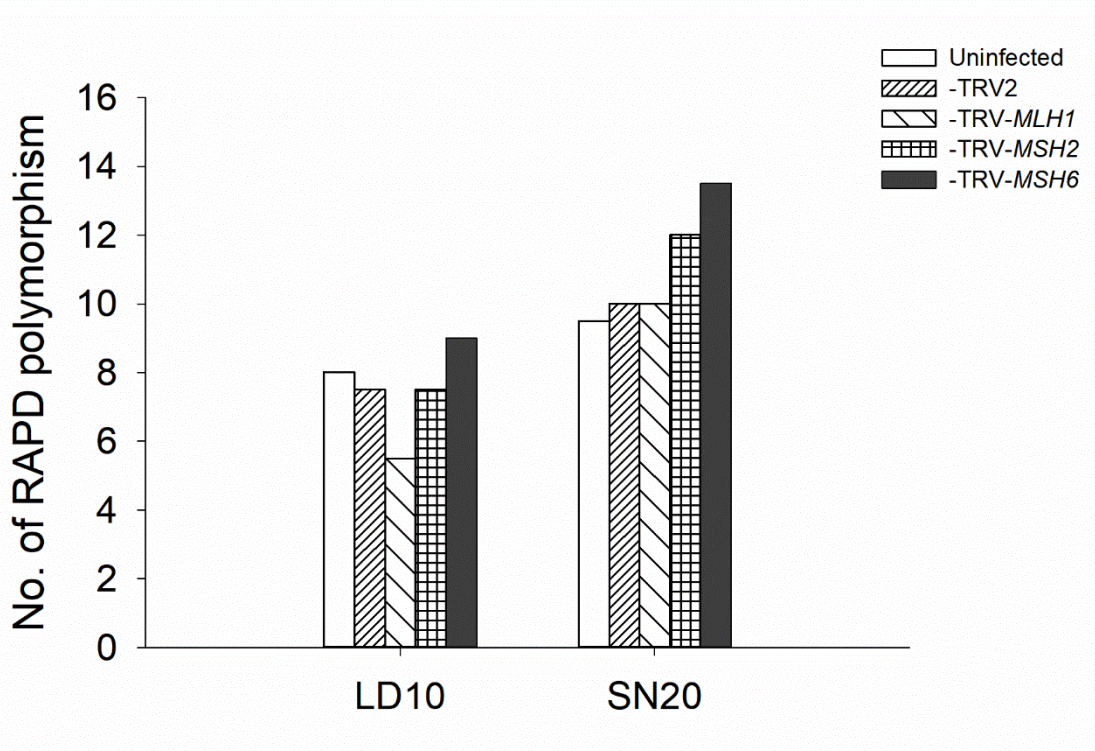


Figure 6. RAPD polymorphism variations in roots of TRV-based gene silencing soybean plantlets exposed to 0.5mg·L^{-1} Cd for 4 days. For all treatments, reproducible bands in at

least two replicates were evaluated and calculated for polymorphism analysis.

Contribution of the soybean MMR system in the Cd-induced cell cycle arrest

To investigate the contribution of the MMR system in Cd-induced cell cycle progression arrest, 1 cm long root tips of LD10 and SN20 lines exposed to 0.5 $mg \cdot L^{-1}$Cd stress for 4 days were harvested for cell cycle progression analysis using FCM. As shown in Figure 7 and Figure S9, Cd-induced G1/S arrest was significantly attenuated in the LD10-TRV-*MLH1* compared with the uninfected LD10 or TRV2 seedling roots under 0.5 $mg \cdot L^{-1}$ Cd stress. The 2C nuclear content decreased by 14.1%, while the 4C nuclear content increased by 14.7% in LD10-TRV-*MLH1* seedling roots relative to the uninfected LD10 under Cd stress. However, in the SN20-TRV-*MSH2* and SN20-TRV-*MSH6* roots, 0.5 $mg \cdot L^{-1}$ Cd stress significantly increased the proportion of cells with 2C nuclear content, which was 19.5% and 17.6% compared with the uninfected SN20 (Figure 7); whereas there was a reduction of 4C nuclear content by 12.9% and 12.1%, respectively. The results indicate that Cd-induced G2/M arrest was attenuated in the SN20-TRV-*MSH2* and TRV-*MSH6* seedling roots. Taken together, the results indicate that *MLH1* is involved in the Cd-induced G1/S phase arrest, while *MSH2* and *MSH6* are involved in the Cd-induced G2/M phase arrest in root tips of soybean seedlings.

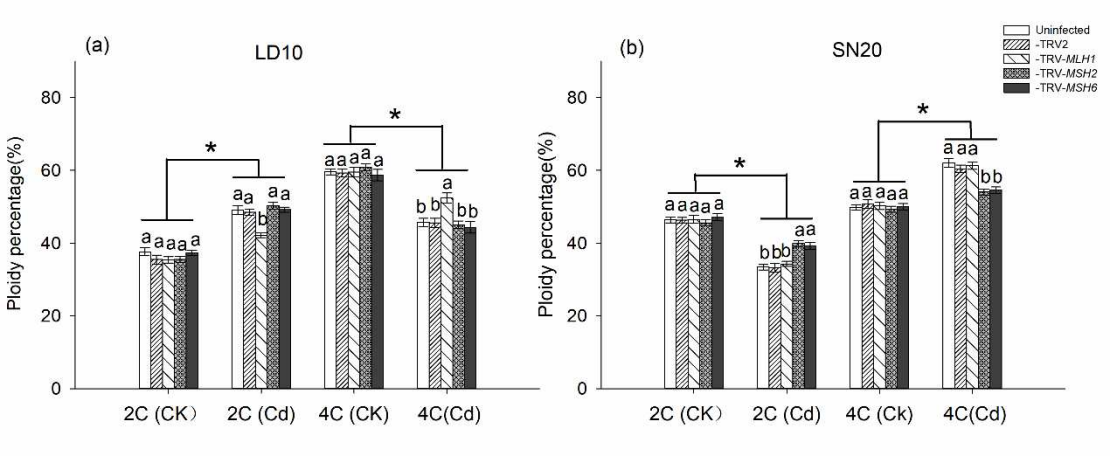


Figure 7. FCM analysis of the nuclear DNA contents of soybean genotypes LD10 (a) and SN20 (b) seedling roots under 0-0.5 mg·L^{-1} Cd stress for 4 days. The percentage of 2C and 4C nuclear content cells in total cells was calculated，respectively. Standard deviations were calculated with three independent experiments. Different letters indicate statistically significant differences ($P < 0.05$). * indicate statistically significant differences ($P < 0.05$)

