AN ABSTRACT OF THE THESIS OF

Aly Mohamed for the degree of Honors Baccalaureate of Science in Bioresource Research, Biotechnology presented December 1, 2006. Title: Repression of DNA Mismatch Repair (MMR) in *Arabidopsis thaliana* through Dominant-Negative MMR Proteins

Abstract approved:_____
John B. Hays

The highly conserved multi-protein mismatch repair (MMR) system is known for its ability to correct post replication errors in genomic DNA. A hallmark of MMR deficiency in all organisms is microsatellite instability. The initiating proteins in the eukaryotic MMR system are hetrodimers formed with an integral MSH2 subunit and one of three other subunits; MSH3, MSH6, or MSH7. Conserved helix-turn-helix and ATPase domains in the MSH2 protein have been identified as necessary for MMR function. Mutations in these domains, followed by over-expression of the mutant proteins, have resulted in dominant-negative phenotypes in *Saccharomyces cerevisiae*. Described here is a project in which Arabidopsis AtMSH2 proteins carrying mutations in their helix-turn-helix and ATPase domain were constructed in "super expression" vectors then transformed into *Arabidopsis thaliana* with the prospect of out-competing wild-type AtMSH2 produced by the plant for hetrodimeric association with MSH3, MSH6, or MSH7. An incapacitated MMR system causes increased post-replicative mutational accumulation in cells, which can be observed through microsatellite instability. Determining the capacity of the MMR system following transformation and over-expression of the mutant At*MSH2* gene constructs is completed through detection of allele shifts in microsatellite sequences by capillary gel electrophoresis.

Key Words: DNA repair**,** mismatch repair, dominant-negative, AtMSH2, MSH2
Corresponding e-mail address: aly.mohamed@gmail.com

Repression of DNA Mismatch Repair (MMR) in *Arabidopsis thaliana* through Dominant-Negative MMR Proteins

by

Aly Mohamed

A PROJECT

submitted to

Oregon State University

University Honors College

and Bioresource Research

in partial fulfillment of
the requirements for the
degree of

Honors Baccalaureate of Science in Bioresource Research, Biotechnology

Presented December 1, 2006
Commencement June 2006

Honors Baccalaureate of Science in Bioresource Research, Biotechnology thesis of Aly Mohamed presented on December 1, 2006

APPROVED:

Mentor, representing Environmental and Molecular Toxicology

Committee Member, representing Genetics and Microbiology

Committee Member, representing Bioresource Research

Dean, University Honors College

I understand that my project will become part of the permanent collection of Oregon State University, University Honors College and Bioresource Research. My signature below authorizes release of my project to any reader upon request.

Aly M. Mohamed, author

ACKNOWLEDGEMENTS

The author extends his sincere gratitude to
the following individuals in recognition
of their perseverance, professional assistance,
and extremely valued friendship.

Dr. Mohamed S. Mohamed

Dr. John B. Hays

Mr. Peter Hoffman

Mr. Awad El-garguri

Mrs. Stephanie Bollman

Dr. Walt Ream

Mrs. Wanda Crannell

Dr. Katharine G. Field

Dr. Kevin Ahern

Dr. Marc Curtis

Mr. Buck Wilcox

Dr. Huixian Wang

Mrs. Laurel Wheeler

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

DEDICATION

Dedicated with unconditional love and affection to the most cherished source of warmth in my life - my mother, Nagat Bishr El-Shinawy.

Repression of DNA Mismatch Repair (MMR) in *Arabidopsis thaliana* through Dominant-Negative MMR Proteins

Introduction

DNA Damage & Repair

Damaged DNA can arise from many endogenous and exogenous cellular stressors. Exposure to UV radiation, gamma radiation, X-rays, and industrial mutagenic chemicals exemplify some exogenous sources that lead to damaged DNA. Detrimental endogenous cellular processes are primarily of four types: oxidation of bases (e.g. 8-oxoguanine), alkylation of bases, hydrolysis of bases (deamination, depurination and depyrimidination), and mismatch of bases due to DNA replication in which a non-complementary DNA base is placed in a newly replicated DNA strand.

The proofreading ability that replicative polymerases provide is one of the most important genome-preserving mechanisms cells have evolved. Proofreading is a first line of defense in maintaining genomic integrity, single-handedly reducing base misinsertion frequency in one cell division to the range of 10^{-6} to 10^{-7} per base pair replicated (Voet and Voet 2004). Yet comparing this conservative misinsertion frequency with a genome size of ~4×10^9 bases (roughly the size of the human genome), a polymerase would generate approximately 400 mutations per cell division ($10^{-7} \times 4 \times 10^9 = 400$), an unacceptable mutation rate.

In human cells, normal metabolic activities and environmental factors can result in as many as one million individual molecular lesions per cell per day (Lodish *et al*,

2004). By employing a diverse array of genome surveillance mechanisms and cell cycle checkpoints, cells are able to minimize mutation rates, suppress genotoxic stress, and ultimately ensure their genomic stability (Adamson *et al*, 2005).

The MMR System

DNA repair pathways are important in suppression of mutational loading (or accumulation), which can lead to serious complications such as tumorigenesis, birth defects, and reduction of lifespan (Holland, 2003). One of the critical DNA damage response systems, known primarily for its ability to correct post replication errors in genomic DNA, is the multiprotein mismatch repair system (MMR). MMR also activates cellular checkpoints, promotes homologous recombination during meiosis, represses homeologous recombination between non-identical heterologous DNA sequences, and signals to cell-cycle-arrest or apoptosis pathways in response to certain types of DNA damaging agents (e.g. methylating agents such as O^6-methylguanine) (Hsieh, 2001). This system has evolved significant capacity to reduce spontaneous mutations by a factor of 10^2-10^3. Coupled with proofreading and base selection capabilities of DNA polymerases, the total error rate in cells is reduced by a factor of 10^{-9} to 10^{-10} per base pair per round of replication (Leonard *et al*, 2003).

In prokaryotic and eukaryotic organisms, MMR systems consist of a set of highly conserved proteins (e.g. Figure 1) (Culligan *et al*, 2000). Although details differ in prokaryotes and eukaryotes, the basic repair pathways are quite similar. MMR pathways can be divided into four stages: *detection* of mismatched base pairs or atypical DNA lesions, *nicking* of the nascent (newly synthesized) strand, *excision* of the DNA from the

nick to beyond the anomaly, and *resynthesis* of DNA. All MMR systems function through a long-patch repair pathway in which an extensive region of DNA, beginning at the nick site, is excised to remove an anomaly.

In the last decade MMR has been extensively studied. Although eukaryotic MMR systems are still less than fully understood, scientists have a remarkably complete understanding of MMR pathways in prokaryotes (Hsieh, 2001). The proteins which facilitate MMR in *Escherichia coli* have been extensively characterized, and therefore, the methyl-directed *E. coli* MMR pathway serves as a general model for more complex MMR systems in eukaryotes.

The protein responsible for initiating MMR in *E. coli* is MutS. This homodimeric protein is able to identify and bind mismatched base pairs with varying affinities, and extra-nucleotide DNA loop-outs resulting from insertion or deletion mutations, frequently caused by polymerase slippage in highly repetitive DNA regions. When bound to a pre-mutagenic DNA substrate, the MutS protein associates with the MutL homodimer, forming a MutS-MutL protein complex (Voet and Voet, 2004).

In *E. coli*, palindromic d(GATC) sequences are methylated on their adenine residues shortly following DNA replication, giving rise to hemimethylated d(GATC) sequences, appearing only transiently immediately after replication. The strand bias occurring as a consequence of post-replication delay of d(GATC)-adenine methylation is used by the *E. coli* MMR system to distinguish the nascent DNA strand.

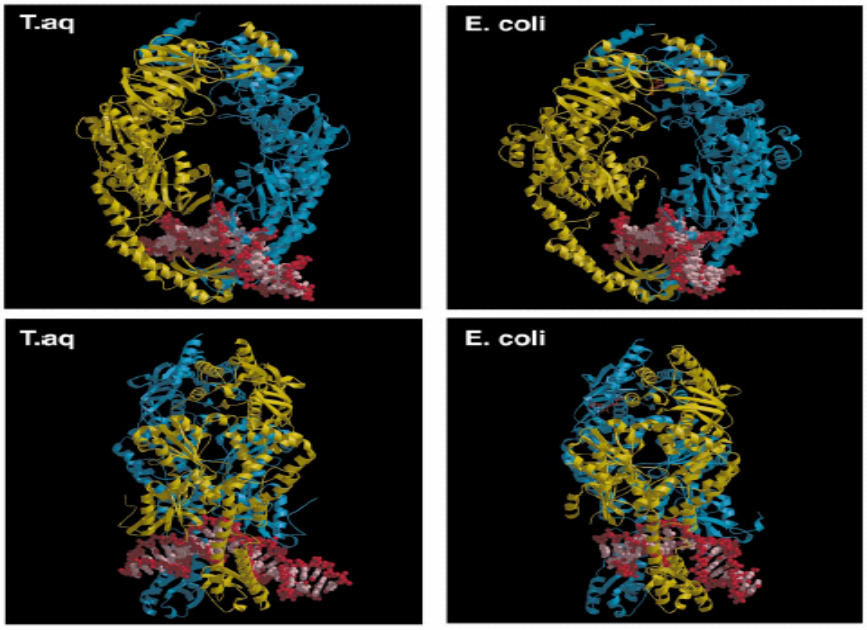


Figure 1: *Center and side views*, DNA and MutS proteins of *T. aquaticus* and *E. coli* as represented by X-ray crystallography reveals striking structural identity (Hsieh, 2001).

