Mol Biotechnol (2011) 47:120–129
DOI 10.1007/s12033-010-9319-9

RESEARCH

High-Level Production of MSH2 from *Arabidopsis thaliana*: A DNA Mismatch Repair System Key Subunit

Rodrigo L. Gomez · Celina Galles · Claudia P. Spampinato

Published online: 5 August 2010

Abstract Biochemical and immunological information concerning DNA mismatch repair proteins from higher plants is currently limited, probably due to their low abundance in vivo. An initial analysis of At*MSH2* gene expression by quantitative real-time RT-PCR indicates that calli and seedlings contain 96.7 and 1.4 cDNA copies per ng RNA, respectively, confirming that this gene is predominantly expressed in rapidly dividing tissues. In order to obtain large quantities of AtMSH2, the protein was efficiently expressed in an *Escherichia coli* system. The expressed gene product has an in-frame N-terminal Trx-His_6-S-tag. The fusion protein represents about 11% of the soluble protein from IPTG-induced *E. coli* cells. After a two-step purification procedure the final yield accounts for 0.7 mg/g cells. Digestion of this electrophoretically homogeneous recombinant protein with enterokinase results in an intact protein with only one extra amino acid introduced at the N-terminal end. Purified intact protein was used to induce polyclonal antibodies in rabbits. These antibodies cross-react with a 110-kDa protein from cauliflower inflorescences. Together, our data describe the transcript level, cloning, expression, purification, and polyclonal antibody preparation of AtMSH2. This work will surely be useful for carrying out plant mismatch repair assays in vitro and analyzing protein expression after the exposure of plants to various stresses.

Keywords MSH2 · Heterologous expression · *E. coli* · Polyclonal antibody

R. L. Gomez · C. Galles · C. P. Spampinato (✉)
Facultad de Ciencias Bioquímicas y Farmacéuticas, Centro de Estudios Fotosintéticos y Bioquímicos (CEFOBI), Universidad Nacional de Rosario, Suipacha 531, 2000 Rosario, Argentina
e-mail: spampinato@cefobi-conicet.gov.ar

Abbreviations

MMR	Mismatch repair
IPTG	β-Isopropyl-D-thiogalactoside
DTT	Dithiothreitol
MSH	MutS homolog
MLH	MutL homolog
PMS	Post-meiotic segregation

Introduction

The mismatch repair (MMR) system is a DNA repair pathway essential for the correct maintenance of genetic information across many generations. The system is best known for its role in the correction of base substitutions and small insertion/deletion loops (IDL) generated during every round of replication. In addition, the MMR system is also implicated in the recognition of modified bases generated in response to various physical, physiological, and environmental factors (see Refs. [1–5] for recent reviews). In particular, plant genomes are constantly challenged with major mutagenic agents, such as solar UV and reactive oxygen species generated by mitochondrial and chloroplastic electron transport chains. Even worse, plants are sessile organisms unable to escape from environmental genotoxins and lack a reserved germline. Therefore, if DNA lesions are not repaired, non-lethal mutations accumulate and increase the amount of genetic variation in the progeny.

The MMR system is highly conserved across evolution [2, 4, 6, 7]. The first step of the pathway involves recognition of a mismatch by MutS homodimers in bacteria or MutS homologs (MSH) heterodimers in eukaryotes. Human and yeast MSH2–MSH6 heterocomplexes (MutSα)

recognize base–base mispairs and small IDL, whereas MSH2–MSH3 heterodimers (MutSβ) participate in the repair of IDL heterologies [8–11]. In addition, plants contain an additional complex implicated in mismatch correction, MSH2–MSH7 (MutSγ) [12, 13]. These three heterodimers have been partially characterized in *A. thaliana* by an in vitro transcription and translation technique. AtMutSα and AtMutSβ heterodimers show a similar recognition pattern as their eukaryotic counterparts, whereas AtMutSγ preferentially binds some base–base mispairs [12, 13]. Apart from mismatch correction, MutSγ also plays an important role during meiosis [14]. Other information about the function of plant MMR proteins has been obtained by monitoring microsatellite instability or recombination events of *msh2*, *mlh1*, or *pms1* deficient plants (see [15] for a recent review). However, information on the biochemistry and biophysics of the MMR system in higher plants is still limited, probably due to their low in vivo abundance and the lack of an efficient large-scale strategy for heterologous protein production.

In this work, we first evaluate the expression of the At*MSH2* gene in calli and seedlings. We then report a high-level expression of AtMSH2 in *Escherichia coli*, purification of this subunit to homogeneity and generation of polyclonal antibodies. These antibodies cross-react with a 110-kDa protein from nuclear extracts prepared from cauliflower inflorescences. Recombinant AtMSH2 protein can be useful for a further reconstitution of plant MMR in vitro while the antibody can provide a tool for the study of protein expression and distribution of AtMSH2 in various tissues after the exposure of plants to different stresses.

Materials and Methods

Chemicals and Reagents

Zero Blunt TOPO PCR Cloning Kit was purchased from Invitrogen. IPTG was obtained from Promega. Restriction enzymes were products from New England Biolabs or Promega. Recombinant enterokinase was obtained from New England Biolabs. Monoclonal anti-His antibody, HisTrap HP columns, and HiTrap Q FF columns were purchased from Amersham Biosciences. Anti-mice and anti-rabbit IgG antibodies conjugated to alkaline phosphatase were obtained from Bio-Rad. All other chemicals and reagents were of molecular biology grade.

Plant Material and Growth Conditions

Arabidopsis thaliana ecotype Columbia (Col-0) seeds were surface sterilized and sown on Murashige and Skoog basal salt (MS) [16] solid medium, containing Phytagel as a solidifying agent. After 3 days of vernalization at 4 °C in the dark, plants were grown in a growth chamber under a 16 h light/8 h dark regime.

Calli were obtained according to procedures described by Mathur and Konz [17], except for: (1) MS medium was supplemented with 3% sucrose, 0.5 mg/L kinetine, and 2.5 mg/L 2,4-dichlorophenoxyacetic acid and (2) calli were derived from germinated seeds. Briefly, seeds were sown in MS medium. After vernalization and subsequent transfer to the growth chamber until germination, plaques were transferred to dark conditions at 22 °C. After approximately 1 month, development of 5 mm diameter calli from each individual seed was evident and from this point on, calli were sectioned into four fragments and subcultivated to fresh media once a month, when their diameter was roughly 1 cm (Fig. 1).