qRT-PCR results showed that exposure to 0.5 mg·L^{-1} Cd stress for 4 days significantly influenced the expression levels of DNA damage response and repair genes in LD10, SN20 and their VIGS-induced gene silencing seedling roots compared with the control (Figure 8). Notably, LD10-TRV-*MLH1* seedling roots showed higher gene expression levels of *MLH1, MSH6*, *BRCA, RAD51* and *KU70*, but lower gene expression levels of *MSH2* and *ATM,* compared with the uninfected LD10 and LD10-TRV seedling roots under 0.5 mg·L^{-1} Cd stress for 4 days. Similar expression levels of the same DNA damage response and repair genes occurred in LD10-TRV-*MSH2 and* LD10-TRV-*MSH6* seedling roots compared with the uninfected LD10 and LD10-TRV2 seedling roots under 0.5 mg·L^{-1} Cd treatment for 4 days, with the exception that *MSH2* and *MSH6* gene expression, which was reduced.

However, in SN20-TRV-*MSH2* seedling roots, expression levels of some genes (i.e. *MLH1*, *MSH2*, *MSH6*, *ATR*, *BRCA1* and *RAD51*) were significantly reduced, while others such as *ATM* and *KU70* were up-regulated (i.e. an increase of 1.13- to 1.23- fold) compared with those in the SN20 or SN20-TRV root tips under Cd stress of 0.5 mg·L^{-1} ($P < 0.05$). A similar trend appeared in SN20-TRV-*MSH6* seedling roots. Furthermore, expression levels of *PCNA1*, *E2FA*, *HISTONE H4* were significantly higher in LD10-TRV-*MLH1* seedling roots than those in uninfected LD10, LD10-TRV2 and LD10-TRV-*MSH2/6* seedling roots exposed to 0.5 mg·L^{-1} Cd stress for 4 days (Figure S10). In contrast, there were no significant differences in the expression levels of *PCNA1*, *E2FA*, *HISTONE H4*, *CYCB1;1*, *CDKA;1*, or *WEE1* genes between the SN20-TRV-*MSH2* and the SN20/SN20-TRV2 seedling roots exposed to Cd stress of 0.5 mg·L^{-1}. A similar trend occurred in SN20-TRV-*MSH6*, SN20-TRV-*MLH1*, LD10-TRV-*MSH2*, and LD10-TRV-*MSH6* seedling roots (Figure S10).

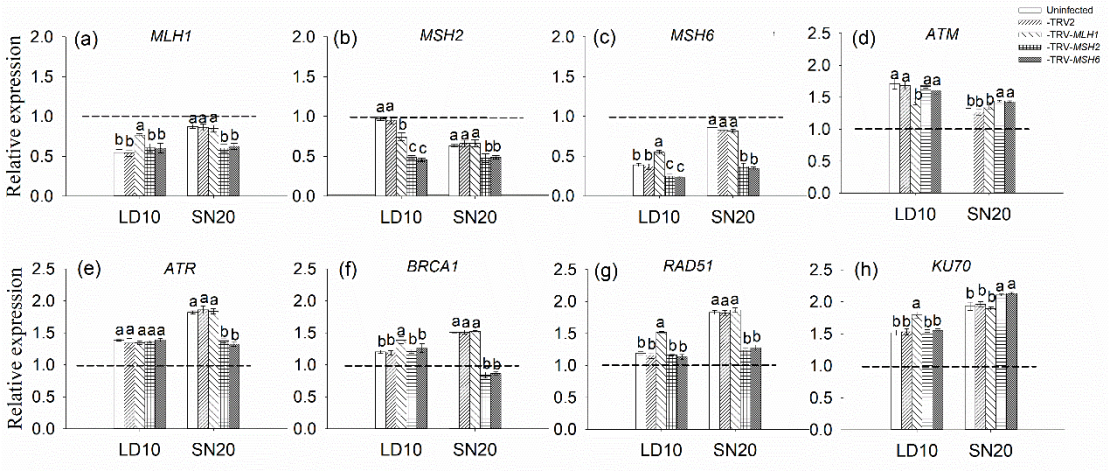


Figure 8. Relative gene expression levels of DNA damage repair genes in seedling roots of LD10 and SN20 genotypes exposed to 0.5 mg·L^{-1} Cd for 4 days. Dashed line indicate

gene expression levels of LD10 seedling roots grown under control conditions were set to 1 as the normalization in qRT-PCR analysis. Standard deviations were calculated with three independent experiments. Different letters indicate statistically significant differences ($P < 0.05$) in a to h.

Discussion

Exposure to Cd stress inhibits plant growth and metabolism, and induces different types of DNA damage including DNA single strand breaks (SSB) and double strand breaks (DSB).[35] DNA damage signals lead to: (1) activation of cell cycle checkpoints resulting in cell cycle arrest, and activation of DNA repair pathways, or (2) induction of apoptosis.[36,37] Previous studies identified the role of the MMR system in Cd toxicology and that *MSH2* and *MSH6* primarily contribute to Cd-induced G_2/M arrest causing suppressed growth of Arabidopsis roots.[4,27] In this study, exposure to (0.25-2.5 $mg \cdot L^{-1}$) Cd stress for 4 days inhibited the growth of soybean seedling roots. Two contrasting soybean cultivars, LD10 (Cd-sensitive) and SN 20 (Cd-tolerant) were used to study the mechanism of cultivar-dependent Cd stress responses in soybean.