The MutS-MutL complex is capable of searching and locating the nearest d(GATC) palindrome, though the mechanism coupling mismatch recognition to site-specific excision is controversial. One model suggests that after binding, the MutS-MutL complex is induced by ATP hydrolysis or binding to slide along the DNA towards the nearest d(GATC) excision sites. A second model contends the MutS-MutL complex remains in place after binding and contacts the nearest d(GATC) excision site by DNA looping.

Located at d(GATC) excision sites are MutH proteins which function specifically in recognizing hemimethylated d(GATC) sequences. In the presence of ATP, the MutS-MutL complex is able to stimulate the endonucleolytic function of the MutH protein which, upon stimulation, introduces a single strand nick in the nascent DNA.

After nicking of the unmethylated d(GATC) sequence, helicase II (UvrD) loads at the nick site and seperates the DNA strands to facilitate excision of the nicked DNA strand by single strand exonucleases which excise to some end beyond the anomaly. Resynthesis of the excised DNA involves DNA Polymerase III and single strand binding proteins, followed by ligation by a DNA ligase protein, restoring DNA integrity.

Eukaryotic MMR proteins have evolved into several homologs sharing many conserved features found in prokaryotic MMR proteins. The eukaryotic counterparts of the MutS and MutL proteins are the MSH (MutS homolog) and MLH (MutL homolog) proteins. Seven different MutS homologs and four MutL homologs are known to date; MSH1 through MSH7 (with MSH7 occurring only in plants), and MLH1, PMS1 (post meiotic segregation 1), PMS2, and MLH3 (Luo *et al*, 2004). There have been no MutH homologs found in eukaryotes (Depeiges *et al*, 2005). The absence of a eukaryotic MutH homolog coupled with the absence of DNA-adenine methylation is indicative of a eukaryotic-specific mechanism for determining repair strand specificity, perhaps involving recognition of growing 3' ends of nascent DNA and 5' ends of Okazaki fragments.

The chief eukaryotic mismatch recognition proteins are heterodimers, formed with an integral MSH2 support subunit. The MSH2 protein can have one of three recognition subunits complexed with it, forming MSH2•MSH3 (MutSβ), MSH2•MSH6 (MutSα), and MSH2•MSH7 (MutSγ).

Substrate specificities of the MutSα and MutSβ complexes are well defined, with slight overlap between the two protein complexes known to exist. MutSα best recognizes single base mismatches and small insertion/deletion loopouts, and MutSβ primarily

recognizes a range of larger extrahelical loop-outs. Despite this difference, both complexes interact with the MutLα (MLH1-PMS2) complex, coupling recognition to excision of the error-containing DNA, and both are found in virtually all eukaryotes including humans, mice, plants and yeast. The function of MutSγ in mediating error correction is less than fully understood, but it thought to recognize a subset of base mismatches that are generally detected by MutSα (Culligan and Hays, 2000).

The specific protein of concern in this project is the AtMSH2 subunit of the model plant *Arabidopsis thaliana*. *Arabidopsis thaliana* is a small flowering plant related to cabbage and mustard. Though *Arabidopsis thaliana* is of little agricultural importance, it is possibly the most frequently used model organism for studying cell and molecular development of flowering plants.

There are several advantages to using *Arabidopsis thaliana* as a model organism. Because of the small size of its genome (five chromosomes, 157 million base pairs), the plant was completely sequenced by the year 2000. *Arabidopsis thaliana* is very small in size and has a notably rapid life cycle of approximately six weeks from germination to progeny seed production. A small and healthy *Arabidopsis thaliana* plant is capable of producing thousands of seeds. The selfing nature of this plant is a helpful feature in many genetic experiments. Transformation of foreign DNA into *Arabidopsis thaliana* genomic DNA using *Agrobacterium tumefaciens* is a simple and routine procedure which does not involve tissue culture or plant regeneration.

Microsatellite Instability

Cells with deficiencies in the MMR pathway typically display mutator phenotypes characterized by an increased incidence of spontaneous mutations (Hsieh, 2001). Microsatellite sequences are simple mono-, di-, or tri- nucleotide repeats found in all eukaryotes. These sequences are substrates for MMR correction because they tend to give rise to extrahelical DNA loop-outs. These loop-outs are formed during DNA replication when the DNA is annealed out-of-frame resulting in template slippage resulting in potential insertion or deletion mutations (Figure 2). Loop-outs are effectively corrected by MMR in wild-type cells, but not as well by polymerase proofreading.

Microsatellite instability (MSI, the expansion or contraction of microsatellite sequences) is a hallmark of MMR deficiency in all organisms. Cell lines that display a loss of function mutation in either MSH2 or MLH1 have yet to be found without MSI (Cejka *et al*, 2003).

A notable advance in the study of MMR was the discovery that DNA repair system failure is associated with certain forms of cancer. In particular, cell lines isolated from individuals with hereditary non-polyposis colorectal carcinoma (HNPCC), also called Lynch Syndrome, display a marked dysfunction in MMR proteins and are characterized by MSI (Charames *et al,* 2003).

Figure 2: Template slippage during DNA replication of microsatellite sequences results in insertion or deletion pre-mutagenic loopouts. If loop-outs are not repaired by the MMR system before the next round of replication, the loop of extra-nucleotides are replicated, resulting in mutation and creating a possibility for frame-shifts. (Modified and used with permission from Stephanie Bollman)

Dominant-Negative MMR Proteins in Arabidopsis

The MSH2 protein and its heterodimeric interactions with MSH3, MSH6, or MSH7 are essential for the function of eukaryotic MMR. Together, these proteins initiate MMR in plants. Previous At*MSH2* knock-out studies in *Arabidopsis thaliana* have shown the deleterious phenotypic effects of completely removing the At*MSH2* gene.

Conserved helix-turn-helix and ATPase domains in the MSH2 protein have been identified as necessary for MMR. Previous studies in *Saccharomyces cerevisiae* have identified point mutations in these domains that do not compromise MSH2 heterodimeric protein-protein interactions or the resultant ability to bind mismatched DNA, but do eliminate capability to carry out complete MMR (Studamire *et al*, 1998)(Alani *et al*, 1997). These so-called "separation-of-function" mutations render the protein complex effective in locating and binding defects in DNA but do not allow it to interact with proteins downstream in the MMR pathway. The over-expression of these mutant MSH2 proteins resulted in a dominant-negative phenotype. A dominant-negative phenotype develops when mutated gene products block some aspect of the function of wild-type gene products within the same cell, usually by competitive interaction with other proteins normally associated with the wild-type gene products.

The purpose of this project is to establish in plants either of two deliberately mutated AtMSH2 proteins that are aberrant in function. The At*MSH2* gene was genetically modified to imitate the two dominant-negative alleles in *Saccharomyces cerevisiae*. At*MSH2* was mutated such that either glycine 671 (of the helix-turn-helix domain) or glycine 833 (of the ATPase domain) was replaced with aspartic acid. These mutant AtMSH2 proteins were over-expressed in *Arabidopsis thaliana* using a binary

vector with an *Agrobacterium* derived 'super-promoter' in front of the gene. The dominant-negative AtMSH2 protein is over-expressed with the expectation of out-competing wild-type AtMSH2 produced by the plant.

A short c-MYC epitope tag was fused to the 3' end of the At*MSH2* gene for expression at the carboxyl termini of the protein. An epitope tag is an antigen detected by a readily available antibody, which targets the epitope independently of the remainder of the fusion protein. For analysis, the AtMSH2 protein sample is first separated according to molecular weight using polyacrylamide gel electrophoresis. In order to make the proteins accessible to antibody detection, they are blotted from within the gel onto a membrane made of Polyvinylidene Fluoride (PVDF). The membrane is placed face-to-face with the gel, and current is applied. The charged proteins move from within the gel and are immobilized on the membrane while maintaining the organization they had within the gel. As a result of this "blotting" process, the proteins are exposed on a thin surface layer for detection.

