American Genetic Association

*Journal of Heredity*, 2016, 266–273
doi:10.1093/jhered/esw003
Original Article
Advance Access publication January 16, 2016

OXFORD

Original Article

# Plastome Mutations and Recombination Events in Barley Chloroplast Mutator Seedlings

**Alejandra Landau, Franco Lencina, María G. Pacheco, and Alberto R. Prina**

From the Instituto de Genética "Ewald A. Favret", CICVyA (Centro de Investigación en Ciencias Veterinarias y Agronómicas), Instituto Nacional de Tecnología Agropecuaria, Nicolás Repetto y de los Reseros s/n (1686) Hurlingham, Buenos Aires, Argentina (Landau, Lencina, Pacheco, and Prina).

Address correspondence to Alberto R. Prina at the address above, or e-mail: prina.albertoraul@inta.gob.ar.

Received September 9, 2015; First decision October 23, 2015; Accepted January 4, 2016.

Corresponding editor: Susan Gabay-Laughnan

## Abstract

The barley chloroplast mutator (*cpm*) is an allele of a nuclear gene that when homozygous induces several types of cytoplasmically inherited chlorophyll deficiencies. In this work, a plastome Targeting Induced Local Lesions in Genomes (TILLING) strategy based on mismatch digestion was used on families that carried the *cpm* genotype through many generations. Extensive scanning of 33 plastome genes and a few intergenic regions was conducted. Numerous polymorphisms were detected on both genic and intergenic regions. The detected polymorphisms can be accounted for by at least 61 independent mutational events. The vast majority of the polymorphisms originated in substitutions and small indels (insertions/deletions) in microsatellites. The *rpl23* and the *rps16* genes were the most polymorphic. Interestingly, the variation observed in the *rpl23* gene consisted of several combinations of 5 different one nucleotide polymorphisms. Besides, 4 large indels that have direct repeats at both ends were also observed, which appear to be originated from recombinational events. The *cpm* mutation spectrum suggests that the CPM gene product is probably involved in plastome mismatch repair. The numerous subtle molecular changes that were localized in a wide range of plastome sites show the *cpm* as a valuable source of plastome variability for plant research and/or plant breeding. Moreover, the *cpm* mutant appears to be an interesting experimental material for investigating the mechanisms responsible for maintaining the stability of plant organelle DNA.

**Subject areas:** Gene action, regulation and transmission, Conservation genetics and biodiversity
**Keywords:** plastome polymorphisms, plastid DNA instability, plastid gene TILLING

Spontaneous genetic variability in the plastid genome or plastome is scarce and the plastome is usually considered as a highly conserved genome (Palmer 1985, 1990; Wolfe et al. 1987; Clegg 1993; Wicke et al. 2011). Results from the artificial induction of plastome mutations are so far very limited, contrasting with the amount of information coming from experiments on nuclear gene-induced mutations (van Harten 1998; Prina et al. 2012a). The use of genetically unstable genotypes as a tool for expanding plastome variability has been pointed out as a promising alternative to artificial mutagenesis, however it has received little attention (Kirk and Tilney-Bassett 1978; Börner and Sears 1986; Gressel and Levy 2010; Prina et al. 2012a). Furthermore, in comparison with our knowledge about DNA repair mechanisms in bacteria and eukaryotic nuclei, little is known about these mechanisms in plant organelles, especially chloroplasts (Rowan et al. 2010; Ruhlman and Jansen 2014). In those regards, the barley chloroplast mutator (*cpm*), previously described as inducing several types of maternally inherited chlorophyll deficiencies (Prina 1992; Prina et al. 2009), appears to be an interesting

experimental material that deserves further investigations. Only a few of the mutants that originated in *cpm* plants, which were previously isolated by their phenotypes, have been experimentally identified as plastome mutants (Rios et al. 2003; Landau et al. 2007, 2009, 2011). To better understand the mechanisms of the *cpm* and to evaluate it as a source of plastome variability, in this study we aimed to obtain more molecular information on the types of induced mutations and their distribution throughout the plastome. Nowadays, the successful use of reverse genetic strategies based on mismatch digestion called Targeting Induced Local Lesions in Genomes (TILLING), (McCallum et al. 2000) for high-throughput searching for induced nuclear DNA polymorphisms, led us to pursue a similar strategy for the screening of polymorphisms in the chloroplast genome of *cpm* plants (cpTILLING).

To fulfill the objectives mentioned above, we took advantage of experimental material previously available that was maintained by natural self-pollination of plants carrying the *cpm* genotype through numerous generations. One group of plants consisted of 2 families that have kept the mutator genotype from 12 to 17 generations, while the other one consisted of 4 families that carried the mutator genotype for 5 generations. We conducted an extensive sequence scan of 33 chloroplast genes and a few intergenic regions. In this way, we detected a substantial number of polymorphisms that originated from at least 61 independent mutational events and were widely distributed in different segments of the plastome and affecting sequences involved in a variety of functions. All results show the *cpm* mutant as a valuable source of plastome variability for plant research and/or plant breeding and suggest that the nuclear gene responsible for the *cpm* syndrome encodes a protein involved in the mismatch repair system of plastid DNA.

## Materials and Methods

### Plant Material

The plant material consisted of chloroplast mutator (*cpm*) (Prina 1992, 1996) and wild-type (WT) barley seedlings. The latter is the parental genotype, which was mutagenized and later used to isolate the *cpm* mutant (Prina 1992). The *cpm* seedlings analyzed in this work belong to 2 groups of progenies. Group A consisted of 2 families that carried the homozygous mutator genotype (*cpm/cpm*) for many generations (from 12 to 17). One of these families (A-$M_2$) originated from the very same *cpm/cpm* $M_2$ plant, which through self-pollination, produced the progenies used for isolating the *cpm* mutant at the $M_4$–$M_5$ generations (Prina 1992). The other one (A-$F_2$) was genotyped as a *cpm/cpm* $F_2$ plant, which came from self-pollination of an $F_1$ derived from a cross of a WT plant as female with a *cpm/cpm* plant as the male parent (see Prina 1992, 1996; Landau et al. 2011). Group B consisted of 4 families that carried the mutator genotype for 5 generations, each one from a different $F_2$ plant (B-$F_2$-1, B-$F_2$-2, B-$F_2$-3, and B-$F_2$-4), obtained as described above for A-$F_2$ progenies (see Supplementary Material 1).