There was a significant difference in DNA damage and cell cycle arrest between LD10 and SN20

RAPD analysis indicated that exposure to Cd stress for 4 days, even at low concentrations (0.25 $mg \cdot L^{-1}$), could induce DNA damage in both Cd-sensitive soybean cultivar LD10 and in Cd-tolerant soybean cultivar SN20 (Figure 1a). This result is consistent with previous researches in Arabidopsis, rice and

543 barley.[4,6,38] Interestingly, although LD10 showed a higher reduction in root
544 length than SN20 when exposed to Cd (0.25-2.5 mg·L^{-1}) stress for 4 days, LD10
545 showed higher genomic stability than SN20 (Figure 1b). Furthermore, the
546 expression levels of DNA DSB repair genes (i.e. *BRCA1*, *RAD51*, *MRE11* and
547 *KU70*) in SN20 seedling roots were significantly higher than those in LD10
548 under corresponding Cd stress, suggesting that SN20 had suffered more
549 serious DSBs induced by Cd stress than LD10 did. Previous studies have
550 shown that RAD51 and BRCA1 are responsible for repair of DSBs via HR.[39]
551 HR needs the homologous sequence of the uninjured sister chromatid as a
552 template for DNA damage repair, which is a complex but precise process for
553 repairing DNA damage. In contrast, MRE11 and KU70 are involved in repairing
554 DSBs via NHEJ.[40] Instead of relying on homologous DNA sequences, the NHEJ
555 pathway directly connects the ends of DSBs using DNA ligase, which is a fast
556 DSBs repair process, but can result in deletions and insertions. Although high-
557 fidelity genetic information is very important for organisms, perhaps it is more
558 beneficial for organisms to tolerate some DNA damage rather than to allow the
559 replication fork to collapse.[29] The DNA damage tolerance (DDT) phenomenon
560 is widespread in eukaryotic cells, allowing the organism to avoid compromised
561 genome integrity or cell death.[29,41] Here we show that although multiple DNA
562 repair systems were more highly activated in the Cd tolerant SN20 cultivar, DNA
563 damage was still greater than in the more sensitive LD10, thus the abiotic
564 stress-induced DNA damage was not fully avoided.

565 The above results might be related to the different points in the cell cycle
566 where the cells arrest: Cd-induced G1/S phase cell cycle progression arrest in
567 LD10 seedling roots and G2/M phase cell cycle arrest in SN20 seedling roots
568 (Figure 2). DNA damage can activate checkpoint pathways at different phases
569 of the cell cycle.[42] Cd-induced G1/S phase cell cycle arrest in LD10 (Cd-
570 sensitive) seedling roots inhibited DNA replication, causing an increase of 2C
571 nuclear content, however. This may contribute to the increased stability of the
572 genomic DNA, due to higher fidelity of DNA replication compared to SN20,
573 although Cd still seriously inhibited root growth (Table 1, Figures 1 and 2). G2/M
574 phase arrest in SN20 seedling roots inhibited mitosis, leading to the increase of
575 4C nuclear content and lower genomic stability. Possibly through translesion
576 DNA synthesis (TLS) mechanisms,[43] DNA replication in S phase is permitted
577 using damaged DNA as a template to keep soybean plantlets growing. This
578 then resulted in DNA damage spreading, as seen by greater RAPD
579 polymorphism in SN20 than in LD10 roots (Table 1, Figures 1 and 2).
580
581 **Differentially expressed and responsive MMR genes determine Cd-**
582 **induced root growth repression in soybean by regulating the cell cycle**
583 Results from the present study (Figure 4) indicate the significantly different
584 basal-expression of MMR genes between LD10 and SN20. This is integrated
585 with cell cycle arrest, accounting for the Cd-tolerant characteristics in soybean.
586 As is known, the MMR system not only corrects biosynthetic errors, but also

surveys DNA damage and participates in the regulation of cell cycle progression in response to abiotic stress induced DNA damage.[16,44,45] This was shown in Arabidopsis *msh2* and *msh6* mutants in our previous work.[46] In this study, the higher expression of *MSH2* and *MSH6* in SN20 indicate their preferential recognition of the DNA damage, resulting in G_2/M phase arrest via MutS-to-ATR/ATRIP signaling (Figures 9 and 10). However, Cd-induced G_2/M phase arrest allows DNA replication, which causes cell volume enlargement and some cell proliferation with post-replication repair, accounting for the Cd-tolerant root growth in SN20. In contrast, the lower expression of *MSH2* and *MSH6* in LD10 bypasses the MMR system monitoring of DNA damage, leading to G_1/S phase arrest. This occurs prevailingly via the MRN complex-to-ATM signaling which is a DSB recognition pathway with participation of MLH1.[26,47] G_1/S phase arrest blocks the cell cycle from entering into S phase to prevent DNA replication, which explains the Cd-sensitive root growth in LD10.

To validate the above hypothesis, MMR genes were knocked down by amiRNAs (artificial miRNAs) using TRV-induced VIGS technology. SN20, the Cd-tolerant cultivar, showed higher expression and responsiveness of *MSH2* and *MSH6* to the Cd stress, resulting in less root growth repression because of G_2/M phase arrest. However, in SN20-TRV-*MSH2* and SN20-TRV-*MSH6* seedlings G_2/M phase arrest was reduced and Cd-induced root growth repression increased. This strongly indicates that it is *MSH2* and *MSH6* expression that decreases root growth repression by regulating G_2/M phase

arrest in SN20. Although root growth was also repressed in SN20-TRV-*MLH1*, the repression was not significant. However, knocking down the expression of *MLH1* in LD10 reduced G_1/S phase arrest. This provides evidence that in LD10, MLH1 that is engaged in the MRN complex-to-ATM pathway regulating G_1/S phase arrest, is responsible for the greater Cd-induced root growth repression compared to SN20.

Taken together the results can be used to build a model for how MMR genes regulate Cd tolerance by regulating the phase of cell cycle arrest and root growth repression (Figure 9). Thus, differentially expression of *MSH2* and *MSH6* play a crucial role in determining the intervarietal Cd tolerance in soybean.

