Substoichiometric shifting in the plant mitochondrial genome is influenced by a gene homologous to MutS

Ricardo V. Abdelnoor  
*University of Nebraska - Lincoln*

Ryan Yule  
*University of Nebraska - Lincoln*

Annakaisa Elo  
*University of Nebraska - Lincoln*

Alan C. Christensen  
*University of Nebraska - Lincoln*, achristensen2@unl.edu

Gilbert Meyer-Gauen  
*University of Nebraska - Lincoln*

*See next page for additional authors*

Follow this and additional works at: [http://digitalcommons.unl.edu/plantscifacpub](http://digitalcommons.unl.edu/plantscifacpub)

Part of the [Plant Sciences Commons](http://digitalcommons.unl.edu/plantscifacpub)

Abdelnoor, Ricardo V.; Yule, Ryan; Elo, Annakaisa; Christensen, Alan C.; Meyer-Gauen, Gilbert; and MacKenzie, Sally A., "Substoichiometric shifting in the plant mitochondrial genome is influenced by a gene homologous to MutS" (2003). *Faculty Publications from the Center for Plant Science Innovation*. Paper 45.  
[http://digitalcommons.unl.edu/plantscifacpub/45](http://digitalcommons.unl.edu/plantscifacpub/45)
Authors
Ricardo V. Abdelnoor, Ryan Yule, Annakaisa Elo, Alan C. Christensen, Gilbert Meyer-Gauen, and Sally A. MacKenzie
Substoichiometric shifting in the plant mitochondrial genome is influenced by a gene homologous to MutS

Ricardo V. Abdelnoor†*, Ryan Yule*, Annakaisa Elo*, Alan C. Christensen‡, Gilbert Meyer-Gaunet*, and Sally A. Mackenzie*§

*Plant Science Initiative, ‡School of Biological Sciences and Beadle Center for Genetics Research, University of Nebraska, Lincoln, NE 68588-0660; and †Embrapa Soja, Caixa Postal 231, CEP 86001-970, Londrina, PR, Brazil

Edited by Richard D. Kolodner, University of California at San Diego, La Jolla, CA, and approved February 20, 2003 (received for review December 16, 2002)

The plant mitochondrial genome is retained in a multipartite structure that arises by a process of repeat-mediated homologous recombination. Low-frequency ectopic recombination also occurs, often producing sequence chimeras, aberrant ORFs, and novel subgenomic DNA molecules. This genomic plasticity may distinguish the plant mitochondrion from mammalian and fungal types. In plants, relative copy number of recombination-derived subgenomic DNA molecules within mitochondria is controlled by nuclear genes, and a genomic shifting process can result in their differential copy number suppression to nearly undetectable levels. We have cloned a nuclear gene that regulates mitochondrial substoichiometric shifting in Arabidopsis. The CHM gene was shown to encode a protein related to the MutS protein of Escherichia coli that is involved in mismatch repair and DNA recombination. We postulate that the process of substoichiometric shifting in plants may be a consequence of ectopic recombination suppression or replication stalling at ectopic recombination sites to effect molecule-specific copy number modulation.