Bacteria Strains and Growth Conditions

Escherichia coli DH5α strain was used as host in cloning experiments, and *E. coli* BL21(DE3) and BL21(DE3)-Codon plus-RIL strains were used for protein expression. Bacteria were grown in Luria broth. Where appropriate, ampicillin (100 μg/mL), chloramphenicol (20 μg/mL), and/or agar 1% (w/v) were added.

Quantitative Analysis of AtMSH2 Transcript by Real-Time PCR

Steady-state mRNA levels were analyzed by quantitative real-time RT-PCR assays. Total RNA was isolated from about 100 mg of seedlings and calli using the TRIzol reagent (Invitrogen, Carlsbad, CA, USA) as described by the manufacture's protocol. Briefly, 100 mg of tissue was

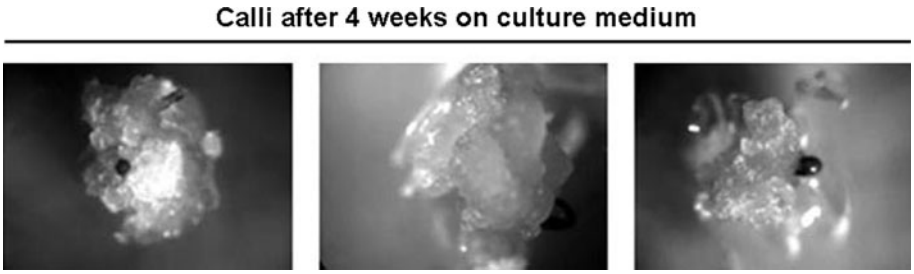

Fig. 1 Calli formation from germinated seeds after 4 weeks on MS medium supplemented with 3% sucrose, 0.5 mg/L kinetine, and 2.5 mg/L 2,4-dichlorophenoxyacetic acid

ground to a fine powder using liquid N_2 followed by addition of 1 mL TRIzol Reagent. After 5-min incubation at room temperature, 0.2 mL of chloroform was added. Tubes were vigorously shaken by hand, incubated 2 to 3 min at room temperature and centrifuged at 12,000*g* for 15 min at 4 °C. Following centrifugation, RNA from the aqueous phase was precipitated using 0.5 mL of isopropyl alcohol. Finally, RNA was dissolved in RNase-free water and treated with RNase-free DNase I (1 U/mL) to remove traces of genomic DNA. Then, 4 μg RNA was converted into first-strand cDNA using SuperScript II reverse transcriptase (Invitrogen) and oligo-dT as a primer in a final volume of 20 μL. After reverse transcription, 1 μL of a 1:3 dilution of the cDNA obtained was used as template for real-time PCR with a Mx 3000P QPCR System (Stratagene, La Jolla, CA, USA) using a reaction mixture containing 332.5 nM forward (TGGTAACACTGTTGG TGTGG) and reverse (TTCTCTTCCCACTCTCCTCA) specific primers spanning a 242-bp region of At*MSH2* cDNA, 20 mM Tris–HCl, pH 8.5, 50 mM KCl, 0.5× SYBR Green I (Invitrogen, Carlsbad, CA, USA), 3.5 mM $MgCl_2$, 0.23 mM dNTPs, and 0.58 U Platinum Taq DNA polymerase in a total volume of 20 μL. The PCR thermal cycling conditions were 2 min denaturation at 94 °C; 50 cycles at 96 °C for 10 s, 57 °C for 15 s, and 72 °C for 20 s, followed by 1 min extension at 72 °C. The correct molecular weights of the amplification products were verified by 2% agarose gel electrophoresis and melting curve analysis.

Quantification of At*MSH2* mRNA levels in plant tissues was estimated by comparing the threshold cycle (C_t) of each sample with a calibration curve. C_t is the number of required cycles to reach the defined threshold level. The C_t values are proportional to the logarithms of the initial target concentrations. A calibration curve was generated by plotting C_t values of dilutions of the standard versus the corresponding concentrations expressed as log moles/20 μL reaction. Standards were generated by serial 10-fold dilutions of the pET32b-AtMSH2 plasmid stock solution. Based on the molecular weight of the cloned DNA, the concentration of the template applied in each reaction ranged from 2.4×10^{-10} to 2.4×10^{-5} pmol. The initial At*MSH2* copy number in the samples was then calculated according to Avogadro's number (6×10^{23}). All reactions were performed in triplicate from two independent sets of biological replicates and each real-time PCR experiment was conducted at least twice.

Vector Constructions

The At*MSH2* cDNA inserted into plasmids pGEM-3Z was generously provided by Dr. John Hays [12]. This cDNA was then used as template for PCR reactions with gene specific primers (forward primer AtM2-5PNco: 5′-GCAGTCC**ATG** GAGGGTAATTTCGAGGAACAGAA-3′; reverse primer AtM2-3PBam: 5′-CTACGGATCC**TTA**TCACAGAAACT GCCTGAGCC-3′. The translation start and stop codons are written in bold face while the *Nco*I and *Bam*HI recognition sites are underlined). The PCR amplification reaction contained 10 ng of pGEM3Z-AtMSH2, 2 μM each gene specific primer, 200 μM dNTPs, 0.04 units of Vent DNA polymerase, 20 mM Tris–HCl, pH 8.8, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, and 0.1% Triton x-100 in a final volume of 50 μL. The amplification procedure included an initial denaturation for 5 min at 94 °C followed by a touch-down PCR consisting of 30 s denaturation at 94 °C, 20 s annealing starting at 70 °C, and 3 min extension at 72 °C. The annealing temperature was lowered 2 °C every second cycle until it reached 60 °C. After the touch down, the annealing temperature was held at 60 °C for another 30 cycles. The final extension step was at 72 °C for 7 min. The *Nco*I and *Bam*HI restriction sites introduced by the primers allowed the amplified product to be cloned into the equivalent sites of the pET32b expression vector (Novagen). In this construction, shown in Fig. 2, the AtMSH2 coding sequence is in frame with the *N*-terminal Trx-His_6-S-tag (154 amino acids) and only one additional amino acid at the *N*-terminal end of the protein remains after enterokinase digestion.