Steps are taken to prevent non-specific protein interactions between the proteins on the PVDF membrane and the antibody used for detection of the target protein. Blocking of non-specific antibody to protein binding is achieved by placing the membrane in a dilute solution of protein from non-fat dry milk, with a minute percentage of detergent such as Tween 20. After blocking, a dilute solution of primary antibody is incubated with the membrane under gentle agitation for several hours and washed repeatedly.

The primary antibody::epitope complex is probed with a secondary antibody, which in turn is linked to a reporter enzyme such as horseradish peroxidase. A

horseradish peroxidase-linked secondary is used in conjunction with a chemiluminescent agent, and the reaction product produces luminescence proportional to the amount of protein present. Photographic film is exposed to the blot and creates an image of the antibodies bound to the blot. This procedure known as "immunoblotting" has been performed on plant protein extracts to quantify protein production in plants.

In *Arabidopsis thaliana*, over-expressed AtMSH2 mutant proteins would hypothetically create a dominant-negative effect by out-competing wild-type AtMSH2 for heterodimeric association with MSH3, MSH6, and MSH7. The objective is to genetically engineer and over-express in plants mutated AtMSH2 proteins that can mask wild-type AtMSH2 proteins produced by the plant in a dominant-negative manner. If this can be verified, these plants might be useful in more strenuously testing the hypothesis that MMR activity is important in maintaining plant genomic integrity and promoting homologous recombination during meiosis. The plants could also be used to generate a variety of mutant progeny which can be back-crossed at will and screened for elimination of dominant-negative At*MSH2* alleles.

A dominant-negative allele is preferred over a knocked-out allele because knock-out plants produce heterozygous progeny containing recessive knock-out alleles. F1 generation *Arabidopsis thaliana* At*MSH2* knock-outs would have the wild-type allele in every single plant. A complete At*MSH2* knock-out is only achieved with plants that are homozygous for the knock-out allele, which can only be achieved in one fourth (25%) of the F2 generation of the plants. In contrast, the dominant-negative allele competes with wild-type At*MSH2* alleles as soon as it is introduced.

Post-Transcriptional RNA Interference in Arabidopsis

In plants, post-transcriptional RNA interference (RNAi) serves to suppress or inactivate target RNAs and is hypothesized to have evolved as an antiviral defense. The efficiency of RNAi in plants varies among transformed plants. In consideration of the possible gene silencing effects, two *Arabidopsis thaliana* plants that are defective in RNAi pathways were obtained to for use in this project.

Small RNAs belong to two general classes, microRNA (miRNA) and short interfering RNA (siRNA). miRNAs are approximately 21–22 nucleotides in length. They develop from imperfectly base-paired foldback structures from non-protein-coding RNA transcripts.

Structurally, siRNAs are similar to miRNAs, although their size ranges between 21 and 24 nucleotides in plants. siRNAs are processed from precursors containing extensive or exclusive double-stranded RNA structure, such as intermediates formed during viral replication. Double-stranded RNA processed into siRNA triggers posttranscriptional gene silencing (PTGS). In plants, PTGS is involved in gene silencing of transgenes.

An RNaseIII helicase-like enzyme named DICER (DICER-LIKE or DCL1 in *Arabidopsis thaliana*) interacts with double-stranded RNA and fold-back structures to excise mature siRNA and miRNA, respectively. RNA-induced silencing complex (RISC) proteins bind small RNAs using an ARGONAUTE (AGO) protein. The binding of the small RNAs to RISC enables the complex to use the small RNAs as a guide sequence to complementary target sequences (Figure 3)**.**

In Arabidopsis, two genes are believed to be involved in transforming single-stranded RNA into double-stranded RNA. Initially identified by screening siRNA - defective Arabidopsis mutants, *SUPPRESSOR OF GENE SILENCING3* (SGS3) is a protein of unknown function, and *SUPPRESSOR OF GENE SILENCING2* (SGS2)/*SILENCING DEFECTIVE1* (SDE1) encodes RNA-dependent RNA polymerase6 (RDR6). Both of these genes are believed to be involved in transforming ssRNA into dsRNA precursors of siRNAs (Peragine *et al*, 2004). The RDR6 protein is known to prime and synthesize double-stranded RNA, establishing it as a crucial component in PTGS that arises from sense-transgene-mediated RNAi. RDR6 does not play a role in creating miRNA precursors from double-stranded RNA hairpins (Xie *et al*, 2004).

Taking into consideration the possibility of At*MSH2* gene silencing, two RDR6-deficient *Arabidopsis thaliana* plants were obtained for use in this project. RDR6-15 is ecotype Columbia and RDR6-4 is ecotype Lansberg. These plants were used in case At*MSH2* dominant-negative alleles were subjected to gene silencing, causing them to be less than fully over-expressed in wild type Columbia transformants.

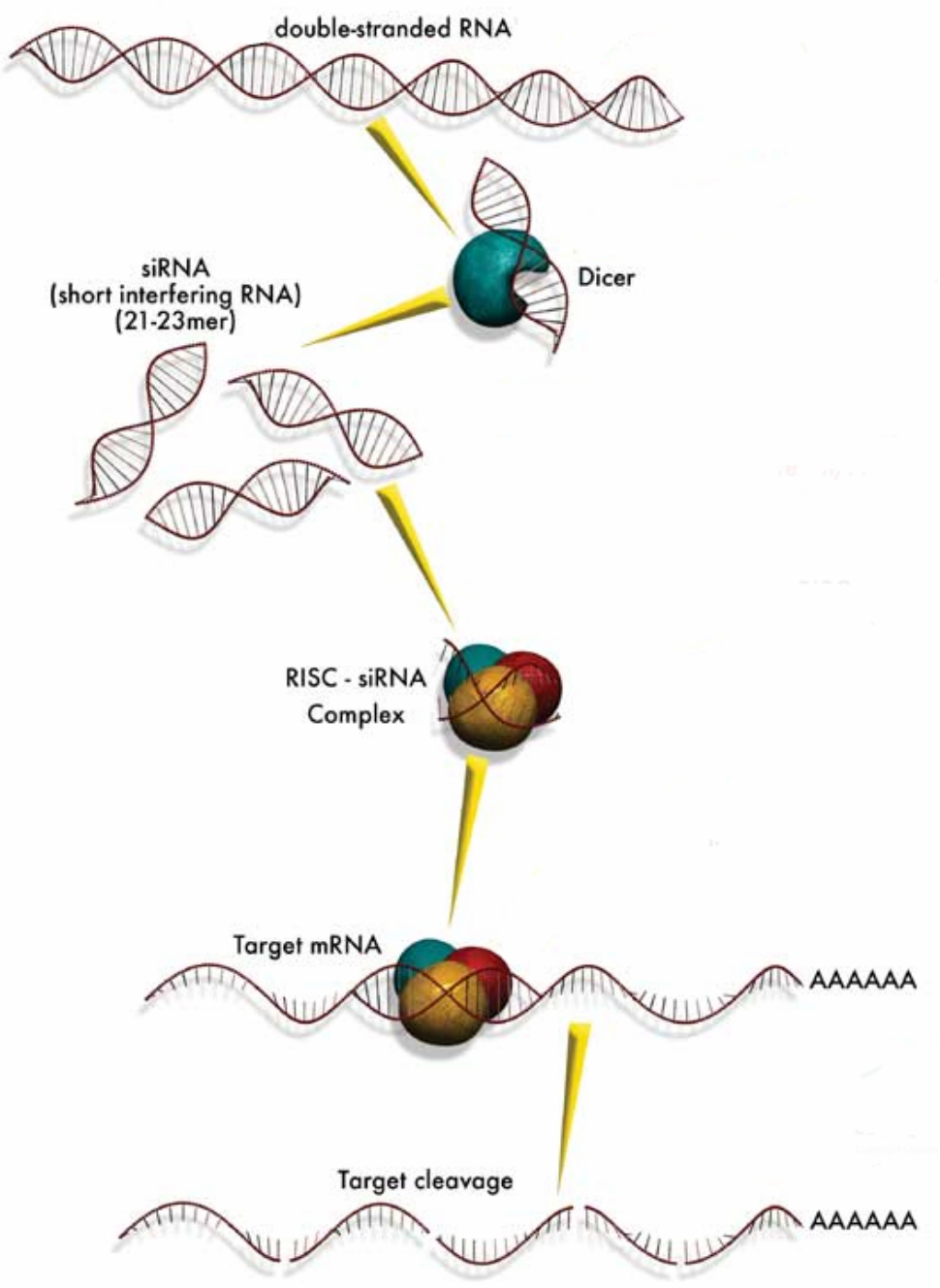


Figure 3: The siRNA pathway (Upstate Inc.)

