AN ABSTRACT OF THE DISSERTATION OF

Stephanie R. Bollmann for the degree of Doctor of Philosophy in Genetics presented on May 21, 2008.

Title: Reversion Reporters in *Arabidopsis thaliana* to Detect all Six Base Substitution Pathways.

Abstract approved:

_____

John B. Hays

Maintaining genome integrity is essential for an organism, as mutation accumulation can lead to cancer, reduced fitness, and heritable diseases in offspring. Therefore the study of mutations, how they are induced, and how they are prevented is vital. Biomonitor systems are useful for understanding the relevant biological effects of a given mutagen, and depending on the system, can even provide information on specific molecular changes induced by the mutagen. Plants are ideal biomonitors, as a sedentary lifestyle allows measurement of mutagens in air, soil, and water, even at low doses. We constructed mutation reporters in *Arabidopsis* designed to restore β-glucuronidase (GUS) activity through one of six base substitution reversions. All six mutant constructs contained inactivating base substitutions in the same codon, which minimized sequence context effects. An AcV5 epitope tag sequence was fused to the 3' end of GUS to allow detection of the inactive protein and selection of sublines

based on levels of GUS protein expression. Initial characterization of these reversion reporters with or without UV-C treatment and exposure to heavy metal ions (Cd^{2+} and Zn^{2+}) supports the ability of the lines to measure different mutations. UV-C radiation induced T → C and C → T reversions, as well as T → G and T → A to a lesser extent. Heavy-metal-ion-mutation induction was inconsistent, showing variable increases in G → T and G → C, and no induction of T → C reversion. Of key interest is the G → T reporter line, as this is the first such reporter in higher eukaryotes. This line showed a large increase in spontaneous mutation as well as potential germinal mutations when compared to the other reporter lines, most likely due to endogenous oxidative damage (i.e. 8-oxoguanine). Further experiments to test the reporter lines with various mutagens as well as ways to improve the ability of the reporters to detect mutation are discussed.

©Copyright by Stephanie R. Bollmann
May 21, 2008
All Rights Reserved

Reversion Reporters in *Arabidopsis thaliana* to Detect all Six Base Substitution Pathways

by
Stephanie R. Bollmann

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented May 21, 2008

Commencement June 2009

Doctor of Philosophy dissertation of Stephanie R. Bollmann presented on May 21, 2008.

APPROVED:

_____

Major Professor, representing Genetics

_____

Director of the Genetics Program

_____

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

_____

Stephanie R. Bollmann, Author

ACKNOWLEDGEMENTS

I would like to express sincere appreciation to Dr. John Hays for being a wonderful mentor and encouraging my creativity, to Pete Hoffman for being a great resource and sounding board for ideas, and to my other colleagues in the lab, thank you for your encouragement and helpful suggestions these past few years. Finally, I would like to thank my graduate committee for being so helpful throughout my doctoral program of study.

TABLE OF CONTENTS

	Page
Specific goals and objectives….	1
Introduction……	3
Mutation and genome maintenance.....	3
Types of mutation.....	3
Sources of mutation.....	4
DNA repair and tolerance systems…….....	11
Mutation reporters……...	18
Bacterial reporters…….....	18
Plant reporters…….	20
Transgene reporters.....	24
Reporter genes.....	24
The Kovalchuk/Hohn reporters.....	25
The Depicker C → T reporters.....	32
Construction of a set of point mutation reporters.....	34
Chimeric GUS…….	34
Codon selection for mutation constructs…….……...	35
Transformation considerations……..…….	35
The Gelvin superpromoter system.....	38
Epitope tags.....	42
Measurement of microsatellite instability……..	44
Materials and Methods……	51
Results…….....	65
Testing and production of mutation constructs.....	65
Construction and testing of codon-11 amino acid substitutions…..…….	65
Construction of mutation reporters.....	68
Characterization of plant lines.....	76
Isolation of plant protein extracts for western blots……	76
Isolation of plant mutation-reporter lines.....	83
Expression patterns of wild-type constructs ……...	87
Testing of mutant plant lines with or without mutagens..……	90
Kovalchuk/Hohn point-mutation reporters……	90
Analysis of mutation events…….	95
Spontaneous mutation…….	98
UV mutagenesis……...	104
Mutagenesis by heavy metal ions...…….	111
Photoperiodicity effects on mutagenesis...……..	113

TABLE OF CONTENTS (Continued)

Page

Variation in mutation with generation and subline...……... 118
Problematic reporter lines……. 121
The M3 (G → C) reporter line……. 121
The M5 (A → T) reporter line……. 123
Analyses of mutation reporters……. 124
Sequence confirmation and southern blotting of reporter lines……... 124
Isolation of tissue for comparison of protein levels……. 124

Discussion and Conclusions……... 127
Limitations of previous transgenic mutation reporters……. 127
Initial observations of properties of reporter lines…… 129
Reversion frequency analyses…….. 131
Future experiments……... 144

Bibliography……... 149

Appendix……. 155

LIST OF FIGURES

Figure	Page
1. X-Gluc cleavage …..…….	26
2. Maps of the T-DNA binary vectors containing the Gelvin superpromoter ..	39
3. Frequency (%) of repeat length-shifted alleles in endogenous microsatellite loci ……...……	47
4. Microsatellite analysis of G5 *Atmsh2-1* plants ……..	49
5. Representative microsatellite genotypes for all *Atmsh2-1* (*AtMSH2*::*TDNA*) lines …...……...	50
6. Activity of mutant GUS proteins synthesized by *in vitro* transcription/ translation ……..	67
7. Maps of the T-DNA binary vectors containing the mutant GUS constructs	73
8. Strategy for sequencing binary-vector constructs ..……..	74
9. Western-blot signals for AcV5-GUS$^+$ protein diluted in plant protein extracts …….……..	78
10. Analysis of AcV5-GUS$^+$ expression by western blotting and histochemical staining …….……..	80
11. Comparison of staining of total protein with actin antibody and Ponceau S …….…….…...	81
12. Standard curve for Ponceau S staining ……..	84
13. Western-blot analysis of GUS expression in plants ……..	85
14. Staining of wild-type GUS protein expressed during development……...	89
15. Examples of GUS staining patterns in wild-type-construct lines ...…..…..	91
16. UV mutagenesis in M1 (TT → TG) sublines ……..……..	94
17. Staining in cells and tissue containing GUS reversion ..……	97