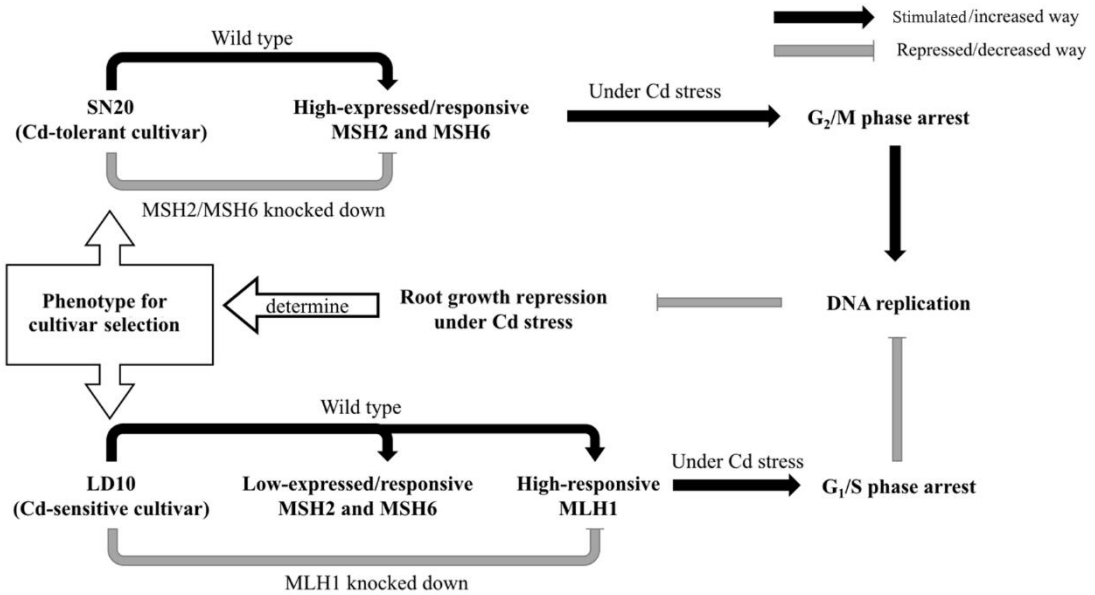


Figure 9. Mechanism of Cd tolerance in soybean SN20 and LD10 cultivars. SN20 and LD10 were respectively Cd-tolerant and Cd-sensitive soybean cultivars selected by Cd-induced root growth repression. In wild type SN20 *MSH2* and *MSH6* are more highly expressed and are more responsive to Cd-stress than in LD10. This causes G_2/M phase

arrest in SN20 but G_1/S phase arrest in LD10 under Cd stress. G_2/M phase arrest in SN20 allows DNA replication leading to cell volume enlargement and some cell proliferation with post-replication repair, but G_1/S phase arrest in LD10 does not. This explains the different Cd-induced root growth repression in SN20 and LD10. This hypothesis was tested by knocking down *MSH2* or *MSH6* from SN20, and by knocking down *MLH1* from LD10 increasing G_1/S phase arrest through blocking the MRN complex-to-ATM signaling.

The MMR system plays multiple roles in Cd-tolerance mechanisms of soybean

Differential expression of *MSH2* and *MSH6* are shown here to influence root growth under Cd stress through regulating the cell cycle. Comparing DNA damage and expression of DNA repair related genes between the LD10 and SN20 wild type and MMR-knocked down plants, another vital role of the MMR system in Cd tolerance is revealed in recruiting DSB repair. BASC, a complex of BRCA1-associated proteins, is involved in several functions such as DNA damage recognition and binding, DNA repair, and downstream activation, in which BRCA1, as the central component of BASC and of HR repair, was found to interact with MSH2 and MSH6.[48] The repressed expression of *BRCA1* and *RAD51* genes in SN20-TRV-*MSH2* and SN20-TRV-*MSH6* seedlings (Figure 8), suggests a recruitment effect of the MMR system, which was also found in our previous study.[27] Indeed, DNA damage was not reduced even though there was decreased G_2/M phase arrest in the TRV lines. In fact, it increased in SN20-TRV-*MSH2/6* seedlings, which indicates weak DNA repair. Also, DNA damage

was not increased as a result of the increased entry into DNA replication caused by reduced G_1/S phase arrest in LD10-TRV-*MLH1*. Indeed DNA damage was reduced in the LD10-TRV-*MLH1* seedling roots, suggesting enhanced DNA repair. MMR and HR repair systems are attributed to post-replication repair and act in the G_2 phase.[46] This explains the increased or decreased DNA damage in SN20-TRV-*MSH2/6* and LD10-TRV-*MLH1* seedlings, respectively. In conclusion, the MMR system not only participates in DNA mismatch repair, DNA error surveillance, and cell cycle regulation, but also recruits HR repair associated proteins in G_2 phase for repairing both SSBs and DSBs,[49] faithfully maintaining genomic integrity and stability (Figure 10). This supplements Cd tolerance and toxicological mechanisms, and moreover provides biomarkers and a molecular basis for selection of Cd-tolerant cultivars.