Expression and Purification of AtMSH2

Overexpression of AtMSH2 was performed in BL21(DE3) and BL21(DE3)-Codon plus-RIL *E. coli*. The latter strain contains extra copies of *argU*, *ileY*, and *leuW* tRNA genes, which encode the RIL codons (AGA/AGG_{Arg}, AUA_{Ile}, and CUA_{Leu}). Bacteria were transformed according to standard electroporation procedures with recombinant pET32b-AtMSH2 plasmid. Following transformation, a single colony was inoculated into 5 mL LB containing ampicillin (100 μg/mL) and chloramphenicol (20 μg/mL) for the

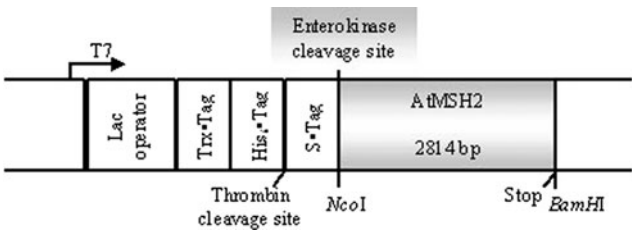

Fig. 2 The pET32b-AtMSH2 construct used for gene overexpression. The coding region of AtMSH2 was amplified by PCR and subcloned into the *Nco*I and *Bam*HI sites of the pET32a expression vector (Novagen) downstream of a Trx-His_6-S-tag under the control of the T7 RNA polymerase promoter (T7). Details of the cloning procedure are described under "Materials and methods". Cleavage sites of thrombin and enterokinase are also shown

strains carrying the particular plasmids. This culture was incubated overnight at 37 °C with vigorous shaking and afterward used to inoculate 200 mL LB medium containing antibiotics (sufficient overnight culture to provide an initial $OD_{600nm} \approx 0.05$). This cell culture was grown at 30 °C with vigorous shaking until the OD_{600nm} reached 0.8–1.0. At this point, uninduced control aliquots were collected and subsequently, IPTG was added to a final concentration of 0.1 mM to allow AtMSH2 expression. The culture was then grown for additional 16 h at 16 °C. Cells were centrifuged at 5,000*g* for 20 min at 4 °C and cell pellet was immediately stored at −70 °C.

Bacteria pellets were resuspended in either lysis buffer containing 25 mM HEPES–KOH, pH 7.5, 0.1 mM EDTA, 1 mM DTT, 1 mM PMSF or binding buffer containing 20 mM Tris–HCl, 30 mM imidazole, 0.5 M NaCl, pH 7.4 (15 mL/g cell wet weight). When indicated cells were disrupted by sonication and then centrifuged at 7,000*g* for 20 min at 4 °C to remove cell debris. Proteins of the uninduced and IPTG-induced cultures (both soluble and insoluble fractions) were analyzed by 8% (w/v) SDS-PAGE and visualized with Coomassie Brilliant Blue stain. In addition, strains carrying the empty pET32b vector were also IPTG-induced and included as controls.

AtMSH2 protein purification was carried out with the soluble fractions. The supernatant was filtered through a 0.2 μm cellulose acetate membrane and loaded onto a HisTrap HP affinity column previously equilibrated with binding buffer. The column was washed with 5 volumes of the same buffer to remove proteins associated non-specifically to the matrix. Imidazole concentration was then increased in the first place to 100 mM to wash away other impurities and finally to 500 mM to allow the elution of bound AtMSH2 protein. Protein fractions were pooled and dialyzed overnight at 4 °C against TG buffer (20 mM Tris–HCl buffer, pH 8.0, 5% (v/v) glycerol) containing 200 mM NaCl and subjected to a HiTrap Q FF column previously equilibrated with the same buffer. After washing with 5 column volumes of TG buffer containing 250 mM NaCl, bound AtMSH2 was eluted increasing the salt concentration up to 300 mM. Fractions containing pure AtMSH2 were pooled, brought to 500 mM NaCl and 30 mM imidazole and concentrated using a HisTrap HP affinity column. Aliquots of the various fractions were analyzed by SDS-PAGE and subsequent staining.

Fractions containing AtMSH2 recombinant protein were pooled, desalted by gel filtration chromatography on Sephadex G-50 (Sigma-Aldrich Co.) previously equilibrated with TG buffer and digested with recombinant enterokinase to remove the *N*-terminal tag. The standard cleavage reaction contained 20 mM Tris–HCl, pH 8.0, 50 mM NaCl, 2 mM $CaCl_2$, 5% glycerol, 1.5 ng enterokinase, and 50 μg of purified AtMSH2 protein, in a total volume of 50 μl. Enzymatic digestion was performed at 23 °C for 16 h.

Antibody Generation and Purification

Rabbit polyclonal antibodies against AtMSH2 protein were generated by standard immunization protocols. Following preimmune serum collection, 200 μg of purified denatured recombinant protein emulsified with equal volume of Freund's complete adjuvant was injected subcutaneously into a 4-month-old white rabbit. Three weeks after the initial immunization, a boost with 100 μg of the same protein was injected. Antiserum was collected after 25–28 days and subsequently AtMSH2 antibodies were affinity purified according to Plaxton's procedure [18]. Concisely, after enterokinase digestion (see above) recombinant AtMSH2 was subjected to 8% (w/v) SDS-PAGE and subsequently transferred onto a nitrocellulose membrane. The membrane was reversibly stained with Ponceau S to excise the band corresponding to AtMSH2. The membrane section was then blocked for 1 h at 4 °C with 5% (w/v) skim milk powder in TBS buffer containing 10 mM Tris–HCl buffer, pH 7.5, and 150 mM NaCl. Then, the membrane section was excised in smaller fragments and all these fragments were incubated in an Eppendorf tube with 0.5 mL of the crude antiserum for 1 h at 4 °C. Following three 5 min washes with TBS buffer, bound antibodies were eluted with 500 μL of 100 mM glycine–HCl, pH 3.0 and 100 mM NaCl. After 1 min incubation, eluted antibodies were rapidly buffered to pH 7.5 with 50 μL of 1 M Tris–HCl, pH 8.0. The antibody elution step was repeated 3 times and the whole affinity purification procedure was performed four times in all. Finally, the antibodies were pooled, brought to 0.02% (w/v) NaN_3, and stored at −80 °C.