Materials and Methods

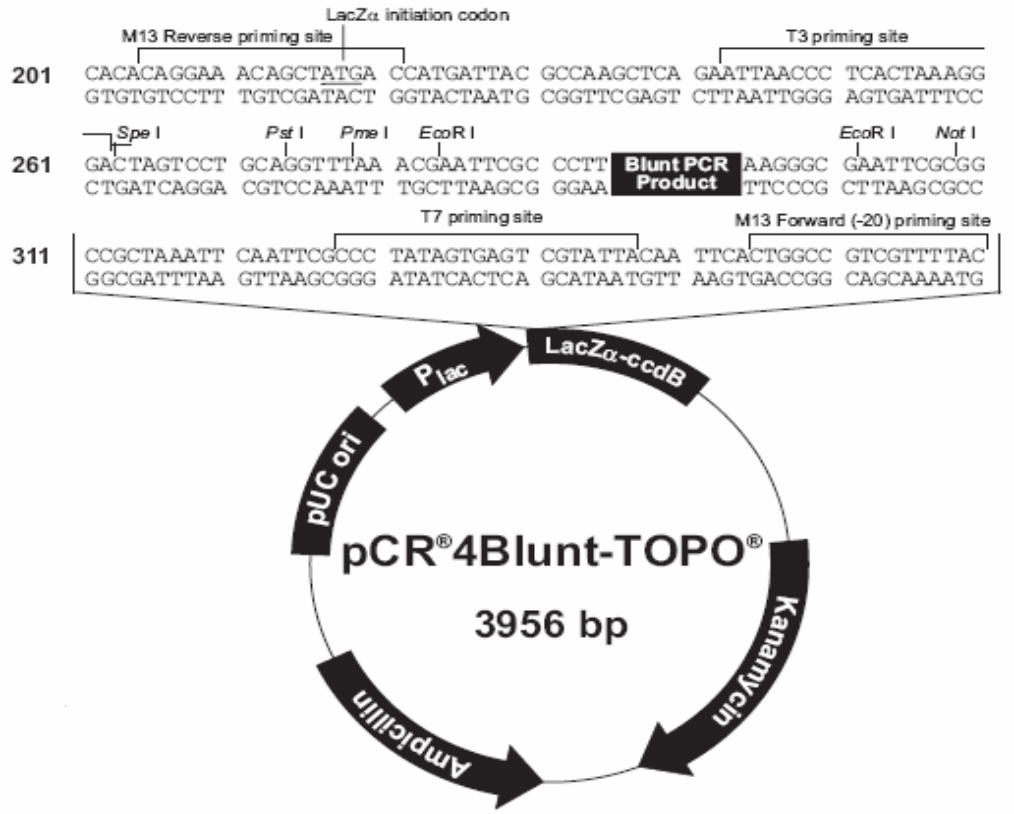
Assembly of At*MSH2* Dominant-Negative Gene Constructs

Site-directed mutagenesis was performed through PCR to create the altered At*MSH2* alleles. The first allele, designated dominant-negative1 (DN-1), was mutated in bases 2021-2022, corresponding to the ATPase domain of AtMSH2. The second allele, designated dominant-negative2 (DN-2), was mutated in bases 2507-2508, corresponding to the helix-turn-helix domain of AtMSH2. In both alleles the point mutations corresponded to a guanine to adenine transition in the first base and guanine to thymine transversion in the second base. The AtMSH2 protein was mutated such that either glycine 671 (of the helix-turn-helix domain) or glycine 833 (of the ATPase domain) was replaced with aspartic acid.

DNA fragments encoding the At*MSH2* dominant-negative alleles were cloned into pBluescript SK+ (Stratagene). At*MSH2* PCR products were generated using primer MSH2-5' (5′-GGTACCACCATGGAGGGTAATTTCGAGGACAGAAC-3′) encoding a *KpnI* restriction site upstream of the initiating ATG codon of the At*MSH2* gene, and MSH2/c-MYC-3' (5′-GCAGACTGCCACTGGCTCAGGCAGTTTCTGGAACAAAAACTTATTT CTGAAGAAGATCTG*TAATGA*CCTAGGCG-3′) encoding a c-MYC epitope tag followed by two stop codons and an *AvrII* restriction site. At*MSH2* PCR products were ligated into the pCR4Blunt-Topo plasmid (Invitrogen; Figure 4a) using standard protocols. Briefly, PCR products were incorporated into a table top reaction mixture containing the linearized plasmid product attached to DNA topoisomerase I, which functions both as a restriction enzyme and as a ligase (Invitrogen; Figure 4b). Topoisomerase I functions *in*

vivo to cleave and rejoin DNA, eliminating the stress of supercoiling incurred during DNA replication. The intermediate At*MSH2*-plasmid products were transformed into *E. coli* strain DH5α and were maintained using kanamycin selection.

(a)



(b)

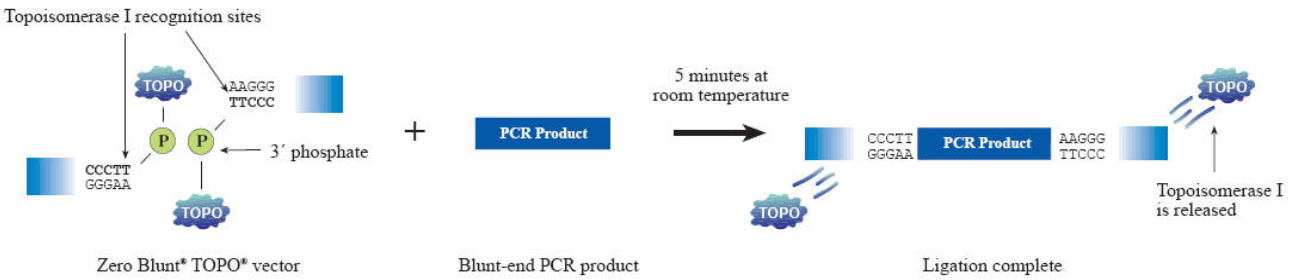


Figure 4: (a) pCR4Blunt-Topo plasmid map (b) depiction of Zero Blunt Topo reaction mechanism. (Invitrogen)

Restriction endonucleases *Bam*HI, *Eco*NI, *Eco*RI, *Ssp*I, *Pst*I, *Not*I, and *Spe*I, were used in preliminary restriction digest analysis of transformed product. Many of these enzymes targeted sites inside of the At*MSH2* gene, while others targeted sites on the intermediate or binary vector. At*MSH2* dominant-negative alleles in the intermediate plasmids were sequenced in their entirety using At*MSH2* gene-specific primers in addition to T3 and T7 priming sites on the pCR4Blunt-Topo plasmid (see Figure 4a for T3 and T7 priming sites).

Restriction of the sequence-confirmed intermediate plasmids with *Avr*II and *Kpn*I endonucleases released ~2.7-kb fragments encoding the mutated At*MSH2* alleles. These fragments were ligated into binary vector pE1803 cleaved with *Avr*II and *Kpn*I (Figure 5).

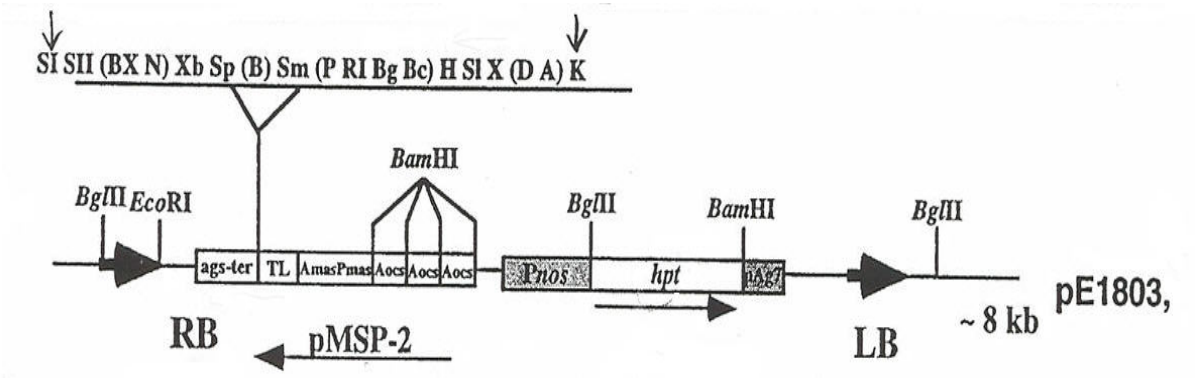


Figure 5: pE1803 binary vector map.