Both groups of families were already available before the start of this work, and we chose this experimental material for the present investigation with the idea that in these families there would be a greater chance to find *cpm*-induced plastome mutants than in families that carried the *cpm* activity during fewer generations.

Regarding the seedlings from which the control DNA was isolated, they belong to families that have gone through many more generations of self-pollination than those that produced the *cpm* seedlings. Indeed, they come from a genotype that had already undergone many generations of natural self-pollination when it was used to mutagenize and to isolate the *cpm* mutant (Prina 1992). Thereafter, this mother genotype was multiplied by natural self-pollination in parallel with the *cpm* analyzed families. It is worth noting that in this material there was no phenotypic evidence (presence of chlorophyll deficiencies and/or conspicuous morphological changes) suggestive of mutational changes. Those analyses were carried out in the glasshouse at second leaf seedling stage and in plants grown at the field nursery until maturity.

### DNA Isolation

Genomic DNA was isolated from 1 or 2 leaves of individual seedlings using the micromethod described in Dellaporta (1994) with modifications. The tissue was ground with Dellaporta isolation buffer in the The Fast Prep®-24 Instrument (MP Biomedicals, USA) and extracted with chloroform before DNA precipitation.

DNA concentrations were measured using a spectrophotometer (Nanodrop, Thermo Scientific, Wilmington, DE, USA) and standardized to a concentration of 80 ng/µL.

### Primer Design

Primers were designed based on the chloroplast genome sequence of barley (*Hordeum vulgare*) [GenBank: NC_008590.1] using Primer3 software (v. 0.4.0, http://frodo.wi.mit.edu/). The primers were designed to have an amplicon size of 1–1.5 kb (see Supplementary Material 2) and checked for single-band amplification on agarose gels before the TILLING screen. We took into account that most of the chloroplast-encoded genes are smaller than the required size for an adequate amplicon for TILLING (1–1.5 kb). For this reason, many amplicons contained 2 or more genes and intergenic regions. On the other hand, due to their larger size, some genes as *16S rRNA*, *23S rRNA*, *ycf3*, *psaA*, and *psaB*, required the design of more than one pair of primers.

The 31 amplicons were distributed throughout different regions of the plastome, i.e. large single copy (LSC), small single copy (SSC), and inverted repeat (IR) regions. They covered 33 genes that according to their function can be grouped as follows: group 1: genes of the photosynthesis apparatus: *psbA*, *psbC*, *psbD* (PSII); *psaA*, *psaB* (PSI); *rbcL* (Rubisco large subunit), and *ycf3* (chaperone involved in PSI assembly); group 2: genes related to the chloroplast translation machinery: *rps2*, *rps3*, *rps4*, *rps7*, *3' rps12*, *rps14*, *rps16*, *rps18*, *rps19*, *rpl2*, *rpl16*, *rpl20*, *rpl22*, *rpl23*, *rpl33* (ribosomal proteins), *infA* (translation initiation factor 1), *16S rRNA*, *23S rRNA* (ribosomal RNAs) and *tRNA His*, *tRNA Arg*, *tRNA Met*, *tRNA Leu* (transfer RNAs), and group 3: genes whose functions are not very well understood and that do not belong clearly to any of these groups: *matK* (maturase K), *ccsA* (cytochrome c biogenesis protein), *clpP* (ATP-dependent Clp protease proteolytic subunit), and *cemA* (envelope membrane protein) (for details see Supplementary Material 2).

### PCR and Screening of Mutants

The DNA samples from individual *cpm* seedlings were mixed with DNA from WT control seedlings in a 1:1 ratio. A PCR reaction was performed in a final volume of 12 µL using 40 ng of genomic DNA, 1× Taq buffer (50 mM KCl, 10 mM Tris–HCl, pH 9.0, 0.1% Triton X-100), 1.25 mM $MgCl_2$, 0.6 µM of each primer, 0.6 mM dNTP mix, and 1.25 units kit T-plus 5 U/µl Taq DNA polymerase (Inbio Highway, Tandil, Argentina). After denaturation at 94 °C for 3 min, the reaction mixtures were heated to 94 °C for 30 s, 60 °C for 1 min, and 72 °C for 2 min, in 35 cycles and 10 min at 72 °C. Denaturation consisted of heating the samples at 99 °C for 10 min;

this was followed by a 70-step re-annealing series starting at 70 °C for 20 s, decreasing 0.3 °C in each step, to favor the formation of a heteroduplex at the end of the PCR reaction. After re-annealing, the whole PCR reaction was digested with celery juice extract (CJE) obtained with the protocol described in Till et al. (2006). Digestions and electrophoresis were performed according to Uauy et al. (2009). Samples were run on 3% polyacrylamide gels (19:1 Acrylamide:bis ratio) in 0.5× TBE running buffer and stained with ethidium bromide after electrophoresis. We used a Protean II xi Cell vertical electrophoresis system (gel size: 20 cm tall by 1.5 mm thick; Bio-Rad Laboratories, CA, USA) and samples were run at 350 V for 90 min. Gel images were acquired with a G:BOX Chemi gel documentation system (Syngene, Cambridge, UK). The putative mutants were identified by the presence of cleaved products whose combined size was similar to that of the original PCR product. These samples were then sequenced together with the same fragment of the WT barley as a PCR product with their specific primers. The sequence alignments were done for the putative mutants, our WT barley and the sequence of reference [GenBank: NC_008590.1]. Vector NTI Software (Life Technologies, USA) was used for contig assembling, alignments, and subsequent sequence analyses. In order to discard errors due to the Taq polymerase and/or sequencing, once a pattern of digestion compatible with a putative polymorphism was found, the PCR and digestion were repeated to confirm the pattern of digestion before resequencing was conducted. Furthermore, the results of sequencing the putative mutants were always compared with both the corresponding sequence found in the WT control and the reference barley sequence.

## Results

The location of the analyzed amplicons within the barley plastome is displayed in Figure 1. Polymorphic amplicons found in the families belonging to group A or B are shown in different colors in Figure 1. The detailed molecular changes composing the polymorphisms and the number of times each polymorphism was observed in independently analyzed seedlings are listed in Tables 1 and 2, for groups A and B, respectively. Due to the peculiar pattern of variation observed in the *rpl23* gene, these data are shown separately in Table 3.