LIST OF FIGURES (Continued)

Figure		Page
18.	Comparison of data to Poisson distributions …….	101
19.	Spontaneous mutation frequencies ……	103
20.	Spontaneous and UVC-induced mutation …….…….	105
21.	Staining of apparent line M4 (G → T) germinal-revertant plants .……	107
22.	Effect of UV treatment on plant growth ..……..	109
23.	UV-specific mutations ……..…….	112
24.	Effect of photoperiodicity and UV-C treatment on plant growth …….	115
25.	Comparison of spontaneous mutation frequencies in successive generations……..…….....	119
26.	Comparison of UV-specific mutation induction between successive generations ……..……..….	120
27.	Protein expression in different plant tissues ….…….	126
28.	Growth of plants on cadmium-containing plates ……...	141
29.	GUS protein expression in successive generations……....……	148

LIST OF TABLES

Table		Page
1.	Frequency of spontaneous and induced reversions in the Kovalchuk/Hohn point-mutation reporters……...……...	30
2.	Determination of activity of mutant GUS proteins synthesized *in vitro*	69
3.	Mutation Reporter Constructs ……..	70
4.	Reversion of Kovalchuk/Hohn mutation-reporter lines ……...	93
5A.	Spontaneous and UV-induced mutation frequencies…….	99
5B.	Spontaneous and cadmium-specific mutation frequencies……	100
6.	Comparison of mutation-frequency calculations……..	102

LIST OF ABBREVIATIONS

2-OH-A	2-Hydroxyadenine
[6-4]s	Pyrimidine-pyrimidone 6-4 photoproducts
8-oxodGTP	7,8-Dihydro-8-oxodGTP
8-oxoG	7,8-Dihydro-8-oxoguanine
AmasPmas	Mannopine synthase promoter/activator region
AMP	Ampicillin
Aocs	Octopine synthase upstream activating sequence
AP	Apurinic/apyrimidinic
AtMMH	*Arabidopsis thaliana* homolog of MutM
AtOGG1	*Arabidopsis thaliana* homolog of 8-oxoguanine glycosylase
BER	Base excision repair
CaMV	Cauliflower mosaic virus
CoYMV	Commelina yellow mottle virus
CPDs	Cyclobutane pyrimidine dimers
DSB	Double-strand break
DSBR	Double-strand-break repair
EMS	Ethyl methane sulfonate
GFP	Green fluorescent protein
GGR	Global/general genomic repair
GUS	β-glucuronidase
HR	Homologous recombination

LIST OF ABBREVIATIONS (Continued)

IR	Ionizing radiation
KAN	Kanamycin
MCN	Micronuclei
MMR	Mismatch repair
MMS	Methyl methane sulfonate
MSH	MutS homolog
MSI	Microsatellite instability
MUG	4-Methylumbelliferyl β-D-glucuronide
NER	Nucleotide excision repair
NHEJ	Nonhomologous end joining
NTR	Nontranslated region
PCR	Polymerase chain reaction
PTGS	Post-transcriptional gene silencing
RISCs	RNA induced silencing complexes
RNAi	RNA interference
ROS	Reactive oxygen species
SDSA	Synthesis-dependent strand annealing
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
siRNAs	Small interfering RNAs
TCR	Transcription-coupled repair
TEV	Tobacco etch virus

LIST OF ABBREVIATIONS (Continued)

TGS	Transcriptional gene silencing
TLS	Translesion synthesis
UV	Ultraviolet
X-gal	5-Bromo-4-chloro-3-indolyl-β-D-galactoside
X-gluc	5-Bromo-4-chloro-3-indolyl-β-D-glucuronic acid
XP	Xeroderma pigmentosum

DEDICATION

I dedicate this work to my husband and my parents, who have supported and encouraged me for many years, and to my son who is helping me rediscover the world.

Reversion Reporters in *Arabidopsis thaliana* to Detect all Six Base Substitution Pathways

Specific goals and objectives

Mutation reporter systems are essential for studies of mutation induction *in vivo*, and plant reporter systems have the advantage of being able to detect damaging agents from air, soil, and water. In addition, as plant cells are fixed in place, mutations are observed as a spot or clonal sector, thus enabling quantitative measurement of every mutation induced. This event isolation is in sharp contrast with bacterial mutation assays, in which mutation reporter strains are grown in liquid culture before selective plating. This leads to the inability to determine if the number of positive-scoring colonies on a plate is the result of multiple independent, late mutations or one early mutation multiplied by cell division.

Previous plant mutation assays have been convenient for measurement of gross chromosomal changes, but do not provide direct indicators of the specific mutations induced. More recent reversion-reporter systems have begun to address this specific issue, but have not covered the entire point mutation spectrum. There is a need, therefore, for a system within which reversions corresponding to all six base substitution pathways can be analyzed. These defined mutations should all lie within the same sequence context, to enable better comparison among lines. However, selection of an ideal codon is required to ensure all six mutant constructs produce inactive protein. High transgene expression is also required, for ease of detection of reversion events. The promoter should be strong and ideally able to be expressed throughout the plant. Resistance to transgene silencing is also essential to promote

equivalent expression across multiple generations, allowing production of a stable mutation reporter line. The transgene itself should also be designed for optimal translation efficiency. Finally, addition of an epitope tag would aid selection of highly expressing lines for analysis.