A phenomenon rendering the plant mitochondrial genome unusually variable in structure is termed substoichiometric shifting. First reported in maize (17) as the stable presence of a subgenomic molecule encoding a chimeric sequence (16). Genetic analysis of this phenomenon revealed two cases to be under nuclear gene control, involving the Fr gene in Phaseolus vulgaris (23) and the CHM gene in Arabidopsis (24, 25). Mutation of the nuclear CHM gene results in a green-white leaf variegation that, in subsequent generations, displays maternal inheritance (25). The appearance of the variegation phenotype is accompanied by a specific rearrangement (24) that includes amplification of a mitochondrial DNA molecule encoding a chimeric sequence (16). Genetic analysis suggests that the wild-type form of CHM actively suppresses copy number of the subgenomic molecule carrying the chimeric sequence. Loss of proper function of the CHM gene, characterized by two available ethyl methanesulfonate-derived mutant alleles, chm1-1 and chm1-2 (25), and a tissue culture-derived mutant allele, chm1-3 (24), results in rapid and specific copy number amplification of the subgenomic molecule, producing the consequent leaf variegation. It is not clear whether the copy number amplification or suppression of a single subgenomic molecule occurs by differential replication or a recombination mechanism. To investigate further this unusual plant phenomenon, we have cloned the nuclear CHM gene in Arabidopsis that affects mitochondrial substoichiometric shifting. This gene, located on chromosome III, was shown to encode a protein that is targeted to mitochondria and that has high homology to a MutS protein. The identified locus seems to represent a plant counterpart to the yeast mitochondrial MutS homolog (MSh1) and was designated AtMSh1.

Materials and Methods

Gene Mapping, Cloning, and Sequence Analysis. A map-based cloning strategy for the isolation of the CHM locus involved the design of

This paper was submitted directly (Track II) to the PNAS office.

Abbreviation: T-DNA, portion of the Ti (tumor-inducing) plasmid that is transferred to plant cells.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession no. AY191303).

†To whom correspondence should be addressed. E-mail: smackenzie2@unl.edu.
PCR-based codominant markers, using the Cereon Arabidopsis Polymorphism Collection (ref. 26; www.arabidopsis.org/cereon/), to distinguish between the Col-0 and Landsberg erecta ecotypes used in the F2 mapping populations. The markers were designed in a 5-Mb region of chromosome III based on information from the classical mapping experiments of CHM (24, 25). The primer sequences for markers are available on request. The F2 mapping population was derived in our laboratory from a cross between the chm1-1 mutant line and Landsberg erecta ecotype (pollen donor). A segregating subpopulation of 172 variegated plants was analyzed. Genomic DNA purification was conducted according to Li and Chory (27). DNA gel blot analysis was conducted by using the protocol of Sambrook et al. (28).

DNA sequencing of the candidate locus in chm1-1, chm1-2, and chm1-3 mutants (24) was conducted in a Beckman Coulter CEQ2000XL eight-capillary DNA sequencer. Two independent PCR samples for each mutant were sequenced. The 5' RACE analysis was done with the GeneRacer kit (Invitrogen). Mutants chm1-1 and chm1-2 were obtained from the Arabidopsis Biological Resource Center, and mutant chm1-3 was graciously provided to us by C. Somerville (Stanford University, Stanford, CA).

The two sequence-indexed T-DNA [portion of the Ti (tumor-inducing) plasmid that is transferred to plant cells] insertion mutations were identified on The Salk Institute Genomic Analysis Laboratory (SIGnAL) web site at http://signal.salk.edu, and seed for the mutants was obtained from the Arabidopsis Biological Resource Center. The T-DNA insertion positions were confirmed by DNA sequencing of the insertion junctions (data not shown).

Plant Transformation and Biolistic Delivery. The amino acid sequence of AtMSH1 was analyzed with MITOPROT (http://mips.gsf.de/cgi-bin/proj/medgen/mitofilter; ref. 29), and the first 213 nucleotides of the gene were PCR-amplified with the primers MSHtranspFor 5'-GGGCCATGGTGTGAATTGCAT-AGTGTGTGAATTGCAT-3' and MSHtranspRev 5'-GGGCCATGGAAAAGTCGTCG-3'. PCR products were ligated to the pGEM-T Easy Vector System (Promega) and digested with Neol to release the insert. Insert fragments were ligated to the pCAMBIA 1302 vector at the NeoI site that resides at the start of gfp. This vector utilizes the cauliflower mosaic virus 35S promoter. Bombardment experiments used 4-week-old leaves of Arabidopsis (Col-0) with tungsten particles and the Biolistic PDS-1000/He system (Bio-Rad). Particles were bombarded into Arabidopsis leaves by using 900-psi (1 psi = 6.9 kPa) rupture discs under a vacuum of 26 inHg (1 inHg = 3.4 kPa). After the bombardment, Arabidopsis leaves were allowed to recover for 18--22 h on Murashige and Skoog media plates at 22°C in 16 h of daylight. Localization of GFP expression was conducted by confocal laser scanning microscopy with Bio-Rad 1024 MRC-ES by using 488-nm excitation and two-channel measurement of emission, 522 (green/GFP) and 680 (red/chlorophyll) nm.