Protein Extracts

Nuclear isolation was performed using cauliflower inflorescences as an abundant source of proteins from an Arabidopsis closely related species essentially as described by Bowler et al. [19]. Briefly, cauliflower buds were ground in a blender with extraction buffer (4 mL/g) containing 0.4 M sucrose, 25 mM HEPES–KOH, pH 7.6, 10 mM $MgCl_2$, 5 mM β-mercaptoethanol, and 0.2 mM PMSF. The homogenate was filtered through 4 layers of Miracloth and centrifuged at 3,000*g* for 20 min at 4 °C. The nuclear pellets were gently resuspended in homogenization buffer (1 mL/g) containing 0.25 M sucrose, 10 mM HEPES–KOH, pH 7.6, 10 mM $MgCl_2$, 5 mM β-mercaptoethanol, 0.15% (v/v) Triton x-100, 0.2 mM PMSF and centrifuged at 2,200*g* for 20 min at 4 °C. Nuclear pellets were washed

three times with homogenization buffer with subsequent centrifugations. Chromatin fractionation was modified as previously reported [20]. The resulting pellet was resuspended in a minimum volume of buffer containing 0.3 M sucrose, 10 mM HEPES–KOH, pH 7.0, 100 mM NaCl, 3 mM $MgCl_2$, 5 mM β-mercaptoethanol, 0.5% (v/v) Triton x-100, 0.1 mM PMSF. After 5–10 min on ice, the nuclear suspension (0.04 mL/g of initial tissue) was centrifuged at 1,400*g* for 10 min to separate soluble proteins from the chromatin fraction. The resulting pellets were resuspended in SDS-PAGE loading buffer and boiled. Alternatively, pellets were resuspended in 0.02 mL of isolation buffer containing 50 mM Tris–HCl, pH 7.6, 1% (w/v) SDS, 5 mM β-mercaptoethanol, 1 mM PMSF followed by sonication at 30% amplitude. After 4 cycles of 20-s pulses, the sheared chromatin was clarified by centrifugation.

Total soluble cauliflower extracts were prepared by grinding 1 g of tissue in liquid nitrogen followed by the addition of 3 mL of isolation buffer. Cell debris was removed by centrifugation and the supernatant was used for protein analysis.

SDS-PAGE and Immunological Analyses

Denaturing gel electrophoresis was carried out using 8% (w/v) polyacrylamide/bisacrylamide gels at 40 mA with the Bio-Rad Mini-Protean 3 system according to Laemmli [21]. Cell proteins were then stained with Coomassie brilliant blue (R250) or electrotransferred to nitrocellulose membranes (Bio-Rad) at 50 V for 1 h for immunoblotting [22]. Either the commercially available anti-His antibodies (1:3,000 dilution) or the affinity purified anti-AtMSH2 polyclonal antibodies (6.5 µg/mL) described above were used for detection. Bound primary antibodies were recognized by rabbit anti-mice or goat anti-rabbit IgG conjugated to alkaline phosphatase and subsequently developed with 5-bromo-4-chloro-3-indolyl phosphate and nitro blue tetrazolium [22].

Bioinformatic Analyses

Published microarray expression data retrieved from the Genevestigator database were used to analyze AtMSH2 expression levels [23].

Analytical Methods

Protein concentration was quantified using Bradford reagent (Bio-Rad) and BSA as the protein standard [24]. Cell growth in liquid media was determined by measuring the absorbance at 600 nm using a UNICAM, HEλI0S β spectrophotometer.

Results and Discussion

The DNA mismatch repair system plays an essential role in preserving genome stability. In eukaryotes, MSH2 forms complexes with MSH6, MSH3, or MSH7 (this last being an exclusive MSH6 homolog reported only in plants). The MSH2–MSH6, MSH2–MSH3, or MSH2–MSH7 heterodimers, named as MutSα, MutSβ, and MutSγ, respectively, are required for the initial recognition of mismatched DNA [4, 6, 13]. Extensive information regarding MutSα and MutSβ heterodimers has been reported in other organisms, such as yeast [8, 11, 25, 26] and human [9, 10, 27, 28]. In plants, however, protein characterization and elucidation of the mechanism of action of this DNA repair system are still in progress. At present, studies in plants are restricted to the evaluation of *msh2*, *mlh1*, and *pms1* knockout mutations on microsatellite instability [29–32] and illegitimate recombination [33–37]. Here, we first analyzed *MSH2* gene expression by quantitative real-time RT-PCR in calli and seedlings from *A. thaliana*. Then, to achieve higher expression levels of soluble AtMSH2, we constructed a prokaryotic system to express the protein as a fusion to cleavable Trx-His_6-S-tag sequence in order to increase AtMSH2 solubility and facilitate protein detection on Western blots and purification by Ni^+ affinity chromatography. Only one attempt for AtMSH2 expression in *E. coli* has been reported with very low expression levels, barely detectable with antibodies directed against the His·Tag, and thus, the amounts obtained were insufficient for detailed biochemical analysis [38]. Our preparation yields large quantities of the protein that can be useful for future biochemical studies.

Analysis of AtMSH2 Transcript Levels

At*MSH2* mRNA expression was evaluated by quantitative real-time RT-PCR. Previous reports indicate that *MSH2* genes are expressed at very low levels in *A. thaliana* and maize as investigated by Northern blot [39, 40]. Here, the starting copy number of the At*MSH2* gene was quantified in calli and seedlings using standard curves generated with the pET32b-MSH2 plasmid as template. Based on the known molecular weight of the plasmid, the amount of template applied in each reaction ranged from 2.4×10^{-10} to 2.4×10^{-5} pmol (126 to 1.26×10^7 copies, respectively). As shown in Fig. 3, the expression of MSH2 was 70-fold higher in calli than in seedlings (96.7 and 1.4 cDNA copies per ng RNA, respectively). These data would be indicating that *MSH2* gene transcription is strongly required to ensure genomic stability in callus tissue (Fig. 1), where active cell division and often endoreduplication [41, 42] and somaclonal variation are known to be taking place [43, 44]. On the other hand, *MSH2* mRNA

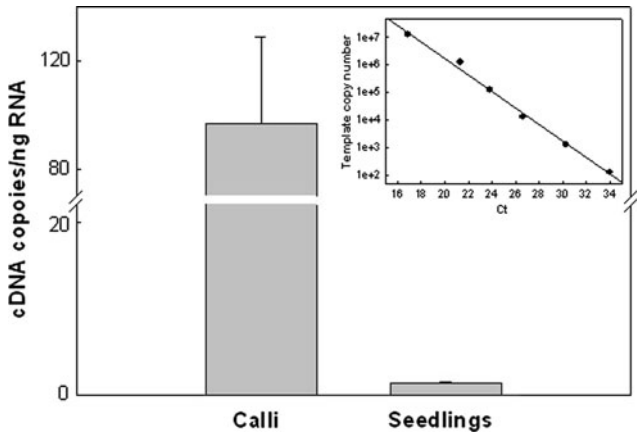

Fig. 3 Expression profiles of At*MSH2* gene. Absolute copy numbers of At*MSH2* transcript in calli and seedlings. *Inset* Calibration curve performed for real-time RT-PCR experiments using serial dilutions of the pET32b-AtMSH2 plasmid as template DNA. Based on the molecular weight of the cloned DNA, the concentration of the template applied in each reaction ranged from 2.4×10^{-10} to 2.4×10^{-5} pmol (126 to 1.27×10^{7} copies, respectively) per 20 μL reaction