A system which is able to measure each of the six transitions and transversions offers the advantage of measuring different mutational responses depending on the type of mutagen. For example, UV-C damage should preferentially yield T → C and C → T transitions when the target base is located in the 3' side of a potential dimer photoproduct. Another use for this system is measurement of certain types of mutation which have proved difficult to assay accurately. Bases are easily oxidized during DNA purification and some oxidation products are biochemically unstable, making it difficult to accurately measure spontaneous base oxidation rates, persistence, and repair [1]. It would therefore be very useful to have a reporter system enabling measurement of oxidative damage *in vivo*, without the need to manipulate DNA directly.

This dissertation describes the construction of a reporter system designed to fulfill the previously stated objectives. Initial experiments with UV radiation and heavy-metal-ion treatment were performed to test the utility of the lines for reporting mutation. Background information provided includes a review of mutation and genome maintenance, previous mutation reporters, and topics relevant to the construction of the reporter lines. A summary of my preliminary graduate research is also provided.

Introduction

Mutation and genome maintenance

DNA is continually exposed to different potential sources of mutation, from replication errors and other endogenous sources to environmental pollutants in water, soil, and air. UV radiation from sunlight and ionizing radiation are also threats to genome integrity. There are protective measures available, either by physically avoiding polluted areas or through removal of damage by DNA repair, but when the level of damage is too high, mutations can accumulate [2]. Mutations can lead to increased genetic diversity, allowing for adaptation to a changing environment. However, they can also lead to embryonic lethality, malformation, cancer, and hereditary diseases, depending on the affected gene and the severity of the mutation [3]. Mutations have also been implicated in cellular aging [1]. It is therefore important to obtain a thorough understanding of how mutagenic lesions occur and how they are repaired.

Types of mutation

There are several different types of genetic mutation. Base substitutions are either transitions (A:T to G:C or G:C to A:T) or transversions (A:T to T:A or C:G, G:C to C:G or T:A). When base substitutions occur within a gene, they are categorized as silent mutations (resulting in a different codon that codes for the same amino acid), missense mutations (code for different amino acids), or nonsense mutations (premature stop codons leading to early protein truncation). The severity of

the mutation depends on the category of mutation, the codon that is mutated, and how essential the gene product is to the cell or organism. For example, silent mutations are generally not harmful whereas nonsense mutations are predominantly harmful. Base substitutions affect a single base, whereas other types of mutation can affect variable numbers of bases. Insertions consist of (an) extra base(s) added within a given sequence, and deletions consist of the loss of (a) base(s) from the DNA. Small insertions or deletions can alter the reading frame of a gene, leading to altered amino acid coding and premature stop codons downstream of the mutation. Large insertions or deletions can create null mutants through the disruption of genes. Duplication of a chromosome section can occur through unequal crossover during meiosis, where DNA is deleted from one chromosome and added to its homolog. Inversions involve the 180º rotation of a fragment of DNA. Inversion and duplication can lead to aneuploidy and reduced gamete formation. Reduced gamete formation also results from reciprocal translocations, which occur when two nonhomologous chromosomes exchange segments [4].

Sources of mutation

Endogenous mutation

The replicative polymerases accurately choose which base to insert across from the templated base, and their proofreading activity can correct the majority of misincorporated bases. Yet replication errors occur on average once every 10^6-10^7 bases, which in *Arabidopsis* would lead to 10-100 mutations per cell per cell division. Additionally, water molecules can react with DNA, inducing spontaneous alteration of

DNA bases. Nonenzymatic cleavage of glycosyl bonds leads to AP (apurinic/apyrimidinic) sites [5, 6]. AP sites can block DNA replication and transcription, and if left unrepaired can lead to cell death. Yet AP sites are recognized and repaired rapidly, and therefore are not a major source of mutation. The more significant hydrolytic damage is deamination of cytosine to uracil and 5-methylcytosine to thymine. Uracil is excised by uracil glycosylase, but CpG and CpNpG methylation is common in plants, in some measurements as much as 10% of the genome. This is problematic, as T:G mispairs resulting from 5-methylcytosine deamination, when left unrepaired before the next round of replication, lead to C → T mutation [1]. Aging of seeds can also lead to chromosome damage, which is dependent on temperature, moisture, and oxygen levels, and can result in loss of seed viability. Environmental stresses, such as drought, temperature extremes, or increased UV-B due to stratospheric ozone depletion, can increase mutation through formation of reactive oxygen species (ROS) [5]. ROS are also induced naturally as byproducts of chloroplast and mitochondrial functions, such as photosynthesis and respiration. Cells have defenses against ROS, such as free radical scavengers, but these systems can be overwhelmed [1]. ROS damage DNA through oxidation or cleavage of bases and induction of single strand breaks, either directly or as a result of excision repair [5]. ROS can also react with DNA precursors in the dNTP pool. One example is 7,8-dihydro-8-oxo-dGTP (8-oxodGTP) which can be incorporated into DNA and induce T → G mutation [7]. The main ROS are hydroxyl radicals, superoxide, and nitric oxide, which can produce such adducts as 8-hydroxyguanine (pairs with adenine or cytosine)

and oxidized cytosine, which when deaminated leads to mutagenic uracil derivatives. The major oxidative damage is single strand breaks, as base stacking protects the bases from attack, but a nearby nick loosens the DNA structure and makes bases more accessible to damaging agents [1].

Oxidative damage induces several base substitution pathways. 7,8-dihydro-8-oxoguanine (8-oxoG) induces G → T mutations when template guanine bases are oxidized, and T → G mutations when 8-oxodGTP is inserted opposite template adenine. 2-hydroxyadenine (2-OH-A) induces predominantly A → T transversions, thymine glycol induces T → C transitions, and uracil glycol induces C → T transitions, to name a few. Although many different types of mutation are possible, some mutations are more prevalent. In studies of spontaneous mutation, human T lymphocytes show mainly C → T mutations, mouse embryonic fibroblasts show C → T and G → T mutations, and *E. coli* shows C → T and G → T mutations, with a smaller proportion of CC → TT tandem mutations. These spectra reflect both oxidative damage and base deamination. Oxidized dNTPs have also been implicated in the induction of small insertions and deletions (frameshifts) during replication [8].