At*MSH2*-pE1803 binary vector ligation products were initially transformed into *E. coli* strain DH5α and were selected and maintained using kanamycin resistance. Restriction endonucleases *Avr*II and *Kpn*I were used in restriction digestion analysis of

transformed product and yielded a 2.7kb fragment. At*MSH2*-pE1803 vector constructs were then transformed into *Agrobacterium tumefaciens* strain EHA105 and maintained by kanamycin selection. The At*MSH2* dominant-negative alleles reconstituted in the binary vector were again sequenced in their entirety using At*MSH2* gene-specific primers in addition to binary-vector-specific priming sites.

Transformation into Plants

Arabidopsis Thaliana lines used were RDR6-15 (ecotype Columbia-0), RDR6-4 (ecotype Lansberg), and Columbia-0. To prevent soil from falling into inoculation medium after seeds were planted the soil was covered with a nylon window screen fabric secured by a rubber band. Plants were grown to flowering stage, occurring one month (30 days) after germination, and were clipped after forming primary bolts to obtain more floral buds per plant and to promote synchronized emergence of secondary bolts.

Agrobacterium tumefaciens cultures were started from a 1:100 dilution of an inoculum grown for ~24 h. All bacteria were grown at 30°C in an incubator shaking at 250 rpm in sterilized TBY broth (10g tryptone, 5g yeast extract, 10g NaCl per litre water) with added kanamycin (50 μg ml^{-1}). Cells were harvested by centrifugation for ~10 min at 4°C at 5500G. Cells were then resuspended in Floral Dip Inoculation Medium of 5.0% sucrose and 0.05% Silwet L-77 (Lehle seeds Inc.) to a final A_{600} of 0.80.

Plants were dipped when secondary bolts were 1–10 cm tall (4–8 days after clipping). The solution of resuspended bacteria was added to a beaker and plants were inverted into the suspension. The plants were promptly removed after 3–5 sec of agitation in solution. For vacuum infiltration, plants submerged in inoculation medium were placed

into a chamber and a vacuum was applied until air bubbles were drawn from plant tissues. The vacuum was held for ~5 min, and then released rapidly. Dipped or vacuum-infiltrated plants were placed in plastic trays and covered with clear-plastic domes to maintain humidity. Plants were left in the dark overnight. Domes were removed approximately 12–24 h after treatment.

Plants were further grown for 3–5 weeks and were dipped in transformation medium 2-3 additional times before collecting seed. Plants were then watered less frequently to promote maturation. Roughly 1,000 seeds were collected for every plant dipped. Care was taken to keep bolts from each pot separated from neighboring pots. Extra plant material was removed by sifting the seeds through a metal screen several times.

Screening for Transformants

Seeds were surface-sterilized by treatment with 95% ethanol for one minute and 50% household bleach containing .05% Tween 20 for 10 min, followed by three or four washes with sterile water.

Primary transformants were selected by plating seeds onto media containing 0.5X Murashige and Skoog salts (Qbiogene), B5 vitamins, 3% Sucrose, 0.75% agar, and 50 mg L^{-1} hygromycin. Seeds were distributed onto petri dishes and incubated at 4°C for 24 h, then placed in a growth chamber at 24°C under constant illumination. Petri plates and lids were sealed with parafilm to prevent fungal contamination of media.

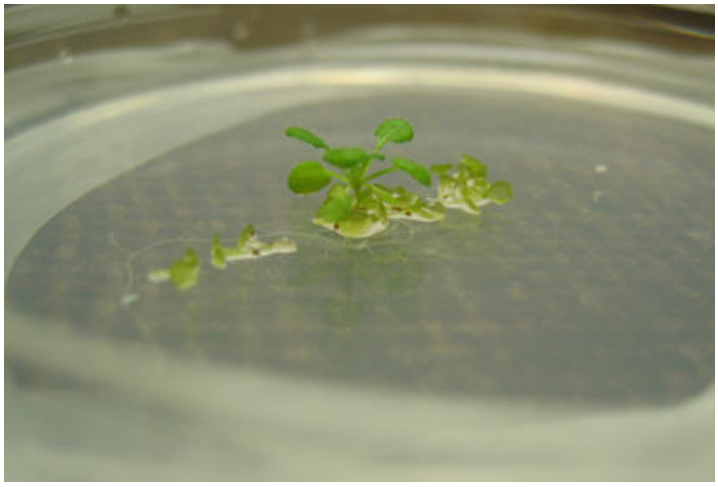


Figure 6: A prospective transgenic plant grown on agar containing hygromycin selection.

Purification of Plant DNA

Plants tissue used for DNA preparations was isolated from progeny screened against hygromycin selection. The tissue was ground with a pestle in a 1.5 ml microtube using Extraction Buffer of 200 mM NaCl, 25 mM EDTA, 50 mM Tris pH 7.5, and 0.5% SDS. After the grinding procedure, the remaining tissue was spun down, and the supernatant was collected in a separate 1.5 ml microtube. DNA in the supernatant was precipitated by mixing 1 ml of 95% isopropyl alcohol into solution then spinning at 12,000 rpm for one minute in the tabletop centrifuge. The resultant pellet of DNA was allowed to dry briefly and was then eluted in 100 μL of DNA buffer (10 mM Tris, pH 7.0 1mM EDTA).

Confirmation of Transgenic Plants

The progeny of prospective transformants were confirmed to be genuine by PCR amplification. One primer sequence targeted the binary vector promoter 350 base pairs upstream of the At*MSH2* dominant-negative start site (primer BVup), while the other primer targeted a site ~850 base pairs down stream of the At*MSH2* start site (primer 850 Reverse Complement). PCR of genuine transformants revealed a predicted band of ~1.2kb while non-transformants showed no product amplification.

Immunoblotting

Protein expression is to be analyzed by standard immunoblot protocol. Equal amounts of plant protein extract are to be resolved by SDSPAGE using 12% polyacrylamide-SDS gels. The proteins are then blotted from the gel onto a thin surface layer of PVDF. AtMSH2 proteins are to be detected using anti-c-MYC (ROCHE pharmaceuticals, San Diego, CA). Protein extracts are to be incubated with anti-c-MYC antibody in PBS for 2 h at room temperature. The peptide-antibody mixture is to be diluted in 5% milk and incubated to prevent non-specific protein interactions with the membrane. The antibody::epitope complex is then probed with a secondary antibody conjugated to a fluorescent, chemiluminescent or colormetric reporter. A sheet of photographic film is placed against the membrane, and exposure to the light from the reaction creates an image of the antibodies bound to the blot.

Analysis of MSI

Samples of plant DNA (see Purification of Plant DNA above) are to be analyzed as previously described (Leonard *et al*, 2003). Briefly, Six dinucleotide microsatellite loci (NGA6, NGA8, NGA139, NGA151, NGA172, and NGA1107) will be used to PCR-amplify DNA using one fluorescently marked primer (labeled with hexachloro-6-carboxyfluoroscein or 6-carboxyfluoroscein fluorescent dyes) and one unlabeled primer for each locus.

The following forward and reverse primers are to be used:

NGA6, 5′-GACTAAAGTGGGTCCCTTGG-3′
and 5′-CACACCCAAAACTCGTAAAGC-3′;

NGA8, 5′-TGGCTTTCGTTTATAAACATCC-3′
and 5′-GAGGGCAAATCTTTATTTCGG-3′;

NGA139, 5′-GGTTTCGTTTCACTATCCAGG-3′
and 5′-AGAGCTACCAGATCCGATGG-3′;

NGA151, 5′-CCAGAGCTTGTTTTGGGAAG-3′
and 5′-TTTGATGAAACGGAATATAGAAAGC-3′;

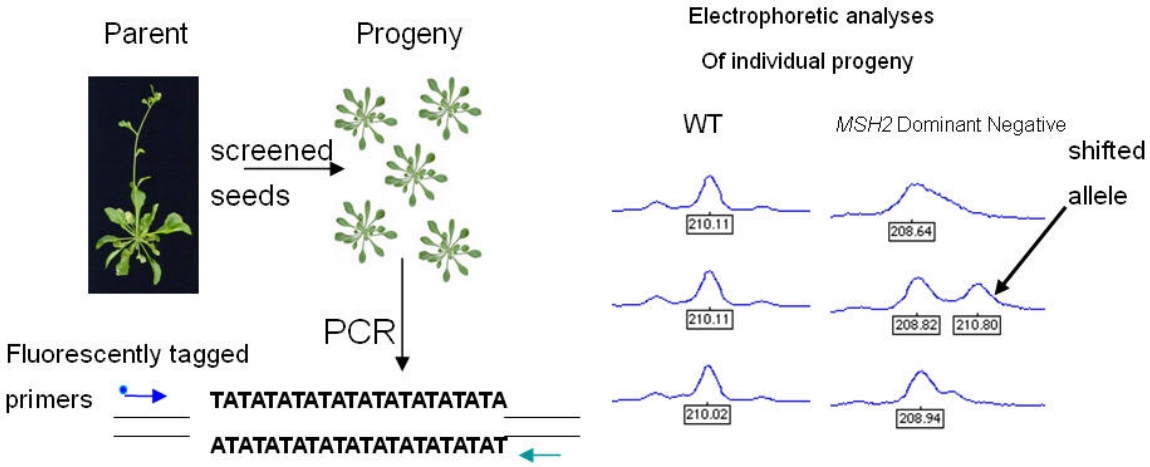
NGA172, 5′-CATCCGAATGCCATTGTTC-3′
and 5′-AGCTGCTTCCTTATAGCGTCC-3′;

NGA1107, 5′-CGACGAATCGACAGAATTAGG-3′
and 5′-GGCTACAATAGTGGGAAAACG-3′.