levels are not as high in total seedling as in calli, since DNA replication is not as frequent. We then compared these results to the signal intensity values from microarray experiments available in the GENEVESTIGATOR Microarray Database [23]. According to this database, higher transcript levels are found in calli than in seedlings. Thus, both RT-PCR experiments and Genevestigator data show the same biologically relevant trend in At*MSH2* gene expression in rapidly dividing tissues. This is consistent with previous reports indicating higher levels of MSH2 activity in actively dividing cells from *A. thaliana* [39], *Zea mays* [40], and *Solanum lycopersicum* [45] as compared to cells in mature tissues.

Overexpression and Purification of the Recombinant AtMSH2

The coding region of *A. thaliana* MSH2 was cloned in pET32b taking advantage of the unique *Nco*I and *BamH*I restriction sites present both in the vector and in the cDNA. The resulting construction named pRG01 encodes a MSH2 fusion protein carrying 154 additional amino acids of the Trx-His_6-S-tag sequence in its *N*-terminal region (Fig. 2). This construction was used to transform different *E. coli* strains and expression of MSH2 was analyzed by SDS-PAGE as shown in Fig. 4a. Low-level expression of the protein was observed in BL21(DE3) strains representing

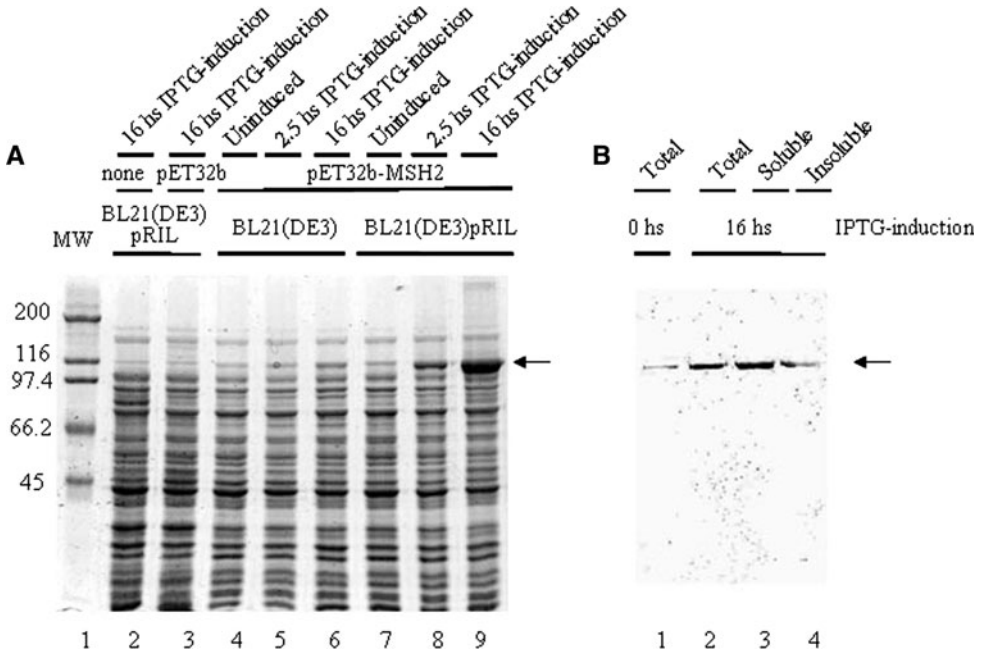

Fig. 4 Expression analysis of pET32b-AtMSH2. **a** SDS-PAGE electropherogram stained with Coomassie brilliant blue. Cell extracts from BL21(DE3) and BL21(DE3)-pRIL *E. coli* strains are indicated above the respective lanes. *Lane 1*: molecular weight markers the values of which indicated in kDa are shown on the left side of the SDS-PAGE; *lane 2*: BL21(DE3)-pRIL strain induced with IPTG; *lane 3*: BL21(DE3)-pRIL strain carrying the empty plasmid pET32b induced with IPTG; *lanes 4* and *7*: uninduced total fraction; *lanes 5* and *8*: total fractions after 150 min induction with IPTG; *lanes 6* and *9*: total fractions after 16 h induction with IPTG. The *arrow* indicates the position of the overexpressed fusion protein at the expected position (122 kDa). **b** Western blot analysis of BL21(DE3)-pRIL cell extracts using monoclonal antibodies directed against the N-terminal poly-histidine tag. *Lanes 1* and *2*: total fractions before and after 16 h induction with IPTG. *Lanes 3* and *4*: soluble and insoluble fractions recovered after sonication of the overnight induced cultures. Approximately 25 μg of total proteins were applied per lane, except for the insoluble fraction where 6 μg of total proteins were loaded. The *arrow* indicates the position of the overexpressed AtMSH2 fusion protein

4% of total cell proteins when induced for 16 h at 16 °C using 0.1 mM IPTG. In contrast, co-expression of the RIL plasmid in the host cells allowed higher level of protein production. This plasmid has been reported to improve expression of several heterologous recombinant genes [46] since it encodes extra copies of certain tRNAs that recognize the rarely used AGA/AGG (arginine), AUA (isoleucine), and CUA (leucine) codons in *E. coli*. These codons are poorly represented in *E. coli* MutS (0.46%), whereas they are commonly used in AtMSH2 cDNA (6.8%). As shown in Fig. 4a, lane 9, BL21(DE3)-pRIL expressed the recombinant protein after 16 h induction with 0.1 mM IPTG at 16 °C. The fusion protein expressed was about 115 kDa, which was in good accordance with the molecular weight deduced from the amino acid sequence of the construct (105.4 kDa) fused to the *N*-terminal tags (17.1 kDa). Neither BL21(DE3)-pRIL strain nor this strain transformed with the empty pET32b plasmid (Fig. 4a, lanes 2 and 3, respectively) showed signal after IPTG induction. After sonication, most of the fusion proteins remained in the soluble fraction. This was detected both by SDS-PAGE and Western blot using monoclonal antibodies directed against the His·tag (Fig. 4b). Typically, AtMSH2 expression yielded as much as 11% of total soluble protein from *E. coli*.