Chemical mutagens

Organisms are routinely exposed to organic and inorganic compounds. Mutagenic inorganic compounds include heavy metal ions. Heavy metal ions can be categorized as essential and nontoxic (e.g. Zn^{2+}), essential and toxic at high concentrations (e.g. Cu^{2+}), and nonessential and toxic at high concentrations (e.g. Cd^{2+}). Heavy metal ion contamination of soil has increased with time, which can lead

to DNA damage and cancer [9]. Common sources of pollution include mining, metal production, fossil fuel combustion, phosphate fertilizers, and sewage sludge [10]. Even non-polluted soil contains 0.04 – 0.32 μM cadmium (moderate pollution is defined as 0.32 – 1 μM cadmium) [11], and accumulation in the edible portion of crop plants is the most common form of exposure for humans. Some heavy metal ions, such as Cd^{2+}, are taken up well by roots and can be spread to leaves and even fruit and seeds [10]. There are two proposed methods of heavy-metal genotoxicity: creation of reactive oxygen species or direct inhibition or interference with DNA repair and replication processes, with the latter specific to cadmium. Heavy metal ions can produce hydroxy radicals and other ROS through Fenton-type reactions [9]. The basic Fenton reaction is $M^{n+} + H_2O_2 \rightarrow M^{(n+1)+} + \bullet OH + OH^-$, where M is a transition metal ion [12]. Carcinogenic heavy metal ions are mostly non-mutagenic in bacteria, and cause various cellular damage in mammalian cells. Heavy-metal-ion-induced genotoxicity can be enhanced when combined with other DNA damaging agents [13].

Cadmium is the most commonly studied heavy metal. It is a potent human carcinogen, with mutagenicity occurring as a result of induction of ROS and inhibition of DNA repair. In addition, the biological half-life of Cd^{2+} is 30 years, making it a cumulative toxin [14]. Cd^{2+} cannot itself participate in the Fenton reaction, but generates free radicals indirectly, by replacing Cu^{2+} and Fe^{2+} in metal-binding proteins, freeing them to participate in Fenton reactions [12]. Cd^{2+} also induces ROS formation through inhibition of anti-oxidant enzymes and free radical scavengers [14]. Cd^{2+} may also disturb DNA repair, even when present at non-cytotoxic concentrations,

through competition with essential metal ions [13]. A study of the effects of heavy metal ions on DNA replication revealed Cd^{2+} reduction of DNA synthesis fidelity [15]. Cd^{2+} is able to displace Zn^{2+} in zinc finger DNA binding domains, affecting replicative polymerases and various excision repair pathway proteins [14] [16] [17]. These inhibitory and mutagenic properties of cadmium have mostly been observed with high, acute doses of Cd^{2+}, which can also induce gross chromosomal changes. Further study of cadmium at biologically relevant, chronic lower doses is essential for understanding cadmium-induced carcinogenicity. When human cells were exposed to low doses of Cd^{2+}, an initial increase in DNA damage was observed, followed by a decrease to background mutation levels after 24 hours. One theory suggests that chronic exposure to low doses of cadmium may lead to adaptation through increased expression of anti-oxidants [14].

Radiation

Organisms are exposed to different forms of radiation. Ultraviolet (UV) radiation induces many types of damage, including oxidative damage (pyrimidine hydrates) and cross-links (DNA to protein and protein to protein). The most prevalent damage is pyrimidine dimers, or photodimers, in which two tandem pyrimidines are covalently bound to each other, thus altering DNA structure [18]. The two major types of dimer induced are cyclobutane pyrimidine dimers (CPDs) and pyrimidine-pyrimidone 6-4 photoproducts ([6-4]s). UV-A (315-400 nm) is the most prevalent form of UV radiation, with the greatest amount of penetrance [1]. Exposure to UV-A results in a photosensitization reaction, with induction of oxidative damage, abasic

sites, strand breaks, and CPDs. UV-B (280-315 nm) additionally induces pyrimidine dimers, with one measure of dimer induction in plant tissue showing nine-fold higher levels of CPDs than [6-4]s. UV-B is harmful to plant growth, development, morphology, and metabolism. Pyrimidine dimers are cytotoxic and mutagenic, as they block both DNA and RNA polymerases [5]. Specialized DNA polymerases that are able to replicate past dimers are prone to misincorporation of bases across from the photodimer, particularly the 3' base.

UV-C (< 280 nm) is commonly used for UV-mutagenesis experiments. UV-C is more efficient than UV-B at dimer formation, but UV-C radiation is blocked by ozone [1]. In one study of UV-C-induced dimers in human cells, the prevalence of lesion induction was T[CPD]T > T[CPD]C > T[6-4]C >> T[6-4]T, C[CPD]T, C[CPD]C. Although T[CPD]T was more highly induced, the major mutation seen was C → T, occurring when adenine was inserted opposite the 3' cytosine of T[CPD]C or C[CPD]C [19]. In agreement with this lesion-induction measurement was a human cell study using high performance liquid chromatography and mass spectrometry to measure CPD and [6-4] induction in all four bipyrimidine sequences. The resulting dimer distribution was T[CPD]T (40-45%) > T[CPD]C, T[6-4]C (20-25%) > C[CPD]T (10%) > T[6-4]T, C[CPD]C, C[6-4]C, C[6-4]T [20]. Dimer-formation spectra vary with cell type, yet the overall trend remains the same [21]. However, the relative levels of induced dimers do not correlate with the observed mutation spectra. Douki and Cadet suggest several possibilities for this difference. TT is not a preferential context (hotspot) for mutation, potentially due to a lower

frequency of misinsertion across from the dimer relative to cytosine-containing dimers. TC is a mutation hotspot due to cytosine deamination, which also contributes to the high incidence of CC → TT despite the CC sequence context showing the lowest dimer induction. Finally, there are differences in repair efficiencies between the two types of photodimer [20]. [6-4]s are more mutagenic than CPDs, but [6-4]s are repaired more efficiently by global genomic nucleotide excision repair (discussed later, page 14); therefore the majority of mutations seen involve CPDs. UV-induced mutation is primarily responsible for the development of skin cancer, as photodimers can lead to mutations in key tumor-suppressor genes such as p53. The most prevalent mutations observed in these genes are TC → TT, CC → TT, and to a lesser extent TT → TC. C → T is seen more than T → C, again presumably due to deamination of cytosines in CPDs [22]. Dimer distribution is dependent on the localization of nucleosomes and DNA binding proteins, and thus the accessibility to DNA is a factor in dimer formation. This also affects UV-damage repair, which is influenced by transcriptional status and the chromatin environment [23].