Standard PCR conditions are as follows: Initial denaturation at 94°C for 3 min, followed by 30 cycles of 94°C for 30 s, 54°C for 30 s, and 72°C for 30 s, and a final one hour incubation at 72°C. Samples are to be analyzed by capillary gel electrophoresis and

quantitative fluorescence detection using an ABI Prism 3100 Genetic Analyzer and associated software (Applied Biosystems). Electrophoretic profiles for each locus for each sample will be used to detect increases or decreases (shifts) in the number of dinucleotide repeats.

(a)



(b)

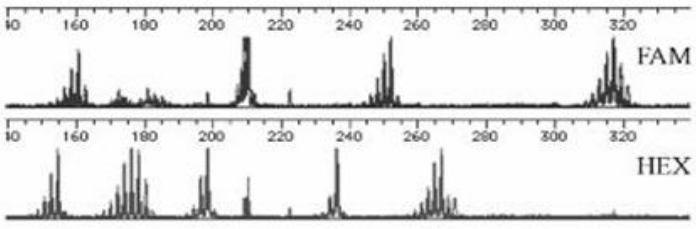


Figure 7: Microsatellite instability detection. (a) Endogenous microsatellite instability detected by length plymorphisms. Primers specific to microsattelite loci carrying dinucleotide repeats are used to amplify DNA prepared from individual progeny of either wild-type or transgenic parents. Fluorescently labeled PCR products were analyzed using the ABI 3100 capillary electrophoresis analyzer and proprietary ABI software. (b) Electropherograms showing PCR products; one of each primer pair was labeled with FAM or HEX fluorescent dye. (Figure 7(a) used and modified with permission from Stephanie Bollman)

Results

Dominant-negative At*MSH2* Alleles in *Arabidopsis*

At the time of writing, this project is slightly past three quarters complete. One set of plants (Col-0, RDR6-15, RDR6-4) have been transformed with the At*MSH2* DN-2 construct. In the set, several independent transformants were obtained for RDR6-15, while only one transformant was obtained for Col-0 and RDR6-4. Although several prospective DN-1 transformants have been screened and are being prepared for verification, to date there have been no verified transformants isolated with this construct.

Confirmation of Microsatellite Instability Testing Protocols

Five DN-2 transformants were propagated to confirm DNA extraction, PCR, and microsatellite instability testing protocols before proceeding with larger scale experiments. DNA was extracted from two separate RDR6-15 transformants (F2) and three individual progeny of a single Col-0 transformant (F2). PCR reactions yielded highly variable results, often producing no quantifiable amplification of microsatellite sequences. PCR performed with NGA6 and NGA172 primers yielded faint PCR products, while NGA8, NGA139, NGA151, and NGA1107 primer sets yielded no PCR products at all (Figure 8).

Plant DNA samples were loaded after dividing a homogenous PCR cocktail consisting of polymerase, dNTPs, 2mM $MgCl_2$, 10X PCR buffer, and water for each reaction. Given that these ingredients were shared for every reaction, they were initially

eliminated as factors contributing to failed PCR. In contrast to the varied microsatellite PCR reactions, PCR amplification of transgene products from plant DNA samples (using primers specific to the binary vector and the dominant-negative At*MSH2* gene) consistently yielded high target sequence amplification. Plant DNA samples used to confirm integration of the At*MSH2* DN-2 transgene into genomic DNA were the same as those used to perform microsatellite PCR amplification. Initially it was suspected the underlying reason for failed PCR amplification was degradation of the primers which were thought to have been subjected to repetitive thawing and freezing, possibly leaving the primers non-functional.

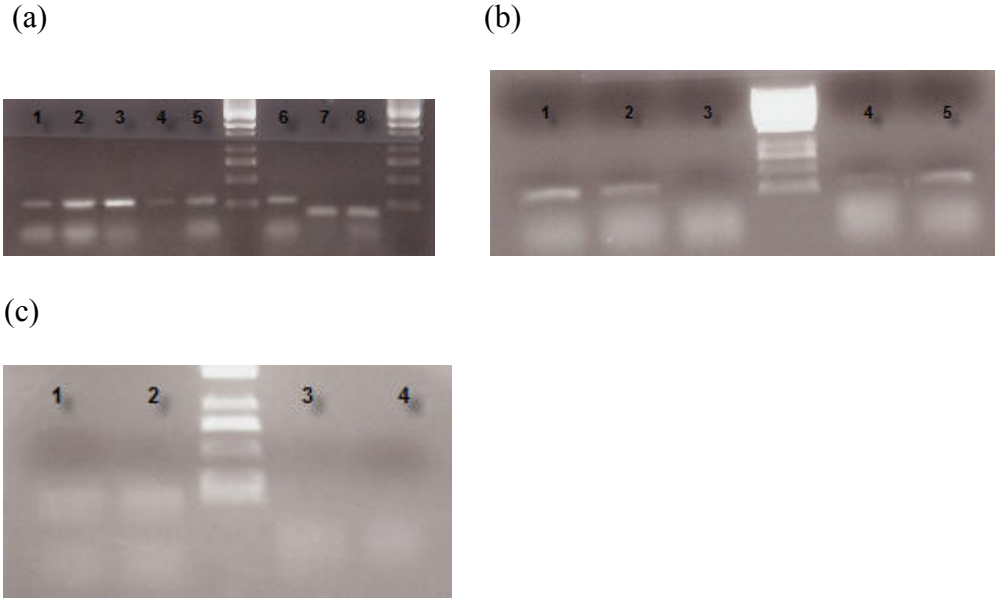


Figure 8: PCR amplification of microsatellite targets shows a range of product output under identical conditions. (a) PCR samples from a DN-2 RDR6-15 transformant using NGA 6-1FAM/NGA6-3 primer set for lanes 1-3, NGA6-1HEX/NGA6-3 primer set for lanes 4-6, and NGA172FAM primer set for lanes 7-8. (b) PCR samples from five separate transformants. Lanes 1-3 are from DN-2 Columbia, plants D, C, and B respectively. Lanes 4 and 5 were isolated from two independent DN-2 RDR6-15 transformants distinguished as plants two and three respectively. All samples were amplified using the NGA6-1FAM/NGA6-3 primer set. PCR using the NGA172FAM/NGA172R primer set yielded no results. No results are observed for plant B in lane three. (c) Lanes 1 and 3 are samples from DN-2 RDR6-15 plant two. Lanes 2 and 4 are DNA samples from DN-2 RDR6-15 plant three. Lanes 1-2 were amplified using the NGA172FAM/NGA172R primer set, while lanes 3-4 used the NGA151-4FAM/NGA151-2 primer set. No results are observed in lanes 3-4.

NGA6 and NGA172 PCR samples were diluted according to their perceived signal strength on gel then analyzed for microsatellite instability via capillary gel electrophoresis (Figure 7). Most samples required further dilution for accurate reading. The few samples which were accurately read displayed no deviations from wild type microsatellite sequences. Upon confirmation of the microsatellite instability detection protocol, analysis of the five samples was abandoned, and despite that four of the six primer sets (NGA8, NGA139, NGA151, and NGA1107) were not yielding PCR amplification of their target sequence, focus was shifted to preparing a larger sampling of transformed plants for complete microsatellite instability analysis.

Microsatellite Instability Sample Preparation

To investigate microsatellite instability in the dominant-negative At*MSH2* transformants, a set of fifty seeds from two separate RDR6-15 transformants (F2) and two progeny of a single Col-0 transformant (F2) were propagated on agar containing hygromycin selection. Segregation analysis indicated that all progeny were homozygous for the dominant-negative At*MSH2* allele.