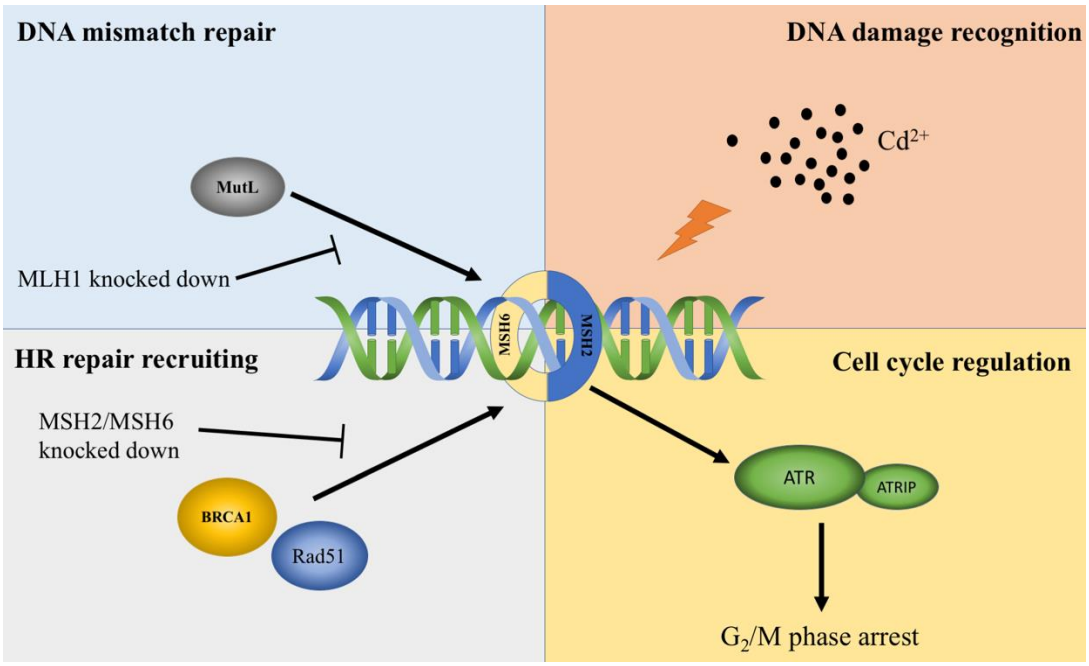


Figure 10. Multiple effects of MMR system on Cd tolerance in soybean. MSH2 and MSH6 forming MutSα regulate the cell cycle by activating the G_2/M checkpoint, leading to cell

662 cycle arrest, when they recognize Cd-induced DNA damage. Furthermore, MutSα can
663 recruit MutL and BRCA1/Rad51 to trigger MMR and HR mediated repair. When *MSH2* or
664 *MSH6* was knocked down, HR repair would be repressed causing increased SSBs and
665 DSBs. Knocking down *MLH1* would suppress MMR, but improve HR repair because of
666 reduced G_1/S arrest leading to enhanced post-replication repair in increased G_2 phase.

667 This is the first report revealing the mechanisms acting in the differential
668 Cd-tolerance of soybean SN20 and LD10 cultivars including Cd-induced DNA
669 damage, DNA repair and cell cycle arrest. Differentially expressed *MSH2* and
670 *MSH6* play a crucial role in Cd-induced root growth repression. A model is
671 proposed in which higher expression of *MSH2* and *MSH6* in SN20 activate
672 MutS-to-ATR/ATRIP signaling, causing G_2/M arrest when Cd-induced DNA
673 damage is detected. This still allows DNA replication, leading to cell volume
674 enlargement and proliferation after post-replication repair mechanisms such as
675 MMR and HR repair are activated. In contrast, LD10 with a lower expression of
676 *MSH2* and *MSH6* bypass the MMR system activating MLH1 that participates in
677 MRN complex-ATM signaling. This causes G_1/S arrest and inhibits DNA
678 replication. In addition, the HR repair system is recruited by MSH2 and MSH6
679 to enhance post-replication repair, thus maintaining genomic integrity and
680 stability under Cd stress. This model explains inter-variety Cd tolerance in
681 soybean and provides both biomarkers and a molecular basis for selection of
682 Cd-tolerant cultivars.

Abbreviations

BER	Base excision repair	MMR	DNA mismatch repair
CDK	Cyclin-dependent kinases	NER	Nucleotide excision repair
DDR	DNA damage response	NHEJ	Non-homologous end joining
DDT	DNA damage tolerance	PAGE	Polyacrylamide gel electrophoresis
DSB	DNA double strand breaks	PDS	Phytoene desaturase
FCM	Flow cytometry method	SSB	Single strand breaks
GTS	Genome template stability	TLS	Translesion DNA synthesis
HR	Homologous recombination	TRV	Tobacco rattle virus
LB	Luria–Bertani medium	VIGS	Virus-induced gene silencing

Author contributions

W.L., F.X. and H. W. designed the experiments. Q.Z., Y.D., Z.W., S.J. and X.Y. conducted the experiments. H.W. and Q.Z. analyzed the data. Q.Z., H.W. and W.L. wrote the manuscript. H.J.R. revised the manuscript. All authors read and approved the manuscript.

Conflict of interest

The authors declare no competing financial interest.

Funding

This work was supported by the National Key Research and Development Plan of Ministry of Science and Technology (2017 YFD 0101306-04), and National Natural Science Foundation of China (NSFC, 21677151, 41807488，41673132).

Supporting Information

Figure S1. Growth performance of soybean seedlings exposed to 0-2.5 mg·L^{-1} Cd stress for 4 days.

Figure S2. RAPD fingerprints of soybean seedlings exposed to 0-2.5 mg·L^{-1} Cd

699 stress for 4 days.

700 **Figure S3.** FCM analysis of the nuclear DNA ploidy in LD10 (a) and SN20 (b)

701 seedlings exposed to 0-2.5 mg·L^{-1} Cd stress for 4 days..

702 **Figure S4.** Vector maps.

703 **Figure S5.** VIGS induced soybean *PDS* gene silencing seedlings

704 **Figure S6.** Relative gene expression levels of cell phase transition regulation

705 genes in seedling roots of soybean VIGS plants under normal cultivation

706 conditions.

707 **Figure S7.** Growth performance of soybean VIGS plant seedlings exposed to

708 0.5 mg·L^{-1} Cd stress for 4 days.

709 **Figure S8.** RAPD fingerprints of wild type and soybean VIGS seedlings.

710 **Figure S9.** FCM analysis of the nuclear DNA ploidy in soybean VIGS seedlings

711 exposed to 0.5 mg·L^{-1} Cd stress for 4 days

712 **Figure S10.** Relative gene expression levels of cell phase transition regulation

713 genes in seedling roots of soybean VIGS plants exposed to 0.5 mg·L^{-1} Cd for 4

714 days.