In order to estimate the number of mutational events that would cause the observed polymorphisms, it should be noted that each of the 6 families came from a different $M_2$ or $F_2$ plant (see Material and Methods) ensuring that the polymorphisms observed in each family would come from mutations that occurred independently from those observed in other families. Thus, identical polymorphisms observed in different families were considered to have originated from independent mutational events, as the case of the A387G transition in the *rps18* gene that was observed in both families of group A, while all identical polymorphisms found within a family were considered as having originated from a single mutational event. According to this reasoning, all the polymorphisms detected after the screening of 182 seedlings belonging to group A were interpreted as coming from at least 35 independent mutational events, while the polymorphisms detected through the analysis of 122 seedlings of group B would correspond to at least 26 mutations. Altogether, 61 independently induced mutations were detected after analyzing 304 seedlings. Fourteen different polymorphisms were found in intergenic regions and, without considering the *rpl23* gene, 47 polymorphisms were located in 19 of the 33 analyzed genes (summarized in Table 4). The latter affected protein coding, intron, tRNA, and rRNA sequences (see Tables 1 and 2).

The largest number of different polymorphisms was found in the *rps16* gene, which was polymorphic in 5 of the 6 analyzed families, with 9 different polymorphisms that were all located in the intron. Identical polymorphisms were repeatedly found in different seedlings within families, a fact that was especially observed in families of group A, which kept the mutator genotype (*cpm*/*cpm*) over many

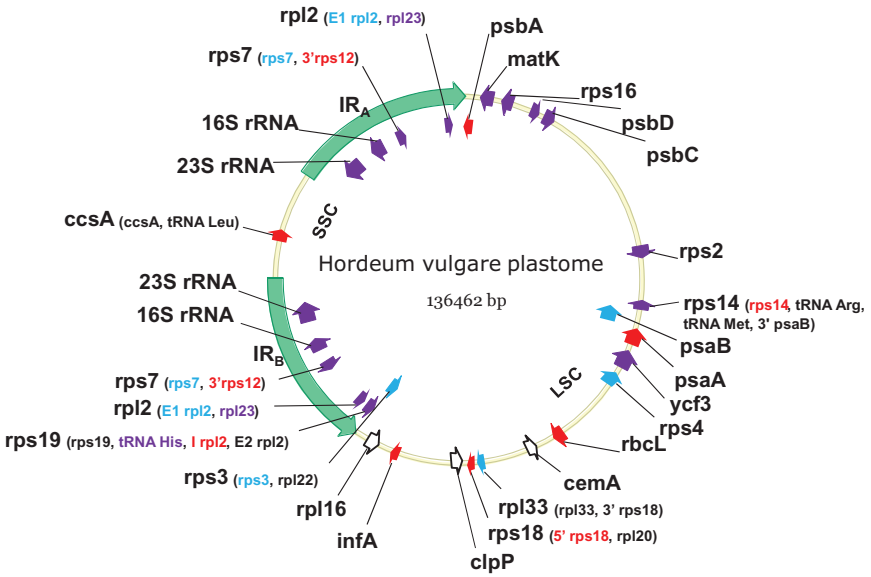


**Figure 1.** The barley plastome. The arrows represent the amplicons assayed in this work. Red arrows indicate mutations found in group A. Blue arrows indicate mutations found in group B. Purple arrows indicate mutations found in both groups. White arrows indicate the amplicons without mutations. Brackets () show the genes included in the amplicons that contain more than one gene. The names of genes or regions of the genes mutated within each amplicon are written in red (mutations in group A), blue (mutations in group B) or purple (mutations in both groups). In the genes containing introns, E indicates exons and I indicates introns. Graphics made with Vector NTI Software (Life Technologies, USA) using the chloroplast genome sequence of barley (*Hordeum vulgare*) [GenBank: NC_008590.1].

**Table 1.** Summary of polymorphisms found in families of Group A

| Gene/amplicon | Nucleotide change | Location |
|---|---|---|
| $F_2$ family | | |
| *rps2* | Ins 1 A 15bp upstream the ATG (×2)[a] | Intergenic |
| | Ins 2 A 15bp upstream the ATG (×3) | Intergenic |
| | G 591 A (×1) | Exon (silent) |
| *3′-rps12* | A 141 G (×1) | Exon (silent) |
| *rps14* | A 35 G (×1) | Exon (missense) |
| | C 703 G of the amplicon (×1) | Intergenic |
| | A 156 G of the amplicon (×1) | Intergenic |
| | T 795 C from primer rps14 R (×4) | Intergenic |
| *rps16* | Ins 1 G 805 (×5) | Intron |
| | G 180 A (×1) | Intron |
| | G 217 A(×1) | Intron |
| *rps18* | C 151 T (×1) | Exon (missense) |
| | A 387 G (×1) | Exon (silent) |
| *rps19* | Ins of 15 n (846 bp of the amplicon) (×3) | Intergenic |
| *rpl2* | A 1019 G (×2) | Intron |
| *psaA* | A 1375 G(×1) | Exon (missense) |
| *psbA* | Del of 620 n (217bp) (×1) | Exon (frameshift) |
| *psbC* | T 573 C (×1) | Exon (silent) |
| | Ins 1 G 311 (×1) | Exon (frameshift) |
| *psbD* | A 644 G (×1) | Exon (missense) |
| *matK* | Ins 2 T (891 bp) (×1) | Exon (frameshift) |
| | Ins 1 T (613 bp) (×3) | Exon (frameshift) |
| *rbcL* | T 456 C (×4) | Exon (silent) |
| *ycf3* | A 1183 G (×1) | Intron 2 |
| *ccsA* | Ins 2 A (1203bp of the amplicon) (×3) | Intergenic |
| *infA* | Del 1 T (418bp of the amplicon) (×1) | Intergenic |
| | A 724 G of the amplicon (×1) | Intergenic |
| *16S rRNA* | A 1136 G (×1) | |
| *23S rRNA* | C 2517 T (×3) | |
| | A 2674 G (×1) | |
| *tRNAHis* | A 47 G (×1) | |
| $M_2$ family | | |
| *3′-rps12* | G 706 A (×2) | Intron 2 |
| *ycf3* | Ins 1 T 151 (×2) | Intron 1 |
| | C 840 T (×2) | Intron 1 |
| *rps18* | A 387 G (×1) | Exon (silent) |

[a]Brackets () show the number of times that the same polymorphism was independently detected. Ins: insertion. Del: deletion.