A set of eight hygromycin-resistant plants was selected from each set of fifty for DNA extraction and microsatellite instability analysis (32 plants total). Instead of using $MgCl_2$ at 2 mM concentrations as previous, NGA6 and NGA172 PCR was performed at a final concentration of 4 mM $MgCl_2$, giving excellent PCR amplification results for every reaction. Unfortunately, a 4 mM $MgCl_2$ concentration did not yield any amplified product with NGA8, NGA139, NGA151, or NGA1107 primer sets. As a last ditch effort, the primers were used in a PCR with final concentration of 6 mM $MgCl_2$. Surprisingly, under

this reaction condition all four primer sets displayed good PCR amplification (Figure 9). Microsatellite PCR amplification of transgenic DN-2 plant DNA has been completed successfully. All samples were diluted according to their perceived signal strength on gel and analyzed for microsatellite instability by capillary gel electrophoresis.

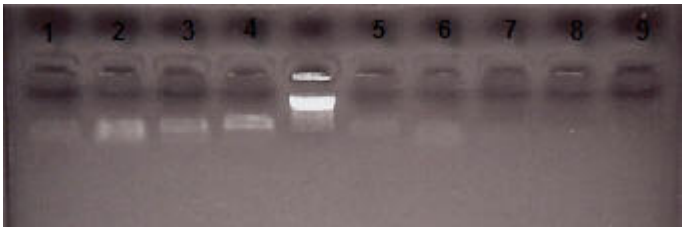


Figure 9: PCR amplification using varied $MgCl_2$ concentrations. DNA in lanes 1-4 were amplified using NGA8, NGA139, NGA151, and NGA1107 respectively, at 6 mM $MgCl_2$. DNA in lanes 5-8 amplified using the same primer set at a 4 mM $MgCl_2$. DNA in lane 9 was amplified using NGA8 at 2 mM $MgCl_2$.

Microsatellite Instability Data Analysis

Although the data obtained from the screening of the 32 dominant-negative transgenic plants is preliminary, it is nonetheless substantial and shows very appreciable microsatellite instability (Table 1). In comparison to wild-type plants, a considerably larger frequency of repeat length-shifted alleles in endogenous microsatellite loci was detected in dominant-negative plants (Figure 10a). The frequency of shifted alleles was calculated by dividing numbers of unique length shifts by the total number of alleles tested in each group. There were no microsatellite shifts detected in any of the NGA6 or NGA172 loci.

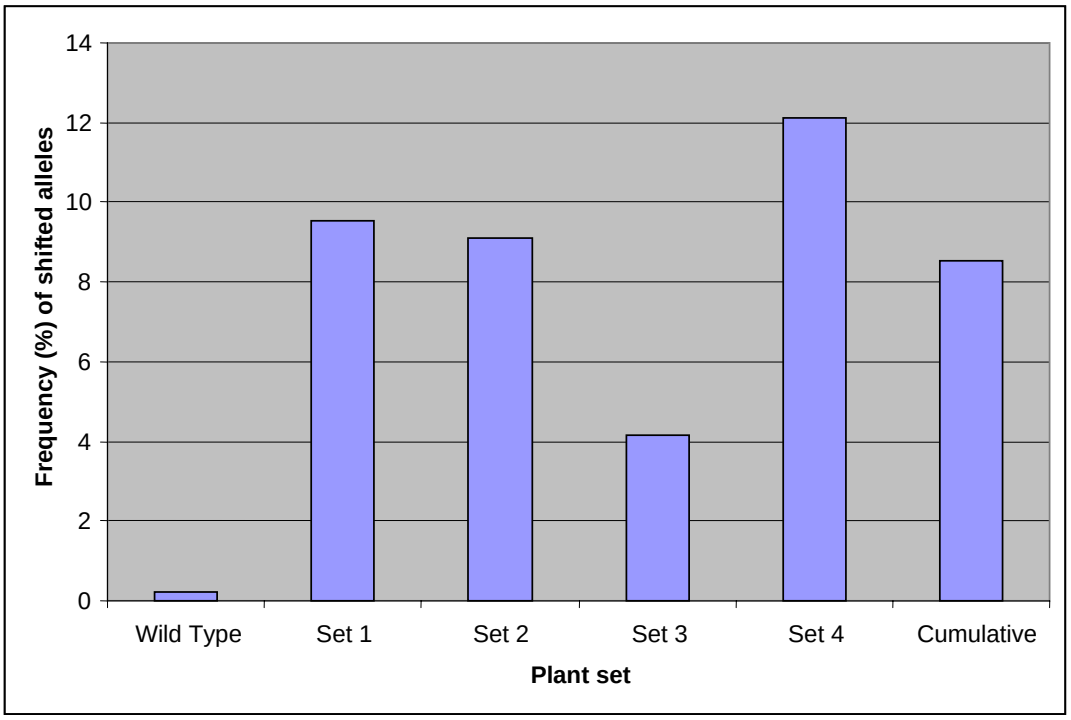
	NGA6	NGA8	NGA139	NGA151	NGA172	NGA1107	Total
Plant Set 1	0/16	4/10	-	-	0/16	-	4/42
	0	0.4	-	-	0	-	0.0952381
Plant Set 2	0/16	1/6	-	2/12	0/16	3/16	6/66
	0	0.1666667	-	0.1666667	0	0.1875	0.0909091
Plant Set 3	0/16	0/12	-	2/12	0/16	1/16	3/72
	0	0	-	0.1666667	0	0.0625	0.0416667
Plant Set 4	0/16	3/14	-	2/6	0/16	3/14	8/66
	0	0.2142857	-	0.3333333	0	0.2142857	0.1212121
Cumulative							21/246
							0.0853659
Wild Type							2/1000
							0.002

Table 1: Effects of dominant-negative AtMSH2 on stability of nucleotide-repeat sequence (microsatellite) alleles. Displayed are the frequencies of unique and total repeat-length shifts at indicated loci in progeny of indicated plants. Frequency of shifted alleles was calculated by dividing numbers of unique length shifts by the total number of alleles tested in each group. Plant Set 1 and Plant Set 2 represent seedlings from two progeny of a single Col-0 transformant. Plant Set 3 and Plant Set 4 represent seedlings from the progeny of two separate RDR6-15 transformants.

Data results obtained for the NGA139 locus were not decipherable and must be repeated. Results for the NGA151 and NGA1107 loci of plant set one were not included in the analysis because the data suggests the microsatellite shifts, identical in every plant in that set, are due to segregation of parental alleles, not microsatellite instability.

Interestingly, the total frequency of shifted alleles in Plant Set 1 and Plant Set 2 correspond very closely (see Table 1 and Figure 10a). This similarity in shifted alleles is probably due to the fact that Plant Set 1 and Plant Set 2 represent seedlings from two progeny of a single Col-0 transformant. With regards to microsatellite loci, the percentage in frequency of shifted alleles was 19% in the NGA8 locus, 16% in the NGA151 locus, and 11% in the NGA1107 locus (Figure 10b).

(a)



(b)

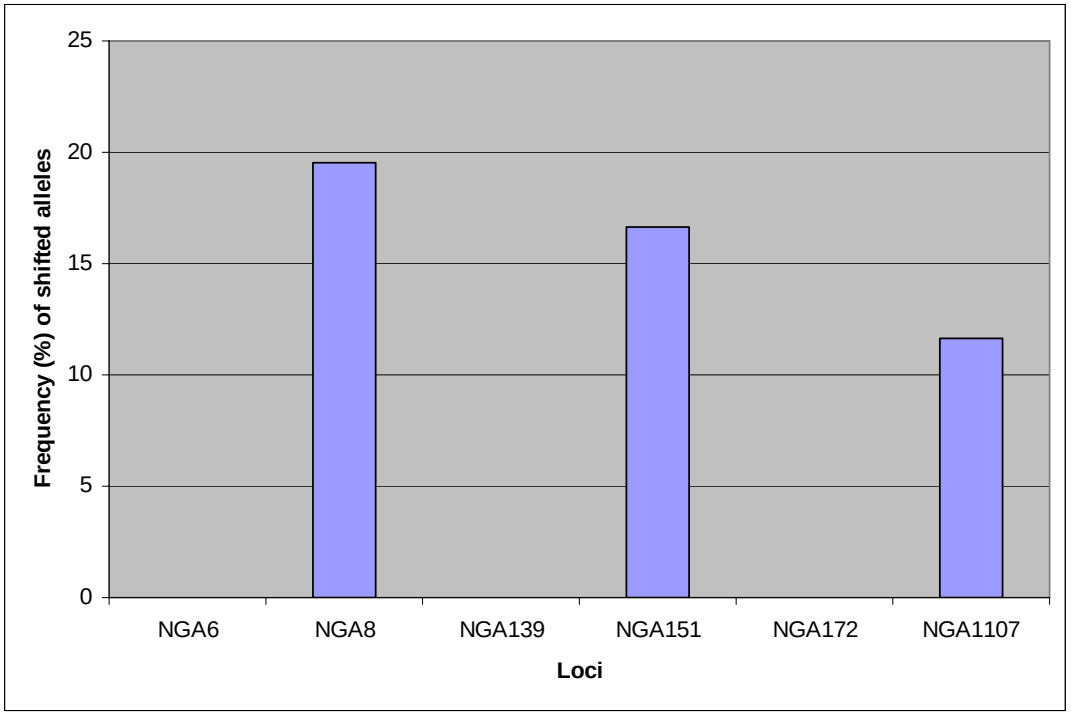


Figure 10: Frequency (%) of repeat length-shifted alleles in endogenous microsatellite loci. Frequencies were calculated by dividing numbers of unique length shifts by the total number of alleles tested in each group. (a) Data represent the average sum in each set of plants for all six loci analyzed. (b) Data represents the average sum of the number of shifts detected in each locus across all plant sets.