Results

Positional Cloning of the CHM Locus. High-resolution mapping of the CHM locus on Arabidopsis chromosome III delimited the gene to an 80-kb interval as shown in Fig. 1. Sequence analysis of the interval revealed a gene candidate with similarity in sequence features to the MutS gene of Escherichia coli (Fig. 2). MutS is a component of the E. coli mismatch repair and DNA recombination apparatus (30). The gene, comprising 22 exons,
was predicted to encode a 43-aa mitochondrial targeting presequence with mitochondrial targeting values of 0.916 (MITOPROT), 0.943 (PREDOTAR, www.inra.fr/Internet/Produits/Predotar/), and 0.856 (TARGETP, www.cbs.dtu.dk/services/TARGETP/). RNA gel blots showed that the transcript derived from this gene was 3.5 kb (data not shown), and the encoded protein was 1,118 aa in length, predicting a 124-kDa polypeptide.

 Confirmation of CHM Identity. To test whether the identified locus might be CHM, we obtained two Arabidopsis lines (Col-0 ecotype) containing T-DNA insertion mutations within the candidate gene [GenBank accession nos. BH748178 (SALK041951) and BH908253 (SALK046763)]. The first insertion was located within the fourth exon and the second within the eighth intron. Analysis of the T-DNA mutants (T2 generation) revealed mild green-white leaf variegation, growing more intense in the selfed generation that followed (Fig. 3). Variegation results were obtained for the second T-DNA (SALK041951) mutation as well (data not shown). None of the segregants hemizygous for, or lacking, the candidate locus is predicted to give rise to highly truncated, inactive peptides.

 Sequence analysis of the chm1-3 allele, derived from a tissue culture line by Martinez-Zapater et al. (24), revealed an amino acid (Cys-Tyr) substitution within the intron–exon junction of exon 2 (Fig. 1E). The mutant phenotype in this case may be due to the substitution of a bulkier amino acid within a site essential for protein function.

 The CHM Candidate Has Features of a Mismatch Repair Component. The MutS-homologous gene identified as a candidate for CHM displayed several features characteristic of a mismatch repair component. These features included an ATP-binding domain (amino acids 761–946) comprised of four well conserved motifs designated M1–M4 (ref. 31; Fig. 2B). In addition to ATPase function, this domain seems to be involved in dimerization of the protein (31, 32), although this has not yet been demonstrated for mitochondrial MutS homologs. A DNA binding domain (amino acids 129–206) was also identified (Figs. 1 and 2) to contain the aromatic doublet (FY) motif that is characteristic of this domain in MutS and MutS-like proteins (Fig. 2A). This doublet was shown to be essential for mismatch recognition and specific DNA binding activity (33, 34). We were unable to detect three other conserved domains characteristic of MutS. A connector domain (involved in interdomain interactions), a core domain, and a clamp domain (involved in nonspecific double-strand DNA binding) did not seem to be well conserved.