**Table 2.** Summary of polymorphisms found in families of Group B

| Gene/amplicon | Nucleotide change | Location |
|---|---|---|
| $F_2$-1 family | | |
| *rps2* | T × C 264 bp of the amplicon (×1)[a] | Intergenic |
| *rps16* | A 517 G (×1) | Intron |
| *psbC* | A 1043 G (×1) | Exon (missense) |
| *psbD* | A 296 G (×1) | Exon (missense) |
| *matK* | Del 1 T (271 bp) (×1) | Exon (frameshift) |
| *23S rRNA* | G 2502 A (×1) | |
| $F_2$-2 family | | |
| *rps16* | A 185 G (×1) | Intron |
| *rpl2* | C 2 T (×1) | Exon (missense) |
| *psbD* | C 245 T (×1) | Exon (missense) |
| *ycf3* | A 886 T (×1) | Intron 1 |
| $F_2$-3 family | | |
| *rps4* | A × C 333 bp of the amplicon (×1) | Intergenic |
| *rps16* | T 348 C (×1) | Intron |
| | A 642 G (×1) | Intron |
| *rpl33* | Del of 79 nt at 505 bp of the amplicon (×1) | Intergenic |
| *tRNA His* | A 14 G (×1) | |
| $F_2$-4 family | | |
| *rps2* | Ins 1 A 407 bp of the amplicon (×1) | Intergenic |
| *rps3* | Del of 45 nt at 211bp (×1) | Exon (loss of 15 aa) |
| | A 105 G (×2) | Exon (silent) |
| *rps14* | A × G at 666bp of the amplicon (×1) | Intergenic |
| *rps7* | G 86 A (×1) | Exon (missense) |
| *rps16* | C 795 T (×2) | Intron |
| | T 84 C (×2) | Intron |
| *16S rRNA* | A 1810 G (×2) | |
| *matK* | T 669 C (×1) | Exon (silent) |
| *psaB* | T 1796 C (×1) | Exon (missense) |
| *psbC* | T 558 C (×1) | Exon (silent) |

[a]Brackets () show the number of times that the same polymorphism was independently detected. Ins: insertion. Del: deletion

generations (see Table 1). Those observed in a larger number of seedlings were found in family A-$F_2$: the insertion in the *rps16* gene of 1 G at position 805 bp, which was independently observed in 5 seedlings; the *rbcL* gene polymorphisms T 456 C; and the polymorphism in an intergenic region near the *rps14* gene T 795 C observed in 4 seedlings each. In group A families, there were also 10 cases of polymorphisms that were independently detected 2 or 3 times in different seedlings of the same family, but only 4 cases of identical polymorphisms were found twice in group B families (see Table 1 and 2).

As mentioned above, the *rpl23* gene showed a peculiar pattern of variation (Table 3). It consisted of different combinations of the very same 5 molecular changes: 2 transitions (G118A: missense; A203G: missense), 2 transversions (G115T: missense; T132A: silent), and 1 nucleotide deletion (G at position 133: frameshift). In some seedlings the 5 polymorphisms were observed together, whereas other different combinations of 4, 3, 2 or only 1 polymorphism were also observed.

In Table 5, the mutations detailed in Tables 1 and 2 are grouped according to the type of molecular changes. Fifty-seven of the 61 polymorphisms detected correspond to point mutations. They consisted of 43 transitions (31 were A/T to G/C vs. 12 that were G/C to A/T), 3 transversions and 11 small insertions/deletions (indels). The small indels consisted of 1 or 2 nucleotides localized in microsatellites (mononucleotide repeats of 9–10 bases).

Besides these, 4 large indels (1 insertion of 15 bp and 3 deletions of 45, 79 and 620 bp) were detected, all of them having direct repeat sequences flanking the inserted/deleted fragments ranging from 7 to 25 bp. The largest deletion of 620 nucleotides was found in the coding sequence of the *psbA* gene. This deletion was detected in heteroplastidic condition and was observed in the gel as 2 bands of PCR products (Figure 2). The higher molecular weight band (1390 bp) corresponded to the entire amplicon and the lower one (770 bp) to the fragment without the deleted 620 bp. The absence of this big

**Table 3.** Combination of polymorphisms found in the *rpl23* gene in families of groups A and B. Presence of a polymorphism is indicated with an X

| Group | Family | G 115 T | G 118 A | T 132 A | Del 1G (133) | A 203 G |
|---|---|---|---|---|---|---|
| A | $M_2$ | — | — | — | — | X |
| | | X | X | X | X | X |
| | $F_2$ | — | — | — | — | X |
| | | — | — | X | X | — |
| | | — | — | X | X | X |
| | | X | X | X | X | — |
| | | X | X | X | X | X |
| B | $F_2$-1 | — | — | — | — | X |
| | | — | — | X | X | — |
| | | X | X | X | X | X |
| | $F_2$-2 | — | — | — | — | X |
| | | — | — | X | X | X |
| | $F_2$-3 | — | — | — | — | X |
| | | — | — | X | X | — |
| | | X | X | X | X | — |
| | | X | X | X | X | X |
| | $F_2$-4 | — | — | — | — | X |
| | | — | — | X | X | — |
| | | X | X | — | — | — |
| | | — | — | X | X | X |
| | | X | X | X | X | — |
| | | X | X | X | X | X |

**Table 4.** Distribution of polymorphisms among the 19 genes affected, with the exception of the *rpl23* gene data that are presented in Table 3

| Genes | Group A | | Group B | | | | Total |
|---|---|---|---|---|---|---|---|
| | $M_2$ | $F_2$ | $F_2$-1 | $F_2$-2 | $F_2$-3 | $F_2$-4 | |
| *rps2* | — | 1 | — | — | — | — | 1 |
| *rps3* | — | — | — | — | — | 2 | 2 |
| *rps7* | — | — | — | — | — | 1 | 1 |
| *3'rps12* | 1 | 1 | — | — | — | — | 2 |
| *rps14* | — | 1 | — | — | — | — | 1 |
| *rps16* | — | 3 | 1 | 1 | 2 | 2 | 9 |
| *rps18* | 1 | 2 | — | — | — | — | 3 |
| *rpl2* | — | 1 | — | 1 | — | — | 2 |
| *psaA* | — | 1 | — | — | — | — | 1 |
| *psaB* | — | — | — | — | — | 1 | 1 |
| *psbA* | — | 1 | — | — | — | — | 1 |
| *psbC* | — | 2 | 1 | — | — | 1 | 4 |
| *psbD* | — | 1 | 1 | 1 | — | — | 3 |
| *ycf3* | 2 | 1 | — | 1 | — | — | 4 |
| *rbcL* | — | 1 | — | — | — | — | 1 |
| *matK* | — | 2 | 1 | — | — | 1 | 4 |
| *16S rRNA* | — | 1 | — | — | — | 1 | 2 |
| *23S rRNA* | — | 2 | 1 | — | — | — | 3 |
| *tRNAHis* | — | 1 | — | — | 1 | — | 2 |
| Total | 4 | 22 | 5 | 4 | 3 | 9 | 47 |

portion of the sequence was confirmed by isolating and sequencing the 770 bp band from the gel. Two additional smaller bands of approximately 430 bp and 340 bp could be observed in the gel corresponding to the digestion of the 770-bp fragment (Figure 2). These fragments probably originated from the digestion of a loop generated during the heteroduplex (770 bp: 1390 bp) formation at the step of re-annealing, which was recognized and cleaved by the CJE.