Another damaging form of radiation is ionizing radiation (IR). Organisms have adapted to natural radiation levels by defensive systems such as DNA repair, which correct the double-strand breaks and modified bases induced by IR. Radioactively contaminated soils are particularly harmful as compared to an acute dose of IR since they provide chronic exposure to plants. This can be damaging to the plants, thus monitoring is important [2]. IR is harmful to cells, inducing single- and double-strand breaks and radiolysis of water, leading to formation of ROS [5]. IR

damage can result in point mutation, inversion, duplication, and translocation [1]. Usually organisms have low levels of exposure to this dangerous mutagen, including cosmic rays and medical x-rays for humans. However, isolated events of higher exposures have occurred [5]. The accident at the Chernobyl nuclear power plant was one such case, where radioactive fallout contaminated over 600 square kilometers of land. Plants are continuing to grow in the affected area. Although the plants have accrued chromosome aberrations, they show adaptation through higher resistance to radiomimetic agents and free-radical-producing agents [24].

DNA repair and tolerance systems

Direct repair

Plants tolerate a majority of UV through absorption by the waxy cuticle, cell walls, and flavonoids. The damage caused by UV radiation that is able to penetrate beyond the first layer of defense must be corrected by DNA repair systems [6]. Photoreactivation, or light repair, is accomplished by photolyases which use the energy of blue light photons (350-450 nm) to break dimers via a cyclic electron transfer reaction [5]. Photolyases employ two chromophores for activity: a flavin cofactor which acts as a transient electron donor to reverse the dimer cross-link and an antenna pigment to excite the electron donor [18]. Most organisms employ CPD photolyases (except placental mammals), and some also contain [6-4] photolyases (insects, plants, fish, amphibians, and reptiles). In *Arabidopsis*, photolyases are expressed in all tissue types, with seedlings expressing less CPD photolyase than mature leaves, and highest expression in floral tissue. [6-4] photolyase expression is

constitutive, whereas CPD photolyase expression is inducible by white light. Photolyases can repair a normal, low level of dimers, but higher levels of dimers require the additional function of excision repair (dark repair) [5]. In one *Arabidopsis* study, photolyase repair was quite rapid. Roughly 50% of both CPDs and [6-4]s were repaired after 24 hours, and 60% of those repairs occurred within the first two hours. In comparison, dark repair was less efficient for both dimers, and only showed 13% repair of CPDs and 23% repair of [6-4]s after 24 hours (most of these repairs occurred in the first two hours) [25]. Overall, the photoreactivation system is the major repair pathway expressed in plants, with expression in both proliferative and nonproliferative tissues, whereas excision repair is mainly limited to proliferative and meristematic tissue. Ozone depletion over the recent decades has led to an increase in UV exposure, especially UV-B, making the investigation of UV-induced damage and repair even more critical [6].

Indirect repair

Base excision repair (BER) involves the removal of damaged bases by DNA glycosylases that are specific for particular base adducts such as 8-oxoG. Bacteria remove 8-oxoG through several pathways, including MutT which hydrolyzes oxidized 8-oxodGTP to prevent incorporation, and MutY which excises an adenine when incorporated opposite 8-oxoG or 2-hydroxyadenine when opposite guanine [7]. 8-oxoG can be excised from DNA in *Arabidopsis* by either the 8-oxoG DNA glycosylase/AP lyase (AtOGG1) or the *Arabidopsis* MutM homolog (AtMMH). DNA glycosylases cleave the glycosidic bonds of damaged bases, resulting in an AP site.

An AP endonuclease (either separate or within the glycosylase) cuts 3' to the AP site. The AP site then undergoes single-strand-break repair, either through a short-patch (1 nucleotide) or long-patch (2-10 nucleotides) repair pathway [5]. Short-patch repair involves removal of the 5'dRP residue followed by single base insertion and ligation. Long-patch repair involves synthesis of a long chain which leaves a single-strand flap to be cleaved, followed by strand ligation [6].

Nucleotide excision repair (NER) detects conformational changes in the DNA strand, and thus can recognize a wider range of damage substrates. There are two types of NER: global/general genomic repair (GGR) and transcription-coupled repair (TCR). In GGR, repair can occur anywhere in the genome at any time. In TCR, repair is limited to actively transcribed DNA strands, and damage is recognized as the RNA polymerase stalls at the site of damage. Beyond damage recognition, both repair pathways converge mechanistically. These well-conserved repair pathways function by removing a damage-containing oligonucleotide 24-32 bases long, followed by synthesis and ligation [5]. Defects in NER in humans lead to the autosomal recessive disease xeroderma pigmentosum (XP). XP patients have extreme UV hypersensitivity and various clinical and genetic changes, depending on the affected gene. The UV hypersensitivity is manifested as high incidence of skin cancer in sun-exposed tissues, on average by the age of eight. UV-B appears to be the primary inducer of carcinogenesis, and UV signature mutations (C → T and CC → TT) are found in mutated genes, such as *ras* oncogenes and *p53* and *PTCH* tumor suppressor genes [26].