715 **Table S1.** Effect of Cd stress (0-2.5 mg·L^{-1}) on growth of soybean seedlings for

716 4 days.

717 **Table S2** Primer list

718 **Table S3.** Variance analysis of gene expression in LD10 and SN20 seedlings

719 exposed to Cd stress for 4 days

720 **Table S4.** Variance analysis of gene expression in wild type and soybean

721 VIGS seedlings exposed to Cd stress for 4 days.

References

[1] Zhou C, Ge N, Guo J, et al. Enterobacter asburiae Reduces Cadmium Toxicity in Maize Plants by Repressing Iron Uptake-Associated Pathways. *J. Agric. Food Chem.* **2019**, 67(36): 10126-10136.

[2] N. Sarwar, M. Imran, M.R. Shaheen, W. Ishaque, M.A. Kamran, A. Matloob, A. Rehim, S. Hussain, Phytoremediation strategies for soils contaminated with heavy metals: modifications and future perspectives, *Chemosphere*. **2017**, 171, 710-721.

[3] K. Qiao, S. Liang, F. Wang, H. Wang, Z. Hu, T. Chai, Effects of cadmium toxicity on diploid wheat (*Triticum urartu*) and the molecular mechanism of the cadmium response, *J. Hazard. Mater.* **2019**, 374,1-10.

[4] W. Cui, H. Wang, J. Song, X. Cao, H.J. Rogers, D. Francis, C. Jia, L. Sun, M. Hou, W. Liu, Cell cycle arrest mediated by Cd-induced DNA damage in Arabidopsis root tips, *Ecotoxicol. Environ. Saf*. **2017**, 145,569-574.

[5] D. Bagchi, S.S. Joshi, M. Bagchi, J. Balmoori, E. Benner, C. Kuszynski, S. Stohs, Cadmium- and chromium-induced oxidative stress, DNA damage, and apoptotic cell death in cultured human chronic myelogenous leukemic K562 cells, promyelocytic leukemic HL-60 cells, and normal human peripheral blood mononuclear cells, *J. Biochem. Mol. Toxicol*. **2000**, 14, 33-41.

[6] A.J. Lin, X.H. Zhang, M.M. Chen, C. Qing, Oxidative stress and DNA damages induced by cadmium accumulation, *J. Environ. Sci*. **2007**, 19, 596-602.

[7] H. Wang, L. He, J. Song, W. Cui, Y. Zhang, C. Jia, D. Francis, H.J. Rogers, L. Sun, P. Tai, Cadmium-induced genomic instability in Arabidopsis: Molecular toxicological biomarkers for early diagnosis of cadmium stress, *Chemosphere.* **2016**, 150, 258-265.

[8] H.S. Kim, Y.J. Kim, Y.R. Seo, An overview of carcinogenic heavy metal: molecular toxicity mechanism and prevention, *J. Cancer Prev*. **2015**, 20, 232.

[9] B.B.S. Zhou, S.J. Elledge, The DNA damage response: putting checkpoints in perspective, *Nature.* **2000**, 408, 433.

[10] L.H. Hartwell, M.B. Kastan, Cell cycle control and cancer, *Science.* **1994**, 266, 1821-1828.

[11] C.P. Spampinato, Protecting DNA from errors and damage: an overview of DNA repair mechanisms in plants compared to mammals, *Cell. Mol. Life Sc*i. **2017**, 74, 1693-1709.

[12] T. Cools, L. De Veylder, DNA stress checkpoint control and plant development, *Curr. Opin. Plant Biol*. **2009**, 12, 23-28.

[13] A.K. Weimer, S. Biedermann, H. Harashima, F. Roodbarkelari, N. Takahashi, J. Foreman, Y. Guan, G. Pochon, M. Heese, D. Van Damme, The plant - specific CDKB1 - CYCB1 complex mediates homologous recombination repair in Arabidopsis, *EMBO J*. **2016**, 35, 2068-2086.

[14] P. Modrich, R. Lahue, Mismatch repair in replication fidelity, genetic recombination, and cancer biology, *Annu Rev. Biochem*. **1996**, 65, 101-133.

[15] D. Gupta, C.D. Heinen, The mismatch repair-dependent DNA damage response: Mechanisms and implications, *DNA Rep*. **2019**, 78, 60-69.

[16] L.D. Lario, E. Ramirez-Para, C. Gutierrez, P. Casati, C.P. Spampinato, Regulation of plant MSH2 and MSH6 genes in the UV-B-induced DNA damage response, *J. Exp. Bot.* **2011**, 62, 2925-2937.