The data presented in Tables 1 and 2 result from the analysis of 42 466 bp per seedling. As mentioned above, by analyzing 182 seedlings corresponding to group A, 35 mutations were detected that accumulated after an average of roughly 14.5 generations carrying the *cpm* syndrome, while in group B, after 5 generations, 26 mutations were detected by analyzing 122 seedlings. Considering only the substitutions, which were 24 for group A or 22 for group B, the substitution rate per bp and generation were $2.1 \times 10^{-7}$ and $8.5 \times 10^{-7}$, respectively.

## Discussion

The barley chloroplast mutator (*cpm*) is to our knowledge the only example in monocots in which a wide spectrum of cytoplasmically inherited mutants is induced (Prina 1992, 1996; Greiner 2012; Prina et al. 2012a). Previous investigations carried out on mutants that were isolated by direct genetics, allowed us to discover only a few plastome mutations caused by the *cpm* (Rios et al. 2003; Landau et al. 2007, 2009, 2011). The main goal of the present work was to detect a more significant number of polymorphisms in *cpm* plants in order to gain knowledge about the molecular types of mutations induced by the *cpm* and its target sites. We applied a TILLING (McCallum et al. 2000) approach based on enzymatic mismatch cleavage for the detection of plastome polymorphisms. This approach, which can be carried out on small samples of tissue, is appropriate for analyzing the experimental material used, that consist of seedlings supposed to be in a wide variety of heteroplastidic conditions. A TILLING strategy was previously used to detect plastome polymorphisms in natural populations (ORG-EcoTILLING, Zeng et al. 2012), but we found no reports of its use for searching "newly induced" mutants in the plastome. The success of the chosen strategy was *a priori* not obvious, because after a plastome mutation occurs, the growth of the mutant sector to an appropriate size to be detected is slower and less predictable than that of a nuclear one (Kirk and Tilney-Bassett 1978; Birky 2001; Prina et al. 2012a, 2012b).

As was pointed out in Materials and Methods, plastome polymorphisms were detected by mixing the DNA sample isolated from each *cpm* seedling with control DNA. In order to discard errors due to the Taq polymerase and/or to sequencing, once a pattern of digestion compatible with a putative polymorphism was detected, the PCR and digestion were repeated to confirm the pattern once again before sequencing. Furthermore, the results of sequencing the putative mutants were always compared with both the corresponding sequence found in the WT control and the reference barley sequence. Interestingly, all the polymorphisms were found in *cpm* DNA samples, but not in those of the WT control. Moreover, the fact that the designed amplicons in many cases comprised more than one gene and also some intergenic regions suggests that it is highly unlikely that we are amplifying nuclear or mitochondrial inserts of plastid DNA (nupts, mipts) with exactly the same size of our amplicons. Besides, we aligned all the amplicons against the *Hordeum vulgare sub sp. vulgare* Ensembl Database and although, we identified part of the amplicons sequences in the nuclear genome, the sequences of both primers of each amplicon never matched together in that small sequences (not more than 400 pb in the majority of the cases).

From the results in Figure 1 and Tables 1, 2, and 4, it can be concluded that the polymorphisms were widely distributed along the plastomes of *cpm* plants and were located in both genic and intergenic regions. Without considering the peculiar situation observed for the *rpl23* gene, 19 genes were found to be polymorphic, comprising sequences involved in a wide variety of functions, i.e. protein

**Table 5.** Summary of the number and type of mutations in seedlings belonging to families of group A (carrying *cpm* during 12–17 generations) and families of group B (carrying *cpm* during 5 generations), excluding the *rpl23* gene data

| | Number of seedlings | Number of mutational events | Indels | | Transitions | Transversions |
|---|---|---|---|---|---|---|
| | | | Small | Large | | |
| Group A | | | | | | |
| $F_2$ | 159 | 31 | 8 | 2 | 20 | 1 |
| $M_2$ | 23 | 4 | 1 | — | 3 | — |
| Subtotal | 182 | 35 | 9 | 2 | 23 | 1 |
| Group B | | | | | | |
| $F_2$-1 | 26 | 6 | 1 | — | 5 | — |
| $F_2$-2 | 16 | 4 | — | — | 3 | 1 |
| $F_2$-3 | 39 | 5 | — | 1 | 3 | 1 |
| $F_2$-4 | 41 | 11 | 1 | 1 | 9 | — |
| Subtotal | 122 | 26 | 2 | 2 | 20 | 2 |
| Total | 304 | 61 | 11 | 4 | 43 | 3 |

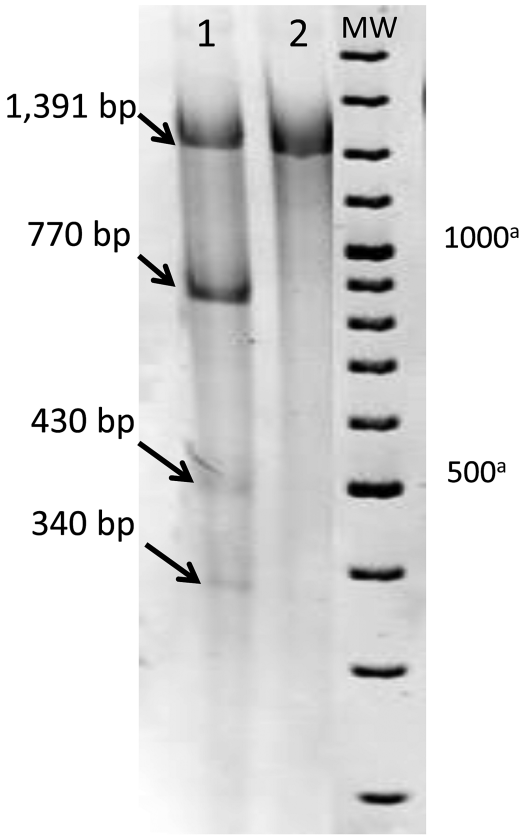


**Figure 2.** CJE digestion of the *psbA* amplicon on a non-denaturing 3% polyacrylamide gel. Lane 1 shows a fragment of 770 bp in a heteroplastidic sample carrying a 620 bp deletion and 2 digestion fragments of 340 and 430 bp. Lane 2 corresponds to a sample without the deletion. [a]The samples run slower than the molecular weight marker (MW) probably due to the components of the CJE digestion mixture.