 The CHM Candidate Protein Likely Localizes to Mitochondria. To confirm that the MutS-like protein localized to the mitochondria,

![Fig. 2. Alignment of ATMSH1 with MutS and MutS homologs.](image-url)
tungsten particle bombardment for transient expression assay. Gene construction was introduced to associated with substoichiometric shifting (24). (1 insertion mutation. Arrowheads indicate DNA band pattern changes asso-

C. Omit digested with BamH I, subjected to gel electrophoresis, blotted, and probed. Lane Wt designates wild-type ecotype Col-0, lane C1 designates mutant chm1-1, and lanes T1 and T2 designate two sister lines containing the T-DNA 1 insertion mutation. A three-primer PCR-based assay to detect substoichiometric shifting (16) was used to assay wild-type Col-0 (WT), mutant chm1-1 (C1), and individual plants segregating for presence of the T-DNA insertion within the candidate CHM locus. All lanes labeled +T represent individual population segregants that are homozygous for the T-DNA insertion, whereas lanes labeled −T include two segregants that are hemizygous for the T-DNA insertion and two segregants that do not contain the T-DNA insertion. M designates the molecular weight marker lane.

Confocal microscopy permitted the confirmation of transgene expression and gfp localization in mitochondria (Fig. 4). The potential for dual localization of the protein to plastids could not be definitively ascertained in these experiments.

Discussion
The predicted features of the candidate CHM-encoded protein suggest that the gene encodes the mitochondrial MSH1 counterpart in higher plants. MSH1 encodes a mitochondrial mismatch repair protein in yeast, although its counterpart in animals has not yet been identified. The CHM candidate sequence showed strongest homology with the Arabidopsis nuclear MSH6 sequence (Fig. 2), consistent with suggestions that nuclear mismatch repair components likely derived from a progenitor to MSH1 (36).

Although the predicted CHM candidate protein displayed several features suggesting its involvement in mismatch repair, lines containing mutations in the locus showed no evidence of mitochondrial point mutation accumulation. The primary effect within the mitochondrion seemed to be the reproducible substoichiometric shifting phenomenon. This assumption is based on the observation of identical mitochondrial DNA restriction fragments arising on substoichiometric shifting in all chm mutants when tested repeatedly (refs. 16 and 24 and this article). Moreover, no evidence of progressive decline in plant growth features has been observed over time. The chm1-1 and chm1-2 mutants, reported in the 1970s (25), appear identical to one another in phenotype and mitochondrial DNA configuration. Although detailed sequence analysis would be required to estimate the incidence of mismatch accumulation in the chm mutants, one would anticipate a random pattern of mitochondrial DNA polymorphism and progressive phenotypic decline in chm mutants were the mismatch accumulation rate enhanced.

Mutation of the MSH1 locus in yeast results in rapid accumulation of mitochondrial genomic rearrangements leading to disruption of mitochondrial function. Interestingly, a reproducible pattern of DNA restriction fragment polymorphism was reported in some of the petite mutants arising in yeast MSH1 mutant strains (37). This observation may be an indication that msh1-associated mitochondrial genomic rearrangements are similar in plants and fungi. Alignment between the yeast MSH1 protein and the Arabidopsis CHM (AtMSH1) candidate shows only 17% amino acid identity overall, with ≈28% identity within the predicted functional domains for ATP and DNA binding, but with well-conserved motifs (Fig. 2). The yeast MSH1 protein has been shown to have both DNA mismatch binding and ATPase activity (38, 39). These properties have not been tested in the CHM1 candidate from Arabidopsis.

The protein seemed to be highly conserved among plant species. We have identified several ESTs that present high homology (>60% amino acid identity) with AtMSH1 in soybean, tomato, potato, rice, wheat, and barley (data not shown). Although we have not yet tested for DNA binding and ATPase activities of the plant protein, identification of the chm1-3 mutation as a cysteine-tyrosine substitution within the predicted ATP binding domain does suggest the importance of this region to protein function. Substitution of the bulkier tyrosine would likely create distortion in the region, affecting ATP binding or hydrolysis.