Recombinant AtMSH2 fusion protein from soluble *E. coli* extracts was then applied to a HisTrap HP affinity column and eluted with binding buffer containing 500 mM imidazole. Fractions containing MSH2 were pooled and dialyzed overnight against 100 volumes of TG buffer containing 200 mM NaCl. Finally, the protein was further purified using an anion exchange chromatography HiTrap Q FF column. Bound protein was eluted at 300 mM NaCl. Aliquots of the various fractions were analyzed by SDS-PAGE (Fig. 5). Typically, 14 mg of AtMSH2 fusion protein was obtained per liter of culture with a 99% estimated purity. Table 1 shows a summary of the purification procedure.

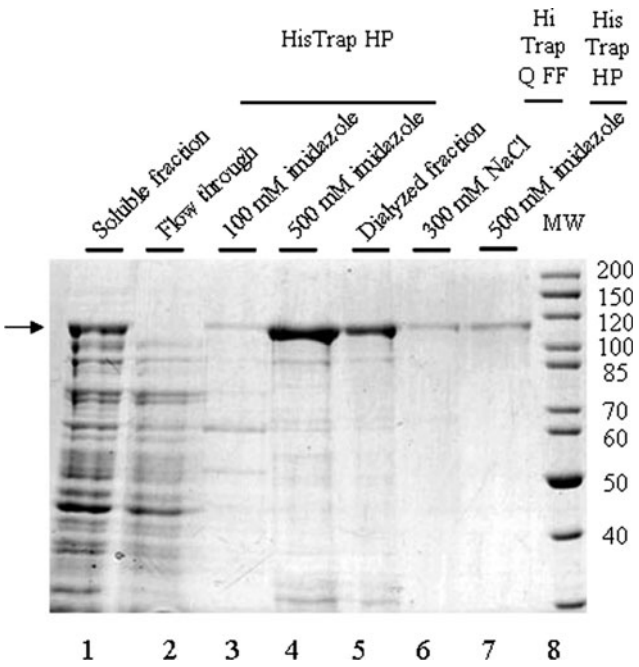

Fig. 5 Purification of AtMSH2 fusion protein. Various purification fractions were analyzed on SDS-PAGE stained with Coomassie brilliant blue. *Lane 1*: soluble fraction recovered after sonication of overnight induced cultures. *Lane 2*: flow-through fraction from the HisTrap HP column. *Lane 3* and *4*: 100 and 500 mM imidazole eluted fractions from the HisTrap HP column, respectively. *Lane 5*: lane 4 dialyzed protein fraction. *Lane 6*: pooled fraction eluted at 300 mM NaCl from the HiTrap Q FF column. *Lane 7*: concentrated pure protein eluted at 500 mM imidazole from a HisTrap HP column. *Lane 8*: molecular weight markers the values of which indicated in kDa are shown on the right side of the SDS-PAGE. The *arrow*: indicates the position of the fusion protein

Immunological Studies

Highly purified AtMSH2 fusion protein was digested with enterokinase to remove the *N*-teminal tags and a protein of about 105 kDa was obtained (Fig. 6a, b, lanes 2 and 1, respectively). This molecular mass corresponds to the intact *A. thaliana* protein. As indicated, the pET32-MSH2 construct was designed in such a way that after enterokinase digestion only one extra amino acid was introduced at the *N*-terminal of AtMSH2. This preparation was used to generate antibodies against the plant protein in rabbits. It should be noted that the generation of antibodies specific to AtMSH2 must rely on recombinant protein technology given the extremely low abundance of this protein across all organisms. As a consideration, different MMR-proficient human cell lines only contain between 60 and 300 fmol of hMSH2 [47].

Anti-AtMSH2 immune serum was affinity purified by the procedure described in "Materials and methods". Purified antibodies clearly detect the recombinant intact (Fig. 6a, b, lanes 2 and 1, respectively) and fusion (Fig. 6a, b, lanes 3 and 2, respectively) proteins on Western blots. No signal was detected with preimmune serum (results not shown). Purified recombinant AtMSH2 antibodies cross-react with a protein of approximately 110 kDa using nuclear extracts from cauliflower (*Brassica oleracea*) inflorescences (Fig. 6a, lanes 1 and 4; Fig. 6b, lane 4). Cauliflower was chosen as a rich source of protein expressed in meristematic tissues from a species closely related to Arabidopsis.

Nuclear extracts were recovered after incubation of isolated nuclei with a physiological strength buffer containing non-ionic detergent (0.5% Triton x-100). Under these extraction conditions, several soluble (Fig. 6c, lane 1) or chromatin associated (Fig. 6c, lane 2) proteins were

Table 1 Purification of recombinant AtMSH2

Step	Volume (mL)	Protein concentration (mg/mL)	Total protein (mg)	MSH2 amount[a] (mg)	Yield (%)	Purification (fold)
Crude extract	20.0	0.90	18.0	2.3	100	1
HisTrap	4.0	0.60	2.4	2.1	91	7
HiTrap Q FF	9.0	0.13	1.2	1.2	52	8
HisTrap	1.5	0.53	0.8	0.8	35	8