coding sequences, introns, tRNAs and rRNAs (see Tables 1 and 2). Table 4 summarizes the number of mutations detected in each gene. The *rps16* gene was the most polymorphic, showing 9 different polymorphisms all of them localized in its intron, a result that is consistent with the use of this intron in phylogenetic studies given its potential for variation (Downie and Katz-Downie 1999; Popp and Oxelman 2004).

The situation observed in the *rpl23* gene was very peculiar and therefore these results were not included in Tables 1 and 2. In the *rpl23* gene, 5 different polymorphisms were observed in a variety of combinations (Table 3) that can be hardly explained by independently induced mutations, but rather is more likely they originated in recombination events between the *rpl23* gene and the *rpl23* pseudogene (NCBI Reference Sequence: NC_008590.1), which already in nature contains those 5 polymorphisms. Interestingly, studies of the grass family have shown that this pseudogene is being maintained by gene conversion with the functional gene (Morton and Clegg 1993).

Regarding the types of DNA molecular changes observed in *cpm* seedlings, the vast majority were substitutions, mostly transitions and small indels (Table 5), suggesting that the *CPM* gene is highly specialized in correcting these kind of mutations.

All the 11 small indels detected consisted of 1 or 2 nucleotides located in single sequence repeats (mononucleotide repeats of 9–10 bases), which are the predominant type of microsatellites in chloroplast genomes (Rajendrakumar et al. 2002). Besides the above-mentioned point mutations 4 big indels were detected, 3 of them exceeding the size of the largest indel previously detected by nuclear TILLING (Comai et al. 2004). Interestingly, direct repeats of variable length were detected at both ends of the 4 big indels, suggesting that recombination events were the cause of these large indels. In this regard, intramolecular recombination mediated by short direct repeats has been reported as a mechanism for producing plastome genetic diversity in wheat species (Ogihara et al. 1988, 1991).

Regarding the mechanism responsible for the *cpm* syndrome, the present results support the early idea about the malfunction of a nuclear encoded enzyme responsible for plastome DNA integrity (Prina 1992), which in accordance with Hasting et al. (1976), when recessive indicates that it corresponds to a gene affecting the DNA repair capabilities. Besides, the mutational spectrum of molecular changes described above strongly suggests that the *CPM* gene is involved in the DNA mismatch repair (MMR) system, which plays a fundamental role in mismatch repair during replication and has also been reported as having anti-recombination activity (Harfe and Jinks-Robertson 2000; Jun et al. 2006; Lin et al. 2007). In this sense, it is widely accepted that an increase in tandem repeat indels is a typical indicator of a failure in the DNA-MMR system (Thibodeau et al. 1993; Vaish and Mittal 2002). It is interesting to note that an increase in short indels associated with microsatellite DNA was observed by Stoike and Sears (1998) in *Oenothera* plants carrying a plastome mutator, which is one of the few mutator genotypes

previously reported as inducing a wide spectrum of cytoplasmically inherited mutants in dicots (Epp 1973).

Another mutator genotype inducing a wide spectrum of mutants was originally reported in *Arabidopsis* as a chloroplast mutator (*chm*) mutant (Redei 1973) and later noted to induce rearrangements in the mitochondrial genome (Martínez-Zapater et al. 1992). This mutant was proven to be homolog of the *MutS* gene, which in *E.coli* is involved in MMR and recombination, and it was renamed as *AtMSH1* (Abdelnoor et al. 2003). Its failure causes mitochondrial genomic shifting involving rapid and dramatic changes in the relative copy number of portions of the mitochondrial genome instead of fixing mismatches (Abdelnoor et al. 2003; Maréchal and Brisson 2010). Despite extensive research, no modifications were detected within the plastome of *chm* plants in the early investigations (Martínez-Zapater et al. 1992; Mourad and White 1992), while more recently a low frequency of DNA rearrangements mediated by recombination were found in the plastome of MSH1 disrupted plants (Xu et al. 2011). Other previous observations taken on MSH1 *Arabidopsis* mutants that differ from our observations in barley *cpm* plants are the faster sorting out of the mutant clones that was inferred based on leaf variegation and the usual presence of plants carrying distorted leaves and high levels of sterility (see Redei 1973; Sakamoto et al. 1996).

Considering the results summarized above, the MSH1 defective plants seem to differ from our observations of *cpm* plants, suggesting different roles of *AtMSH1* and *CPM*. Our present results strongly suggest that the defective protein encoded by the *CPM* nuclear gene plays a role in the chloroplast MMR system. However, an extensive study of the influence of the *CPM* on the mitochondrion is still lacking in order to determine if the CPM protein has dual targeting to plastids and mitochondria.

In conclusion, through a TILLING strategy we demonstrated that *cpm* seedlings carried numerous plastome polymorphisms, mostly point mutations that were widely distributed over the plastome. The slight molecular impact of the majority of the *cpm* induced mutations could be helpful if we aim to obtain allelic series that could include vital alleles that retain functionality and not only gene knock-outs. In this regard, advances in plastome gene functionality based on plastome mutations are so far scarce (Greiner 2012), whereas most of the current knowledge about plastome gene functions came from reverse genetics studies based on gene knock-outs (reviewed by Rochaix 1997; Rochaix 2003; Leister 2003; Day 2012). In this sense, it is remarkable that from *cpm* seedlings it was possible to isolate mutants in essential genes like in *matK* or in ribosomal subunits genes, for which no gene knock-outs so far exist (Scharff and Bock 2014). Genetically unstable genotypes like the *cpm* can be a useful alternative to artificial mutagenesis for expanding plastome variability, as it has been previously pointed out (Kirk and Tilney-Basset 1978; Börner and Sears 1986; Gressel and Levy 2010; Greiner 2012; Prina et al. 2012a). It appears as a very valuable source of plastome variability for research, classical plant breeding and/or biotechnology, either for use in direct selection experiments (see Rios et al. 2003) or by applying an adequate reverse genetic strategy by TILLING or by massive sequencing. Further research on *cpm* can greatly contribute to improve the limited knowledge about DNA repair mechanisms that maintain the integrity of plant organellar DNA (Rowan et al. 2010; Ruhlman and Jansen 2014). Finally, it is worth mentioning that organelle mutators like *cpm* could play an important role in plant evolution, as it was pointed out by Gressel and Levy (2010) especially with regard to the strategies of plant adaptation to environmental stresses.