[17] M.R. AbdelGawwad, A. Marić, A.A. Al-Ghamdi, A.A. Hatamleh, Interactome Analysis and Docking Sites of MutS Homologs Reveal New Physiological Roles in Arabidopsis thaliana, *Molecules*.**2019**, 24, 2493.
[18] A. Alou, M. Jean, O. Domingue, F. Belzile, Structure and expression of AtPMS1, the Arabidopsis ortholog of the yeast DNA repair gene PMS1, *Plant Sci*. **2004**, 167, 447-456.
[19] N. Jackson, E. Sanchez - Moran, E. Buckling, S.J. Armstrong, G.H. Jones, F.C.H. Franklin, Reduced meiotic crossovers and delayed prophase I progression in AtMLH3 - deficient Arabidopsis, *EMBO J*. **2006**, 25, 1315-1323.
[20] R. Gómez, C.P. Spampinato, Mismatch recognition function of Arabidopsis thaliana MutSγ, *DNA Rep*. **2013**, 12, 257-264.
[21] A.H. Lloyd, A.S. Milligan, P. Langridge, J.A. Able, TaMSH7: a cereal mismatch repair gene that affects fertility in transgenic barley (*Hordeum vulgare* L.), *BMC Plant Biol.* **2007**, 7, 67.
[22] C.M. Bray, C.E. West, DNA repair mechanisms in plants: crucial sensors and effectors for the maintenance of genome integrity, *New Phytol*. **2005**, 168, 511-528.
[23] S Hassen, A A Ali, S P Kilaparty, et al, Interdependence of DNA mismatch repair proteins MLH1 and MSH2 in apoptosis in human colorectal carcinoma cell lines, *Mol Cell Biochem*. **2016**,412, 297–305. .
[24] M Elez, M Radman, I Matic, Stoichiometry of MutS and MutL at unrepaired mismatches in vivo suggests a mechanism of repair, *Nucleic Acids Res*. **2012**, 40, 3929-3938.
[25] K. Yamane, J.E. Schupp, T.J. Kinsella, BRCA1 Activates a G2-M Cell Cycle Checkpoint following 6 - Thioguanine - Induced DNA Mismatch Damage, *Cancer Res*. **2007**, 67, 6286-6292.
[26] Y. Qi, N.W. Schoene, F.M. Lartey, W.-H. Cheng, Selenium compounds activate ATM-dependent DNA damage response via the mismatch repair protein hMLH1 in colorectal cancer cells, *J. Biol. Chem*. **2010**, 285, 33010-33017.
[27] X. Cao, H. Wang, D. Zhuang, H. Zhu, Y. Du, Z. Cheng, W. Cui, H.J. Rogers, Q. Zhang, C. Jia, Roles of MSH2 and MSH6 in cadmium-induced G2/M checkpoint arrest in Arabidopsis roots, *Chemosphere.* **2018**, 201, 586-594.
[28] R. Sobkowiak, J. Deckert, Cadmium-induced changes in growth and cell cycle gene expression in suspension-culture cells of soybean, *Plant Physiol. Biochem*. **2003**, 41, 767-772.
[29] P.L. Andersen, F. Xu, W. Xiao, Eukaryotic DNA damage tolerance and translesion synthesis through covalent modifications of PCNA, *Cell Res*. **2008**, 18, 162.
[30] W. Liu, P. Li, X. Qi, Q. Zhou, L. Zheng, T. Sun, Y. Yang, DNA changes in barley (*Hordeum vulgare*) seedlings induced by cadmium pollution using RAPD analysis, *Chemosphere.* **2005**, 61, 158-167.
[31] K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method, *Methods.* **2001**, 25, 402-408.
[32] Y. Liu, M. Schiff, S. Dinesh-Kumar, Virus-induced gene silencing in tomato, *Plant J*. **2002**, 31, 777-786.
[33] H.X. Yan, D.Q. Fu, B.Z. Zhu, H.P. Liu, X.Y. Shen, Y.B. Luo, Sprout vacuum-infiltration: a simple and efficient agroinoculation method for virus-induced gene silencing in diverse solanaceous species, *Plant Cell Rep*. **2012**, 31, 1713-1722.
[34] D. Weigel, J. Glazebrook, Transformation of agrobacterium using the freeze-thaw method,

Cold Spring Harb. Protoc. **2006**, 2006,1031-1036.
[35] M. Viau, J. Gastaldo, Z. Bencokova, A. Joubert, N. Foray, Cadmium inhibits non-homologous end-joining and over-activates the MRE11-dependent repair pathway, *Mutat. Res*. **2008**, 654, 13-21.
[36] A. Cilla, M.J. Rodrigo, L. Zacarías, B. De Ancos, C. Sánchez-Moreno, R. Barberá, A. Alegría, Protective effect of bioaccessible fractions of citrus fruit pulps against H_2O_2-induced oxidative stress in Caco-2 cells, *Food Res. Int.* **2018**, 103, 335-344.
[37] D. de Paula Borges, A.W.A. dos Santos, C.R.K. Paier, H.L.R. Júnior, M.B. Costa, I.R. Farias, R.T.G. de Oliveira, I.G. da Frota França, G.M. Cavalcante, S.M.M. Magalhães, Prognostic importance of aurora kinases and mitotic spindle genes transcript levels in myelodysplastic syndrome, *Leuk. Res*. **2018**, 64, 61-70.
[38] W. Liu, L. Sun, M. Zhong, Q. Zhou, Z. Gong, P. Li, P. Tai, X. Li, Cadmium-induced DNA damage and mutations in Arabidopsis plantlet shoots identified by DNA fingerprinting, *Chemosphere.* **2012**, 89, 1048-1055.
[39] I. Brandsma, D.C. van Gent, Pathway choice in DNA double strand break repair: observations of a balancing act, *Genome integrity.* **2012**, 3, 9.
[40] A. Ciccia, S.J. Elledge, The DNA damage response: making it safe to play with knives, *Mol. Cell.* **2010**, 40, 179-204.
[41] P. Karran, M. Bignami, DNA damage tolerance, mismatch repair and genome instability, *Bioessays.* **1994**, 16, 833-839.
[42] J. Bartek, J. Lukas, DNA damage checkpoints: from initiation to recovery or adaptation, Curr. Opin. *Plant Biol*. **2007**, 19, 238-245.
[43] W. Yang, R. Woodgate, What a difference a decade makes: insights into translesion DNA synthesis, *Proc. Natl. Acad. Sci. USA*. **2007**, 104, 15591-15598.
[44] M. Seifert, S.J. Scherer, W. Edelmann, M. Böhm, V. Meineke, M. Löbrich, W. Tilgen, J. Reichrath, The DNA-mismatch repair enzyme hMSH2 modulates UV-B-induced cell cycle arrest and apoptosis in melanoma cells, *J. Invest. Dermatol*. **2008**, 128, 203-213.
[45] M.L. Sottile, A.D. Losinno, M.A. Fanelli, F.D. Cuello-Carrión, M.M. Montt-Guevara, L.M. Vargas-Roig, S.B. Nadin, Hyperthermia effects on Hsp27 and Hsp72 associations with mismatch repair (MMR) proteins and cisplatin toxicity in MMR-deficient/proficient colon cancer cell lines, *Int. J. Hyperth*. **2015**, 31, 464-475.
[46] V. O'Brien, R. Brown, Signalling cell cycle arrest and cell death through the MMR System, *Carcinogenesis.* **2005**, 27, 682-692.
[47] Y. Situ, L. Chung, C.S. Lee, V. Ho, MRN (MRE11-RAD50-NBS1) complex in human cancer and prognostic implications in colorectal cancer, *Int. J. Mol. Sci*. **2019**, 20, 816.
[48] Y. Wang, D. Cortez, P. Yazdi, N. Neff, S.J. Elledge, J. Qin, BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures, *Genes & Dev*. **2000**, 14, 927-939.
[49] A. Datta, R.M. Brosh Jr, Holding All the Cards - How Fanconi Anemia Proteins Deal with Replication Stress and Preserve Genomic Stability, *Genes*. **2019**, 10, 170.