[a] AtMSH2 amount was measured by densitometry

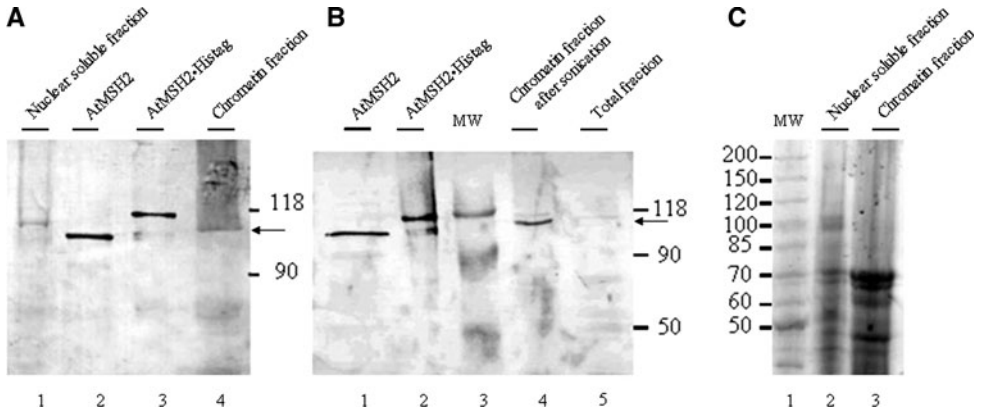

Fig. 6 Reactivity of the anti-AtMSH2 antibody to recombinant and native protein. **a** and **b** Western blot using affinity purified polyclonal antibodies directed against purified AtMSH2 protein. **a** *Lane 1*: nuclear soluble proteins from cauliflower buds. *Lanes 2* and *3*: purified recombinant fusion AtMSH2 protein with and without enterokinase treatment, respectively. *Lane 4*: chromatin-associated proteins from cauliflower buds. Values shown on the right side of the immunoblot correspond to prestained molecular weight markers. **b** *Lanes 1* and *2*: purified recombinant fusion AtMSH2 protein with and without enterokinase treatment, respectively. *Lane 3*: prestained molecular weight markers whose values indicated in kDa are shown on the right side of the immunoblot. *Lane 4*: sonicated chromatin fraction from cauliflower buds. *Lane 5*: total protein cell extracts from cauliflower buds. The arrow indicates the position of the putative MSH2 native protein from cauliflower buds. **c** SDS-PAGE stained with Coomassie brilliant blue. *Lane 1*: molecular weight markers whose values indicated in kDa are shown on the left side of the SDS-PAGE. *Lanes 2* and *3*: soluble proteins and chromatin fraction of nuclear extracts from cauliflower buds, respectively. Half as much material was loaded on the gel to be then stained with Coomassie brilliant blue relative to the immunoblot

released. Results of the immunoblotting indicate that the cauliflower nuclear extract contains a protein that specifically binds to antibodies against AtMSH2 protein in the soluble fraction (Fig. 6a, lane 1) and in the chromatin fraction resuspended directly in SDS-PAGE loading buffer (Fig. 6a, lane 4) or previously sonicated to solubilize and shear the chromatin (Fig. 6b, lane 4). Only a very faint immunoreactive band was observed with whole cell extracts (Fig. 6b, lane 5). The molecular weight of the putative cauliflower MSH2 roughly corresponds to the known values of other eukaryotic species, e.g. the molecular weight of human and yeast MSH2 are 105 and 109 kDa, respectively.

In summary, in the present study we first evaluated the expression of the At*MSH2* gene in calli and seedlings. We then successfully expressed the *A. thaliana* MSH2 protein as an *N*-terminal fusion to Trx-His_6-S-tag at high levels with good solubility in *E. coli*. The fusion protein was purified to near homogeneity by a two-step procedure. The intact protein obtained after removal of the peptide tag was used to generate polyclonal antibodies. These antibodies cross-reacted with a 110-kDa protein from cauliflower inflorescences. The protein and antibodies should be helpful for future biochemical and structural studies on plant DNA mismatch repair, such as the analysis of protein expression after the exposure of plants to various stresses, including endogenous and exogenous mutagens.

Acknowledgments We acknowledge research support from Fundación Antorchas and the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). CPS is a member of the Researcher Career of CONICET. RLG and CG are fellows of the same institution.

References

1. Hsieh, P., & Yamane, K. (2008). DNA mismatch repair: Molecular mechanism, cancer, and ageing. *Mechanisms of Ageing and Development, 129*, 391–407.
2. Iyer, R., Pluciennik, A., Burdett, V., & Modrich, P. (2006). DNA mismatch repair: Functions and mechanisms. *Chemical Reviews, 106*, 302–323.