## Supplementary Material

Supplementary material can be found at http://www.jhered.oxfordjournals.org/.

## Funding

International Atomic Energy Agency Research Contract No 15671: Isolation and Characterization of Genes Involved in Chloroplast Genes Mutagenesis. Agencia Nacional de Promoción Científica y Tecnológica, Fondo para la Investigación Científica y Tecnológica 2007 Nº 620: The barley chloroplast mutator as a tool to originate plastome genetic variability. Instituto Nacional de Tecnología Agropecuaria, Proyecto Específico Area Estratégica Biotecnología 244631: Mutagenesis techniques for diversity generation on characters of agricultural and/or agro-industrial interest.

## Acknowledgments

We are very grateful to Profs Barbara Sears, Barbara Hohn, Anne Britt, and Thomas Börner for the invaluable suggestions on the original manuscript. We also want to thank Mr. Abel Mario Moglie for skillful handling of the plant material for such a long time since the *cpm* mutant was evident for the first time in 1984.

## Data Availability

Data deposited at Dryad: http://dx.doi.org/doi:10.5061/dryad.j0bd4

## References

Abdelnoor RV, Yule R, Elo A, Christensen AC, Meyer-Gauen G, Mackenzie SA. 2003. Substoichiometric shifting in the plant mitochondrial genome is influenced by a gene homologous to MutS. *Proc Natl Acad Sci USA*. 100:5968–5973.

Birky CW Jr. 2001. The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. *Annu Rev Genet*. 35:125–148.

Börner T, Sears B. 1986. Plastome mutants. *Plant Mol Biol Rep*. 4:69–72.

Clegg MT. 1993. Chloroplast gene sequences and the study of plant evolution. *Proc Natl Acad Sci USA*. 90:363–367.

Comai L, Young K, Till BJ, Reynolds SH, Greene EA, Codomo CA, Enns LC, Johnson JE, Burtner C, Odden AR, *et al*. 2004. Efficient discovery of DNA polymorphisms in natural populations by Ecotilling. *Plant J*. 37:778–786.

Day A. 2012. Reverse Genetics in Flowering Plant Plastids. In: Bock R, Knoop V, editors. *Genomics of Chloroplasts and Mitochondria, Advances in Photosynthesis and Respiration*. The Netherlands: Springer, vol 35. p. 415–441.

Dellaporta S. 1994. Plant DNA miniprep and microprep: versions 2.1–2.3. In: Freeling M, Walbot V, editors. *The Maize Handbook*. New York: Springer-Verlag N. Y. Inc. p. 522–525.

Downie SR, Katz-Downie DS. 1999. Phylogenetic analysis of chloroplast rps16 intron sequences reveals relationships within the woody southern African Apiaceae subfamily Apioideae. *Can J Bot*. 77: 1120–1135.

Epp MD. 1973. Nuclear gene-induced plastome mutations in *Oenothera hookeri*. I. genetic analysis. *Genetics*. 75:465–483.

Greiner S. 2012. Plastome mutants of higher plants. In: Bock R, Knoop V, editors. *Genomics of Chloroplasts and Mitochondria, Advances in Photosynthesis and Respiration*. The Netherlands: Springer, vol 35, p. 237–266.

Gressel J, Levy AA. 2010. Stress, mutators, mutations and stress resistance. In: Pareek A *et al*. editors. *Abiotic Stress Adaptation in Plants. Physiological, Molecular, and Genomic Foundation*. Dordrecht: Springer Science. p. 471–483.

Harfe BD, Jinks-Robertson S. 2000. DNA mismatch repair and genetic instability. *Annu Rev Genet*. 34:359–399.

van Harten AM. 1998. *Mutation Breeding: Theory and Practical Applications*., Cambridge (UK): Cambridge University Press.

Hasting PJ, Quah SK, Borstel RC. 1976. Spontaneous mutations by mutagenic repair of spontaneous lesions in DNA. *Nature*. 264:719–722.

Jun SH, Kim TG, Ban C. 2006. DNA mismatch repair system. Classical and fresh roles. *FEBS J*. 273:1609–1619.

Kirk JTO, Tilney-Bassett RAE. 1978. *The Plastids*. Amsterdam: Elsevier.

Landau A, Díaz Paleo A, Civitillo R, Jaureguialzo M, Prina AR. 2007. Two infA gene mutations independently originated from a mutator genotype in barley. *J Hered*. 98:272–276.

Landau AM, Lokstein H, Scheller HV, Lainez V, Maldonado S, Prina AR. 2009. A cytoplasmically inherited barley mutant is defective in photosystem I assembly due to a temperature-sensitive defect in ycf3 splicing. *Plant Physiol*. 151:1802–1811.

Landau AM, Pacheco MG, Prina AR. 2011. A second infA plastid gene point mutation shows a compensatory effect on the expression of the cytoplasmic line 2 (CL2) syndrome in barley. *J Hered*. 102:633–639.

Leister D. 2003. Chloroplast research in the genomic age. *Trends Genetics*. 19:47–56.

Lin Z, Nei M, Ma H. 2007. The origins and early evolution of DNA mismatch repair genes–multiple horizontal gene transfers and co-evolution. *Nucleic Acids Res*. 35:7591–7603.

McCallum CM, Comai L, Greene EA, Henikoff S. 2000. Targeting induced local lesions IN genomes (TILLING) for plant functional genomics. *Plant Physiol*. 123:439–442.

Maréchal A, Brisson N. 2010. Recombination and the maintenance of plant organelle genome stability. *New Phytol*. 186:299–317.

Martínez-Zapater JM, Gil P, Capel J, Somerville CR. 1992. Mutations at the Arabidopsis CHM locus promote rearrangements of the mitochondrial genome. *Plant Cell*. 4:889–899.

Morton BR, Clegg MT. 1993. A chloroplast DNA mutational hotspot and gene conversion in a noncoding region near rbcL in the grass family (Poaceae). *Curr Genet*. 24:357–365.