NER repairs different bipyrimidine dimers with varying efficiency. The general trend for repair is T[6-4]T, T[6-4]C (larger distortion, therefore easily recognized) >> C[CPD]T > C[CPD]C > T[CPD]C > T[CPD]T. As a result, one study using human skin cells showed roughly 50% of CPDs remain 24 hours after UV treatment (~64% T[CPD]T, ~45% T[CPD]C), whereas virtually all of the [6-4]s are repaired. After 48 hours, 50% of T[CPD]T and 25% of T[CPD]C remained. The variability in repair efficiencies contributes to the observed UV-induced-mutation spectra. Few or no mutations are observed at TT contexts, due to the efficient repair of T[6-4]T as well as the high propensity of adenine incorporation across from a TT dimer during synthesis. Few CT mutations are observed, due to efficient repair, potentially before the cytosine has a chance to deaminate. A large number of mutations at TC are observed, due to less efficient repair and thus more time available for cytosine deamination to a more mutagenic lesion. Finally, a large number of CC → TT mutations occur (20% in wild-type cells and 80% in XP cells), due to cytosine deamination despite fairly efficient repair [21].

The major repair pathway for replication errors is the highly conserved mismatch repair (MMR) system, which brings the spontaneous mispair error rate from 10^{-6}-10^{-7} to 10^{-9}-10^{-10}(which would correspond to 0.01-0.1 mutations per cell per cell division in *Arabidopsis*) [5]. MMR also recognizes oxidized bases when within a mismatch [7], and MMR inhibits homologous recombination between mismatch-containing substrates, minimizing illegitimate recombination. In bacteria, MutS is responsible for lesion recognition and MutL binds MutS and activates MutH, which is

responsible for nicking the strand containing the incorrect base. The nicked strand is excised beyond the mismatch, allowing for resynthesis and insertion of the correct sequence of bases. Plants encode MutS homologs MSH1, 2, 3, 4, 5, 6, and 7, with MSH7 being unique to plants. The MSH proteins form a heterodimer which recognizes and binds a mismatch, with different heterodimer pairs having varied specificities. The main heterodimers are MSH2/3, which recognizes extrahelical loopouts, and MSH2/6 and MSH2/7, which recognize small loopouts and mismatched bases with different specificities. MSH1 is involved in mitochondrial repair, and MSH4/5 have unknown functions during meiosis, potentially early in meiotic recombination. The MutL homolog, the MLH1/PMS2 heterodimer, discriminates damaged versus template strand and leads to excision by an exonuclease of the nascent damage-containing DNA past the mismatched base. DNA polymerase and DNA ligase I resynthesize the excised strand and ligate it to the nascent strand [5].

Defects in MMR lead to increased spontaneous mutation, in particular insertions or deletions of repeat units in repetitive DNA sequences (microsatellite instability). A study in yeast revealed inhibition of MMR as a pathway for Cd^{2+}-induced mutagenesis. Chronic exposure to Cd^{2+} (up to 5 μM) led to an increase in microsatellite instability and base substitutions, at a rate from 20-50% of an *msh2* deficient strain. Both mutational assays and biochemical assays indicated a concentration-dependent inhibition of MMR. This result is important, since even small decreases in MMR activity can be a risk factor for developing cancer [17].

Homologous recombination (HR) and nonhomologous end joining (NHEJ) are pathways for the indirect repair of double-strand breaks (DSB). There are three models of HR repair. In DNA double-strand break repair (DSBR), which is important during meiotic recombination, broken DNA ends are resected to form 3' tails which are used to invade homologous chromosomes for DNA synthesis templates. The looped out strand of the homolog then anneals to the other exposed tail. DNA synthesis continues until ligation occurs between the invading strand and the other end of the DSB, which results in the formation of a joint molecule with two Holliday junctions [5]. These junctions can be resolved by nuclease incision [6] with or without crossover formation. Synthesis-dependent strand annealing (SDSA) occurs mainly in somatic cells, and can use sister chromatids, homologous chromosomes, or ectopic regions of homology within the genome as recombination substrates. As in DSBR, a 3' resected tail invades the recombination substrate and primes DNA synthesis, but in SDSA the invading strand reanneals with the other side of the DSB without formation of a joint molecule, thereby decreasing the chance for formation of crossovers. Single-strand annealing occurs between tandemly repeated sequences. Resection of the DSB exposes homologous sequences which anneal to each other, leading to the loss of the intervening sequence. NHEJ is the main DSB repair pathway in higher eukaryotes. The DNA ends of DSB are processed into substrates for alignment at microhomologies, followed by trimming of flaps and ligation. This repair pathway generally results in deletions [5].

Toleration of damage

Plants tolerate low levels of persistent CPDs through translesion synthesis (TLS) [5]. When replicative DNA polymerases δ or ε are blocked by damage, a translesion polymerase (DNA polymerase ζ, η, ι, κ, or Rev1) is loaded onto the DNA and bypasses the lesion. Different translesion polymerases have various fidelities. Some translesion polymerases are fairly accurate, such as Pol η which correctly inserts AA across from a TT photodimer, while others are more error-prone [6]. Pol η usually inserts adenine across from photodimers, making T[CPD]T relatively nonmutagenic, and dimers containing cytosine mutagenic. This group of specialized polymerases do not appear to be expressed in floral tissue, where point mutations introduced into gamete genomes would be more deleterious [1]. Plants rely on translesion polymerases for bypass of various types of DNA damage. For example, *Arabidopsis* plants deficient for polymerase ζ are sensitive to UV (inhibition of root elongation), γ-radiation, and the cross-linking reagent mitomycin C [6]. Another form of damage tolerance is recombinational 'repair' (i.e. template switching). When a lesion is unable to be repaired or bypassed, a homologous region can be used as a template for replication past the lesion, but the lesion itself remains unrepaired. Cell cycle checkpoint activation, another tolerance mechanism, allows more time for DNA repair or tolerance pathways to act [1].