3. Jiricny, J. (2006). The multifaceted mismatch-repair system. *Nature Review Molecular Cell Biology, 7*, 335–346.
4. Kunkel, T., & Erie, D. (2005). DNA mismatch repair. *Annual Review of Biochemistry, 74*, 681–710.
5. Li, G.-M. (2008). Mechanisms and functions of DNA mismatch repair. *Cell Research, 18*, 85–98.
6. Modrich, P. (2006). Mechanisms in eukaryotic mismatch repair. *Journal of Biological Chemistry, 281*, 30305–30309.
7. Marti, T., Kunz, C., & Fleck, O. (2002). DNA mismatch repair and mutation avoidance pathways. *Journal of Cellular Physiology, 191*, 28–41.
8. Johnson, R., Kovvali, G., Guzder, S., Amin, N., Holm, C., Habraken, Y., et al. (1996). Evidence for involvement of yeast proliferating cell nuclear antigen in DNA mismatch repair. *Journal of Biological Chemistry Review, 271*, 27987–27990.
9. Drummond, J. T., Li, G.-M., Longley, M. J., & Modrich, P. (1995). Isolation of an hMSH2–p160 heterodimer that restores DNA mismatch repair to tumor cells. *Science, 268*, 1909–1912.
10. Genschel, J., Littman, S., Drummond, J., & Modrich, P. (1998). Isolation of MutSβ from human cells and comparison of the mismatch repair specificities of MutSβ and MutSα. *Journal of Biological Chemistry, 273*, 19895–19901.
11. Alani, E. (1996). The *Saccharomyces cerevisiae* Msh2 and Msh6 proteins form a complex that specifically binds to duplex oligonucleotides containing mismatched DNA base pairs. *Molecular and Cellular Biology, 16*, 5604–5615.
12. Culligan, K., & Hays, J. (2000). *Arabidopsis* MutS homologs-AtMSH2, AtMSH3, AtMSH6, and a novel AtMSH7-form three distinct protein heterodimers with different specificities for mismatched DNA. *Plant Cell, 12*, 991–1002.
13. Wu, S.-Y., Culligan, K., Lamers, M., & Hays, J. (2003). Dissimilar mispair-recognition spectra of *Arabidopsis* DNA-mismatch-repair proteins MSH2*MSH6 (MutSα) and MSH2*MSH7 (MutSγ). *Nucleic Acids Research, 31*, 6027–6034.
14. Lloyd, A. H., Milligan, A. S., Langridge, P., & Able, J. A. (2007). *TaMSH7*: A cereal mismatch repair gene that affects fertility in transgenic barley (*Hordeum vulgare* L.). *BMC Plant Biology, 7*, 67.
15. Spampinato, C. P., Gomez, R. L., Galles, C., & Lario, L. D. (2009). From bacteria to plants: A compendium of mismatch repair assays. *Mutation Research: Reviews in Mutation Research, 682*, 110–128.
16. Murashige, T., & Skoog, F. (1962). A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiologia Plantarum, 15*, 473–497.
17. Mathur, J., & Koncz, C. (1998). Callus culture and regeneration. In J. Martinez-Zapater & J. Salinas (Eds.), *Arabidopsis protocols: Methods in molecular biology* (Vol. 82, pp. 31–34). Totowa, New Jersey: Humana Press, Inc.
18. Plaxton, W. C. (1989). Molecular and immunological characterization of plastid and cytosolic pyruvate kinase isozymes from castor-oil-plant endosperm and leaf. *European Journal of Biochemistry, 181*, 443–451.
19. Bowler, C., Benvenuto, G., Laflamme, P., Molino, D., Probst, A., Tariq, M., et al. (2004). Chromatin techniques for plant cells. *The Plant Journal, 39*, 776–789.
20. He, D., Nickerson, J. A., & Penman, S. (1990). Core filaments of the nuclear matrix. *Journal of Cell Biology, 110*, 569–580.
21. Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature, 227*, 680–685.
22. Burnette, W. N. (1981). "Western Blotting": Electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Analytical Biochemistry, 112*, 195–203.
23. Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L., & Gruissem, W. (2004). GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. *Plant Physiology, 136*, 2621–2632.
24. Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding. *Analytical Biochemistry, 72*, 248–254.
25. Marsischky, G., Filosi, N., Kane, M., & Kolodner, R. (1996). Redundancy of *Saccharomyces cerevisiae* MSH3 and MSH6 in MSH2-dependent mismatch repair. *Genes and Development, 10*, 407–410.
26. Johnson, R., Kovvali, G., Prakash, L., & Prakash, S. (1996). Requirement of the yeast MSH3 and MSH6 genes for MSH2-dependent genomic stability. *Journal of Biological Chemistry, 271*, 7285–7288.
27. Palombo, F., Gallinari, P., Iaccarino, I., Lettieri, T., Hughes, M., D'Arrigo, A., et al. (1995). GTBP a 160-kilodalton protein essential for mismatch-binding activity in human cells. *Science, 268*, 1912–1914.
28. Palombo, F., Iaccarino, I., Nakajima, E., Ikejima, M., Shimada, T., & Jiricny, J. (1996). hMutSß, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion loops in DNA. *Current Biology, 6*, 1181–1184.
29. Alou, A., Azaiez, A., Jean, M., & Belzile, F. (2004). Involvement of the *Arabidopsis thaliana* AtPMS1 gene in somatic repeat instability. *Plant Molecular Biology, 56*, 339–349.
30. Depeiges, A., Farget, S., Degroote, F., & Picard, G. (2005). A new transgene assay to study microsatellite instability in wild-type and mismatch-repair defective plant progenies. *Plant Science, 168*, 939–947.
31. Leonard, J., Bollmann, S., & Hays, J. (2003). Reduction of stability of *Arabidopsis* genomic and transgenic DNA-repeat sequences (microsatellites) by inactivation of AtMSH2 mismatch-repair function. *Plant Physiology, 133*, 328–338.
32. Hoffman, P., Leonard, J., Lindberg, G., Bollmann, S., & Hays, J. (2004). Rapid accumulation of mutations during seed-to-seed propagation of mismatch-repair-defective *Arabidopsis*. *Genes and Development, 18*, 2676–2685.
33. Lafleuriel, J., Degroote, F., Depeiges, A., & Picard, G. (2007). Impact of the loss of AtMSH2 on double-strand break-induced recombination between highly diverged homeologous sequences in *Arabidopsis thaliana* germinal tissues. *Plant Molecular Biology, 63*, 833–846.
34. Li, L., Jean, M., & Belzile, F. (2006). The impact of sequence divergence and DNA mismatch repair on homeologous recombination in Arabidopsis. *The Plant Journal, 45*, 908–916.
35. Emmanuel, E., Yehuda, E., Melamed-Bessudo, C., Avivi-Ragolsky, N., & Levy, A. A. (2006). The role of *AtMSH2* in homologous recombination in *Arabidopsis thaliana*. *EMBO Report, 7*, 100–105.
36. Dion, E., Li, L., Jean, M., & Belzile, F. (2007). An Arabidopsis MLH1 mutant exhibits reproductive defects and reveals a dual role for this gene in mitotic recombination. *The Plant Journal, 51*, 431–440.
37. Li, L., Dion, E., Richard, G., Domingue, O., Jean, M., & Belzile, F. J. (2008). The *Arabidopsis* DNA mismatch repair gene PMS1 restricts somatic recombination between homeologous sequences. *Plant Molecular Biology, 69*, 675–684.
38. Ade, J., Haffani, Y., & Belzile, F. (2001). Functional analysis of the *Arabidopsis thaliana* mismatch repair gene MSH2. *Genome, 44*, 651–657.
39. Ade, J., Belzile, F., Philippe, H., & Doutriaux, M. (1999). Four mismatch repair paralogues coexist in Arabidopsis thaliana: AtMSH2, AtMSH3, AtMSH6–1 and AtMSH6–2. *Molecular and General Genetics, 262*, 239–249.

40. Horwath, M., Kramer, W., & Kunze, R. (2002). Structure and expression of the Zea mays mutS-homologs Mus1 and Mus2. *Theoretical and Applied Genetics, 105*, 423–430.
41. Larkins, B., Dilkes, B., Dante, R., Coehlo, C., Woo, Y.-M., & Liu, Y. (2001). Investigating the hows and whys of DNA endoreduplication. *Journal of Experimental Botany, 52*, 183–192.
42. Fras, A., Juchimiuk, J., Siwinska, D., & Maluszynska, J. (2007). Cytological events in explants of *Arabidopsis thaliana* during early callogenesis. *Plant Cell Reports, 26*, 1933–1939.
43. Polanco, C., & Ruiz, M. L. (2002). AFLP analysis of somaclonal variation in *Arabidopsis thaliana* regenerated plants. *Plant Science, 162*, 817–824.
44. Fras, A., & Maluszynska, J. (2005). The correlation between the chromosome variation in callus and genotype of explants of *Arabidopsis thaliana*. *Genetica, 121*, 145–154.
45. Tam, S. M., Samipak, S., Britt, A., & Chetelat, R. T. (2009). Characterization and comparative sequence analysis of the DNA mismatch repair MSH2 and MSH7 genes from tomato. *Genetica, 137*, 1–14.
46. Gustafson, C., Govindarajan, S., & Minshull, J. (2004). Codon bias and heterologous protein expression. *Trends in Biotechnology, 22*, 346–353.
47. Chang, D. K., Ricciardiello, L., Goel, A., Chang, C. L., & Boland, C. R. (2000). Steady-state regulation of the human DNA mismatch repair system. *Journal of Biological Chemistry, 275*, 18424–18431.