Mourad GS, White JA. 1992. The isolation of apparently homoplastidic mutants induced by a nuclear recessive gene in Arabidopsis thaliana. *Theor Appl Genet*. 84:906–914.

Ogihara Y, Terachi T and Sasakuma T. 1988. Intramolecular recombination of chloroplast genome mediated by short direct-repeat sequences in wheat species. *Proc Natl Acad Sci Genetics*. 85: 8573–8577.

Ogihara Y, Terachi T, Sasakuma T. 1991. Molecular analysis of the hot spot region related to length mutations in wheat chloroplast DNAs. I. Nucleotide divergence of genes and intergenic spacer regions located in the hot spot region. *Genetics*. 129:873–884.

Palmer JD. 1985. Comparative organization of chloroplast genomes. *Annu Rev Genet*. 19:325–354.

Palmer JD. 1990. Contrasting modes and tempos of genome evolution in land plant organelles. *Trends Genet*. 6:115–120.

Popp M, Oxelman B. 2004. Evolution of a RNA polymerase gene family in Silene (Caryophyllaceae)-incomplete concerted evolution and topological congruence among paralogues. *Syst Biol*. 53:914–932.

Prina AR. 1992. A mutator nuclear gene inducing a wide spectrum of cytoplasmically inherited chlorophyll deficiences in barley. *Theor Appl Genet*. 85:245–251.

Prina AR. 1996. Mutator-induced cytoplasmic mutants in barley: genetic evidence of activation of a putative chloroplast transposon. *J Hered*. 87:385–389.

Prina AR, Landau A, Colombo N, Jaureguialzo M, Arias MC, Rios RD, Pacheco MG. 2009. Genetically unstable mutants as novel sources of genetic variability: the chloroplast mutator genotype in barley as a tool for exploring the plastid genome. In: Shu QY, editor. *Induced Plant Mutations in the Genomics Era*. Rome: FAO. p. 255–256.

Prina AR, Pacheco MG, Landau AM. 2012a. Mutation Induction in Cytoplasmic Genomes. In: Shu QS, Foster BP, Nakagawa H, editors. *Plant Mutation Breeding and Biotechnology*. Rome: FAO-IAEA. p. 201–206.

Prina AR, Landau AM, Pacheco MG. 2012b. Chimeras and mutant gene transmission. In: Shu QS, Foster BP, Nakagawa H, editors. *Plant Mutation Breeding and Biotechnology*. Rome: FAO-IAEA. p. 179–187.

Rajendrakumar P, Biswal AK, Balachandran SM, Sundaram RM. 2002. In silico analysis of microsatellites in organellar genomes of major cereals for understanding their phylogenetic relationships. *Mol Biol Evol*. 19:2084–91.

Redei GP. 1973. Extrachromosomal mutability determined by a nuclear gene locus in *Arabidopsis*. *Mutat Res*. 18:149–162.

Rios RD, Saione H, Robredo C, Acevedo A, Colombo N, Prina AR. 2003. Isolation and molecular characterization of atrazine tolerant barley mutants. *Theor Appl Genet*. 106:696–702.

Rochaix JD. 1997. Chloroplast reverse genetics: new insights into the function of plastid genes. *Trends Plant Sci*. 2:419–425.

Rochaix JD. 2003. Functional analysis of plastid genes through chloroplast reverse genetics in *Chlamydomonas*. In: Larkum AW, Douglas SE, Raven JA, editors. *Photosynthesis in Algae*. Dordrecht: Kluwer Academic Publishers. p. 83–94.

Rowan BA, Oldenburg DJ, Bendich AJ. 2010. RecA maintains the integrity of chloroplast DNA molecules in Arabidopsis. *J Exp Bot*. 61:2575–2588.

Ruhlman T, Jansen R. 2014. The plastid genomes of flowering plants. In: Maliga P, editor. *Chloroplast biotechnology: Methods and Protocols, Methods in Molecular Biology*, Vol. 1132. New York: Humana Press. p. 3–38.

Sakamoto W, Kondo H, Murata M, Motoyoshi F. 1996. Altered mitochondrial gene expression in a maternal distorted leaf mutant of Arabidopsis induced by chloroplast mutator. *Plant Cell*. 8:1377–1390.

Scharff LB, Bock R. 2014. Synthetic biology in plastids. *Plant J*. 78:783–798.

Stoike LL, Sears BB. 1998. Plastome mutator-induced alterations arise in Oenothera chloroplast DNA through template slippage. *Genetics*. 149:347–353.

Till BJ, Zerr T, Comai L, Henikoff S. 2006. A protocol for TILLING and Ecotilling in plants and animals. *Nat Protoc*. 1:2465–2477.

Thibodeau SN, Bren G, Schaid D. 1993. Microsatellite instability in cancer of the proximal colon. *Science*. 260:816–819.

Uauy C, Paraiso F, Colasuonno P, Tran RK, Tsai H, Berardi S, Comai L, Dubcovsky J. 2009. A modified TILLING approach to detect induced mutations in tetraploid and hexaploid wheat. *BMC Plant Biol*. 9:115.

Vaish M, Mittal B. 2002. DNA mismatch repair, microsatellite instability and cancer. *Indian J Exp Biol*. 40:989–994.

Wicke S, Schneeweiss GM, dePamphilis CW, Müller KF, Quandt D. 2011. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant Mol Biol*. 76:273–297.

Wolfe KH, Li WH, Sharp PM. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc Natl Acad Sci USA*. 84:9054–9058.

Wu SY, Culligan K, Lamers M, Hays J. 2003. Dissimilar mispair-recognition spectra of Arabidopsis DNA-mismatch-repair proteins MSH2*MSH6 (MutSalpha) and MSH2*MSH7 (MutSgamma). *Nucleic Acids Res*. 31:6027–6034.

Xu YZ, Arrieta-Montiel MP, Virdi KS, de Paula WB, Widhalm JR, Basset GJ, Davila JI, Elthon, TE, Elowsky CG, Sato SJ, *et al*. 2011. MSH1 is a nucleoid protein that alters mitochondrial and plastid properties and plant response to high light. *Plant Cell*. 23:3428–41.

Zeng C-L, Wang G-Y, Wang J-B, Yan G-X, Chen B-Y, Xu K, Li J, Gao G-Z, Wu X-M, Zhao B, *et al*. 2012. High-throughput discovery of chloroplast and mitochondrial DNA polymorphisms in Brassicaceae species by ORG-EcoTILLING.*PLoS One*. 7: e47284.