Mutation reporters

Genome alteration can be measured directly or indirectly. Indirect mutation assays include forward mutation assays which measure mutation through inactivation of a target transgene product. An example of a target transgene is an integrated phage or plasmid within the genome of an organism, such as a mouse. These targets are excised, cloned into bacteria, positively selected for mutation and sequenced. Forward assays are able to show many different types of mutation, but only reveal inactivating mutations. As well, quantification of mutation events can be influenced by early mutation and cell division leading to multiple counts of one event (jackpotting). Direct measures of mutation often involve reverse mutation assays which measure mutation through a defined inactive to active reversion of a reporter gene. The active reporter gene product is either detected visually or selected for drug resistance. In multicellular organisms such as plants, which have immobile cells, a reversion mutation is seen as a spot or sector of active transgene expression, thus avoiding jackpotting. However, these types of assays are often labor-intensive, as analysis of large numbers of isolates is required for quantifiable data. Additionally, a specific line is required for each desired mutation pathway [27].

Bacterial reporters

The most common bacterial mutation reporter assay is the Ames test. In this assay, histidine-auxotrophic *Salmonella typhimurium* are exposed to a test compound, and ability of the compound to induce mutation is quantified by counting the number

of histidine^{+} revertants [4]. In a study by Levin and Ames, the *Salmonella* mutagenicity assay was modified to enable examination of specific transitions and transversions, thus allowing identification of which base substitution pathways are induced by a given mutagen [28].

An innovative reverse-mutation reporter system was developed by Cupples and Miller [29]. A set of *lacZ* (β-galactosidase) constructs which were inactivated with one of six base substitutions were introduced into *E. coli*. These reporter transgenes could each revert to wild-type activity through a defined transition or transversion pathway. Reversion was observed as blue staining of colonies when grown on 5-bromo-4-chloro-3-indolyl β-D-galactoside (X-gal) containing media. Only the wild-type glutamic acid residue (GAG/CTC), which was located within the active site, allowed galactosidase activity. Therefore any mutation other than the defined reversion would not result in recovery of visible activity. These different reversion lines were then treated with various known mutagens to confirm the ability of the reporters to respond appropriately to different types of DNA damage. Treatment with ethyl methane sulfonate (EMS), an alkylating agent, mainly induced G:C → A:T mutations, with low induction of G:C → T:A, T:A → C:G, T:A → G:C, and T:A → A:T. *N*-methyl-*N*'-nitro-*N*-nitrosoguanidine treatment also mainly induced G:C → A:T reversions, with less induction of T:A → C:G, T:A → G:C, and fewer T:A → A:T and C:G → G:C reversions. UV treatment induced various levels of reversion, with CT<u>T</u> → CT<u>C</u> > C<u>C</u>C → C<u>T</u>C > T:A → A:T > T<u>T</u> → T<u>G</u> > G:C → T:A. 5-azacytidine induced mostly G:C → G:C transversions, as well as G:C → T:A to a lesser extent. 2-

aminopurine induced C:G → T:A > T:A → C:G >> G:C → T:A reversions. The six reversion reporters were also introduced into several repair-deficient backgrounds. Deficiency in MutT, which is responsible for blocking replicative incorporation of 8-oxodGTP, led to increased T → G reversion. Deficiency in MutY, which removes an adenine when paired with 8-oxoG, induced G → T reversions. Finally, deficiency in MutH, which is a key MMR protein in *E. coli*, led to an increase in transitions, with more C → T than T → C. Overall, the six mutation reporters reacted to mutagen treatments as expected from previous studies, with the exception of UV-induced T → C and C → T (discussed later, page 137) [29].

Plant reporters

A biomonitor is an organism designed for detection of toxins or mutagens in the environment. Ideally, a biomonitor should be able to detect pollutants from air, water, and soil sources, with the bioavailability of the mutagenic compound and the duration of contact with the organism being critical factors [2]. Bacterial biomonitor assays include the Ames test, the alkaline single-gel electrophoresis assay, and the sister chromatid exchange assay. These tests work well, but do not respond to all potential mutagens, such as heavy metal ions [9]. Eukaryotic biomonitors include the somatic-eye-mutation test from Drosophila and animal systems such as transgenic mice, transgenic zebrafish (*lacI*), and land snails. These tests are more powerful, as the effect of a compound on a multicellular organism can be studied, but there are drawbacks. Not all sources of contamination can be tested. Zebrafish can only be used to assay water quality. The land snails cover a limited territory and are able to

respond to soil toxicity, but only to high concentrations [2]. Results from mammalian assays can effectively estimate the impact of a compound in humans, but have a history of being controversial and expensive [9]. Overall, laboratory test results can be used to estimate the toxicity of a compound, but laboratory tests cannot completely mimic exposure in a natural system. The conditions of a laboratory are not as variable as a natural environment, the duration of exposure is shorter, and there is a lack of interaction with other environmental factors [2].

Plants, on the other hand, make excellent biomonitors. They are immobile, thus they are constantly exposed to contaminants, and they can respond to compounds in air, soil, and water. Most of the established plant bioassays reveal gross chromosomal aberrations. These biomonitors include *Allium cepa*, *Tradescantia*, and *Vicia faba*, which have been used to sample air, water, and chemical toxicity [2]. While these bioassays do not provide DNA-level information [9], they are more sensitive to mutagens than other assays, such as the Ames test and the mouse microscreen test [2]. This is likely due to higher accumulation of mutagenic compounds than other organisms, especially in somatic tissues, as well as the sedentary lifestyle of plants that leads to inability to escape from sources of mutation [9]. The more recent plant biomonitors have included *Arabidopsis thaliana*, *Hordeum vulgare*, and *Zea mays*, and the detection capabilities have expanded to cover cytotoxic, cytogenic, and mutagenic events [30].