Mismatch repair components seem to be involved in not only the binding and excision of nucleotide mismatches during the replication process, but also suppression of ectopic recombination (40, 41). Investigation of the mitochondrial substoichiometric shifting phenomenon suggests two alternative models for the influence of MSH1. It is conceivable that the AtMSH1 gene has shared or relinquished its mismatch repair function, such that its primary role in the plant mitochondrial genome is to regulate nonhomologous recombination. Disruption of MSH1 could thus
result in the enhancement of intramolecular ectopic recombination activity detected as apparent amplification of novel mitochondrial DNA forms. A possible weakness in this model arises in reports that several plant systems with mitochondrial DNA molecules susceptible to shifting seem to be derived from a DNA exchange that involved at least one molecular form no longer present in high copy number. Some also seemed to contain unique sequences. Therefore, the shifted molecules were thought to replicate autonomously (14, 20, 42).

If mitochondrial DNA molecules that undergo shifting are, in fact, replicated autonomously, an alternative model for molecule-specific substoichiometric shifting might apply. The Arabidopsis MSH1 product likely participates as a component of the DNA replication apparatus. Mitochondrial DNA molecules subject to copy number shifting may have originated by earlier ectopic recombination events during the evolution of the lineage. In this case, the resulting chimeric sites might serve to trigger a process of site-specific replication stalling by the MSH1 protein during vegetative growth.

Both models assume that the replicative form of the mitochondrial genome within meristematic (undifferentiated) tissues differs from that within vegetative (somatic) tissues. Hence, stoichiometric shifting events in vegetative tissues do not contain reversible loss of the suppressed genetic information. Presumably, the complete mitochondrial genetic complement is retained within the transmitting (meristematic) tissues (19, 42).

During higher plant evolution, natural allelic variation for the MSH1 locus likely revealed the adaptive advantage that arises from sporadic copy number modulation of mitochondrial genomic variants. Some of these variants, when amplified, condition male sterility that could facilitate advantageous outcrossing activity in natural populations (19). It merits noting that the mitochondrial DNA metabolism apparatus also seems to be important to the incidence of mitochondrial heteroplasmy in humans. Reports of nuclear, autosomal mutations leading to human mitochondrial genetic disorders are numerous. Autosomal dominant progressive external ophthalmoplegia (AdPEO; ref. 43) is associated with two loci on chromosomes 10 and 4 (44, 45). Interestingly, two nonsense mutations in the adenine nucleotide translocator (ANT) gene were identified on chromosome 4, and mutations in a gene encoding a mitochondrial protein with homology to the primase/helicase of the phage T7 were identified on chromosome 10 (46). Moreover, cases of autosomal recessive progressive external ophthalmoplegia (ArPEO) have been associated with the nuclear gene for mitochondrial gamma DNA polymerase on chromosome 15 (47). Finally, cases of mitochondrial neurogastrointestinal encephalomyopathy (MNGIE; ref. 48), an autosomal recessive mitochondrial disorder that leads to mitochondrial deletions, have been associated with mutations in a gene on chromosome 22 encoding thymidine phosphorylase (49).

It remains to be determined whether the substoichiometric intermediates common to higher plant mitochondrial genomes are maintained by replicative or recombinational mechanisms. This question may be resolved by the identification of proteins with which the Arabidopsis MSH1 protein interacts, as well as mitochondrial DNA sequences to which it binds. Experiments are currently underway to down-regulate the MSH1-homologous genes of other plant species to assess mitochondrial genome influence. With this article we identify an intriguing variation on the rule of eukaryotic MutS homologs that has contributed adaptive advantage by a strategy that very well may be unique to the plant kingdom.

We thank Ms. Sophia Alvarez for her efforts in determining MSH1 transcript size and the University of Nebraska Center for Biotechnology Microscopy Core Facility for assistance with confocal microscopy. This work was supported by grants from the National Science Foundation and the Department of Energy (to S.A.M.). R.V.A. was supported by a predoctoral fellowship provided by the Conselho Nacional de Pesquisas (Brazil). Funding for The Salk Institute Genomic Analysis Laboratory-indexed insertion mutant collection was provided by the National Science Foundation.