Discussion

Plants differ from animals in that they lack reserved germ lines - gametes arise from somatic cells which have divided numerous times before developing gametophytes. It has been suggested that the gametophytic stage offers a chance to expose recessive deleterious alleles – a phenomenon termed "haplosufficiency quality checking" (Wallabot and Evans, 2003). Previous At*MSH2* knock-out studies in *Arabidopsis thaliana* argue MSH2 mediated plant genomic maintenance mechanisms are equally important as haplosufficiency quality checking in maintaining genomic integrity (Hoffman *et al*, 2004). There have been no studies documenting the effect of a dominant negative MSH2 protein in *Arabidopsis thaliana*, although studies in *Saccharomyces cerevisiae* have shown observable dominant-negative phenotypes induced after deliberate over-expression of *MSH2* alleles bearing mutations in the ATPase domain and the helix-turn-helix domain of MSH2 protein.

The results presented here are encouraging in that they undoubtedly present strong evidence for microsatellite instability in the dominant-negative AtMSH2 plants. There are many aspects of this project that must be addressed before completion. Transformation of the DN-1 construct must continue until a complete set of Col-0, RDR6-4, and RDR6-15 transgenic plants is obtained. Appropriate microsatellite primers specific to the Lansberg ecotype are currently being pursued in order to begin analysis of RDR6-4 transformants. Microsatellite instability data must be collected from a larger sampling of plants to increase the statistical significance of the results. It must be determined why NGA6 and NGA172 loci contained no shifted alleles in all dominant-

negative plants. Could this possibly have been due to a procedural or technical error in data analysis, and if not, than why is it only these loci that display a marked lack of microsatellite instability?

Another area of great importance is the verification of dominant-negative AtMSH2 protein expression in the transgenic plants. The expression levels of the dominant-negative AtMSH2 protein must be investigated to ensure the expected robust expression of the protein. Previously, protein extracts from transgenic plants were isolated using a P-PER plant protein extraction kit (Pierce Biotechnology). Unfortunately, immunoblotting of the protein extracts using anti-c-MYC antibodies yielded no results. A colleague in our laboratory making use of the same binary vector and c-MYC tag has determined that immunoblotting using protein samples isolated from P-PER plant protein extracts does not yield quantifiable results. In comparison, a separate plant protein extraction protocol has yielded successful immuno-detection of an ACV-5 epitope fusion protein, but limited immuno-detection of c-MYC protein fusions.

These results indicate immuno-detection of c-MYC AtMSH2 protein fusion from plant extracts may not be possible. If immuno-detection of c-MYC AtMSH2 protein fails, RNA-detection protocols such as northern blotting or RT-PCR will be employed to verify the production and determine the amount of full length dominant-negative At*MSH2* mRNA being produced in plants.

It is conceivable that these MMR-defective plants may offer several advantages towards the basic study of plant genetics and development. The plants may have the potential to generate a wide variety of mutations which would accumulate gradually during propagation, giving rise to a variety of mutant progeny. When a desired mutant is

obtained, a plant line can be back-crossed with a wild-type plant and screened for the eradication of dominant-negative At*MSH2* alleles.

Dominant-negative AtMSH2 plants could also provide a model for long-term mutational loading in plant populations under sustained genotoxic stress. These plants offer a significant advantage over induced mutagenesis by heavy treatment with chemical or physical mutagens, because early-appearing mutations in the dominant-negative plants are probably not accompanied by as many confounding or potentially lethal mutations, as is often the case when treating with chemical or physical mutagens. There is prospect that because dominant-negatives mutants may be obtained without innumerable deleterious mutations, they might be an especially important tool in the identification of multiple locus traits. Importantly, every Arabidopsis plant produces thousands of seeds which are stable over several years at room temperature, ensuring that this analysis of mutational accumulation can be repeated and studied at will.

Bibliography

Adamson, A.W., Beardsley, D.I., Kim, W., Gao, Y., Baskaran, R., and Brown, K.D. 2005. Methylator-induced, Mismatch Repair-dependent G2 Arrest Is Activated through Chk1 and Chk2. *Molecular Biology of the Cell*. 16: 1513–1526

Alani, E., Sokolsky, T., Studamire, B., Miret, J.J., and Lahue, R.S. 1997. Genetic and Biochemical Analysis of Msh2p-Msh6p: Role of ATP Hydrolysis and Msh2p-Msh6p Subunit Interactions in Mismatch Base Pair Recognition*. Molecular and Cellular Biology*. 17[5]: 2436-2447

Cejka, P., Stojic, L., Mojas, N., Russell, A.M., Heinimann, K., Cannavo, E., di Pietro, M., Marra, G., and Jirlcny, J. 2003. Methylation-induced G2/M arrest requires a full complement of the mismatch repair protein hMLH1. *The EMBO Journal*. 22 [9]: 2245-2254

Charames ,G.S., and Bapat, B. 2003. Genomic Instability and Cancer. *Current Molecular Medicine.* 3: 589-596.

Culligan, K.M., and Hays, J.B. 2000. *Arabidopsis thaliana* MutS-homolog proteins—AtMSH2, AtMSH3, AtMSH6, and a novel AtMSH7 protein—form three distinct heterodimers with different specificities for mismatched DNA. *Plant Cell.* 12**:** 991–1002.

Culligan, K.M, Meyer-Gauen, G, Lyons-Weiler, J, and Hays, J.B. 2000.Evolutionary origin, diversification, and specialization of Eukaryotic Mut S homolog mismatch repair proteins. *Nucleic Acid Research*. 28: 463-471

Depeiges, A., Farget, M., Degroote, F., and 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

Hoffman, P.D., Leonard, J.M., Lindenberg, G.E., Bollman, S.R., Hays, J.B. 2004. Rapid accumulation of mutations during seed-to-seed propagation of mismatch-repair-defective Arabidopsis. *Genes & Development*. 18:2676-2685

Holland, J.F. 2003. *Cancer Medicine 6*. BC Decker: Hamilton, Ontario

Hsieh, P. 2001. Molecular mechanisms of DNA mismatch repair. *DNA Repair.* 486: 71–87.

Leonard, J.M., Bollman, S.R., and Hays, J.B. 2003.Reduction of stability of Arabidopsis genomic and transgenic DNA-Repeat Sequences. *Plant Physiology.* 133[1]: 328–338.

Lodish H, Berk A, Matsudaira P, Kaiser CA, Krieger M, Scott MP, Zipursky SL, and Darnell J. 2004. *Molecular Biology of the Cell*. Pg 963. WH Freeman: New York, NY. 5th ed.

Luo, Y., Lin, F., and Lin, W. 2004. ATM-Mediated Stabilization of hMutL DNA Mismatch Repair Proteins Augments p53 Activation during DNA Damage. *Molecular and Cellular Biology*. 24 [14]: 6430–6444

Peragine, A., Manabu, Y., Wu, G., Albrecht, H.L., and Poethig, R.S. 2004. SGS3 and SGS2/SDE1/RDR6/ are required for juvenile development and the production of *trans*-acting siRNAs in *Arabidopsis*. *Genes & Development*. 18:2368-2379

Studamire, B., Quach, T., and Alani, A. 1998. ***Saccharomyces cerevisiae*** Msh2p and Msh6p ATPase Activities Are Both Required during Mismatch Repair. Molecular and Cellular Biology. 18[12]:7590-7601

Voet, J.G., and Voet, D. 2004. *Biochemistry*. John Wiley & Sons: New Jersey. 3rd ed.

Walbot, V. and Evans, M.M.S. 2003. Unique features of the plant life cycle and their consequences. Nat. Rev. Genet. 4: 369-379.

Xie, Z., Johansen, L.K., Gustafson, A.M., Kasschau, K.D., Lessis, A.D., Zillberman, D., Jacobsen, S.E., and Carrington, J.C. 2004. Genetic and Fuctional Diversification of Small RNA Pathways in Plants. *PLOS Biology*. 2:0642-0652