Figure captions

Figure 1. Cd stress induced DNA damage in LD10 and SN20 seedling root tips when grown under 0-2.5 mg·L^{-1} Cd stress for 4 days. (a) RAPD polymorphism variation; (b) The GTS; (c-f) The relative expression level of DNA damage repair genes. Gene expression levels of the LD10 under control conditions were set to 1 as the normalization in qRT-PCR analysis. Standard deviations were calculated with three independent experiments. Different letters indicate statistically significantly differences ($P < 0.05$).

Figure 2. FCM analysis of the nuclear DNA contents of soybean genotypes LD10 (a) and SN20 (b) seedling roots under Cd stress for 4 days. The percentage of 2C and 4C nuclear content cells in the total cell population was calculated. Standard deviations were calculated with three independent experiments. Different letters indicate statistically significant differences ($P < 0.05$).

Figure 3. Relative gene expression levels in root tips of LD10 and SN20 exposed to 0-2.5 mg·L^{-1} Cd for 4 days. In a-h, G1/S phase transition regulation / marker genes *PCNA1*, *E2FA*, *HISTONE H4*; DNA damage response genes *ATM*, *ATR*; G2/M phase transition regulation / marker genes *CYCB1;1*, *CDKA;1*, *WEE1*. Gene expression levels of the LD10 seedling root tips under normal condition were set to 1 as the normalization for qRT-PCR analysis. Standard deviations were calculated with three independent experiments. Different letters

873 indicate statistically significant differences ($P < 0.05$).

874

875 **Figure 4.** Relative gene expression levels of *MLH1* (a), *MSH2* (b), and *MSH6*
876 (c) genes in root tips of LD10 and SN20 seedling exposed to 0-2.5 $mg \cdot L^{-1}$ Cd
877 for 4 days. Gene expression levels of the LD10 seedling roots under normal
878 condition were set to 1 as the normalization in the qRT-PCR analysis. Standard
879 deviations were calculated with three independent experiments. Different letters
880 indicate statistically significant differences ($P < 0.05$).

881

882 **Figure 5.** VIGS induced soybean MMR gene silencing. Relative gene
883 expression levels of *MLH1* (a), *MSH2* (b), and *MSH6* (c) genes in root tips of
884 TRV-based MMR gene silencing soybean plantlet under normal culture
885 conditions. Gene expression levels of the uninfected LD10 seedling roots were
886 set to 1 as the normalization in the qRT-PCR analysis. Standard deviations
887 were calculated with three independent experiments. Different letters indicate
888 statistically significant differences ($P < 0.05$).

889

890 **Figure 6.** RAPD polymorphism variations in roots of TRV-based gene silencing
891 soybean plantlets exposed to 0.5$mg \cdot L^{-1}$ Cd for 4 days. For all treatments,
892 reproducible bands in at least two replicates were evaluated and calculated for
893 polymorphism analysis.

894

895 **Figure 7.** FCM analysis of the nuclear DNA contents of soybean genotypes

LD10 (a) and SN20 (b) seedling roots under 0-0.5 mg·L^{-1} Cd stress for 4 days. The percentage of 2C and 4C nuclear content cells in total cells was calculated， respectively. Standard deviations were calculated with three independent experiments. Different letters indicate statistically significant differences ($P < 0.05$). * indicate statistically significant differences ($P < 0.05$)

Figure 8. Relative gene expression levels of DNA damage repair genes in seedling roots of LD10 and SN20 genotypes exposed to 0.5 mg·L^{-1} Cd for 4 days. Dashed line indicate gene expression levels of LD10 seedling roots grown under control conditions were set to 1 as the normalization in qRT-PCR analysis. Standard deviations were calculated with three independent experiments. Different letters indicate statistically significant differences ($P < 0.05$) in a to h.

Figure 9. Mechanism of Cd tolerance in soybean SN20 and LD10 cultivars. SN20 and LD10 were respectively Cd-tolerant and Cd-sensitive soybean cultivars selected by Cd-induced root growth repression. In wild type SN20 *MSH2* and *MSH6* are more highly expressed and are more responsive to Cd-stress than in LD10. This causes G_2/M phase arrest in SN20 but G_1/S phase arrest in LD10 under Cd stress. G_2/M phase arrest in SN20 allows DNA replication leading to cell volume enlargement and some cell proliferation with post-replication repair, but G_1/S phase arrest in LD10 does not. This explains the different Cd-induced root growth repression in SN20 and LD10. This

hypothesis was tested by knocking down *MSH2* or *MSH6* from SN20, and by knocking down *MLH1* from LD10 increasing G_1/S phase arrest through blocking the MRN complex-to-ATM signaling.

Figure 10. Multiple effects of MMR system on Cd tolerance in soybean. MSH2 and MSH6 forming MutSα regulate the cell cycle by activating the G_2/M checkpoint, leading to cell cycle arrest, when they recognize Cd-induced DNA damage. Furthermore, MutSα can recruit MutL and BRCA1/Rad51 to trigger MMR and HR mediated repair. When *MSH2* or *MSH6* was knocked down, HR repair would be repressed causing increased SSBs and DSBs. Knocking down *MLH1* would suppress MMR, but improve HR repair because of reduced G_1/S arrest leading to enhanced post-replication repair in increased G_2 phase.