Previous plant reporters

One of the earliest plant bioassays was the *Allium cepa* assay of root mitosis. This test was used to study radiation exposure and later chemical mutagenesis. In the 1970s plants became the biomonitor of choice for screening mutagens. Many of the current plant biomonitors are employed for root tip assays, similar to *Allium cepa*, where the endpoints are mutation and chromosomal aberration [31].

Tradescantia palludosa can be used for different cytogenetic tests, including the *Tradescantia* micronucleus bioassay. Acentric chromosome fragments resulting from damaging agents lead to the formation of pollen mother cell micronuclei (MCN). The assay is simple, rapid, and responds to different mutagens. However, mutations such as translocations, inversions, and other rearrangements do not induce MCN formation. Besides cytogenetic assays, *Tradescantia* can also be used for mutagenic tests. The *Tradescantia* stamen hair system reveals mutation through a change in pigmentation from blue to pink. Plants that are heterozygous for flower color (with blue as the dominant allele) are chronically exposed to mutagens, either radiation, air pollutants, or vapors of mutagens. A mutation induced early in floral tissue results in isolated pink sectors, and stamen hairs are ideal for detecting late mutations which only contain one or a few cells. This assay is very sensitive to IR and has a linear response to mutagens up to three weeks, but does not respond to all contaminants [32]. *Tradescantia* has also been used to study complex mixtures: X-rays and chemical mutagens, as well as mine-dump material, which included heavy metal ions and organic toxicants [2].

Micronucleus tests have been used in the past to study the genotoxic effects of heavy metal ions in plants. In one study using contaminated soil, pollen tetrad cells of *Tradescantia* showed increased micronuclei with high heavy metal ion presence, but no effect was seen in meristematic root tip cells of *Vicia faba* [33]. In a similar study, micronucleus assays were used to observe the effects of different metal ions at various concentrations. *Tradescantia* pollen mother cells showed micronucleus formation at low concentrations, and higher concentrations were required for micronucleus formation in *Allium cepa* and *Vicia faba* meristematic root tip cells. Ranked in order of effectiveness at inducing micronuclei were As^{3+}, Pb^{2+}, Cd^{2+}, Zn^{2+}, and Cu^{2+}, with Cu^{2+} causing little or no effect. Interestingly, aqueous soil extracts did not produce the same effect as the corresponding contaminated soil sample [34].

A forward mutation reporter was developed in *Arabidopsis* by Yoshihara, Nakane, and Takemoto to parallel reporters in animals [3]. A plasmid encoding the *E. coli rpsL* gene was integrated into the *Arabidopsis* genome to serve as a target for mutagenesis, similar to previous work in mice and zebrafish [35, 36]. Plants harboring the transcriptionally inactive insertion were treated with mutagens and allowed to grow, then DNA from the plants was isolated. The integrated plasmid DNA was rescued and transformed into *E. coli*, which was subsequently positively selected for streptomycin resistance. Any inactivating mutation in the *rpsL* gene yielded resistance. In the initial characterization of the lines, the plants were treated with EMS, and a 20-fold increase in mutation was observed, with predominantly G:C → A:T mutations. The spontaneous level of mutation observed was 2.5×10^{-5}, similar

to that seen in animals with the *rpsL* target gene. EMS treatment yielded a mutation frequency of 5.7 x 10^{-4}, with the majority of mutants G:C → A:T. This mutation pattern reflects the preference for EMS-induced O^6-ethylguanine to pair with thymine. The integrated *rpsL* gene system is useful for comparing mutation specificities of plants and animals but the process is complex and is only able to measure inactivating mutations [3].

Transgene reporters

A more effective plant biomonitor than those previously discussed would show sensitivity at the molecular level, similar to transgenic animals that contain a target marker gene for mutation. A marker either goes from active to inactive or inactive to active form upon mutation. The marker gene product has to be nonessential to avoid selection bias and easy to visualize. The promoter used for these marker genes should be strong and active in the tissues of interest [2].

Reporter genes

Reporter genes are useful tools for observing various events, from mutation to protein expression and localization. There have been several reporter genes used in plants, each with advantages and disadvantages. *LacZ*, which encodes β-galactosidase, is a common reporter gene. The β-galactosidase assay is straightforward and uses inexpensive substrates, but can be difficult to quantify in plants due to high endogenous activity [37]. *E. coli* β-glucuronidase (GUS), from the gene *uidA* [38], uses several substrates (such as 5-bromo-4-chloro-3-indolyl-B-D-

glucuronic acid, or X-Gluc, Figure 1) that are fairly inexpensive and the assay is also straightforward. Unlike β-galactosidase, GUS is not affected by background protein levels in the plant. Although there is an endogenous *Arabidopsis* GUS-homolog, it is most active at pH 4.0, versus the *E. coli* GUS enzyme which is assayed at pH 7. At neutral pH, the endogenous plant GUS is inactive [39]. GUS activity is not sensitive to N-terminal fusions, and the half-life of GUS in living mesophyll protoplasts was measured to be 50 hours. Additionally, fluorescent substrate 4-methylumbelliferyl β-D-glucuronide (MUG) is sensitive enough to be visualized in single cells [37].

The Kovalchuk/Hohn reporters

Kovalchuk, Kovalchuk, and Hohn [40] produced transgenic *Arabidopsis* plants with a 35S-promoter-driven *GUS* gene which had been inactivated by point mutation or was a substrate for activation by homologous recombination. When a GUS-reversion point mutation or recombination event occurs in a cell, that cell and all resulting daughter cells express active GUS enzyme, which is visualized as a blue spot or sector under X-Gluc staining conditions. Each event, therefore, is isolated and able to be quantified for mutation or recombination rates. These mutation reporters have been used to analyze effects of various mutagens. Ionizing radiation caused both increased point mutation and homologous recombination, with chronic doses being more mutagenic than acute doses [2]. Plants grown on soil sampled from Chernobyl, a site of radioactive contamination, showed a dose-dependent increase in mutation [41]. These plants also provided evidence for adaptation to chronic irradiation, as they

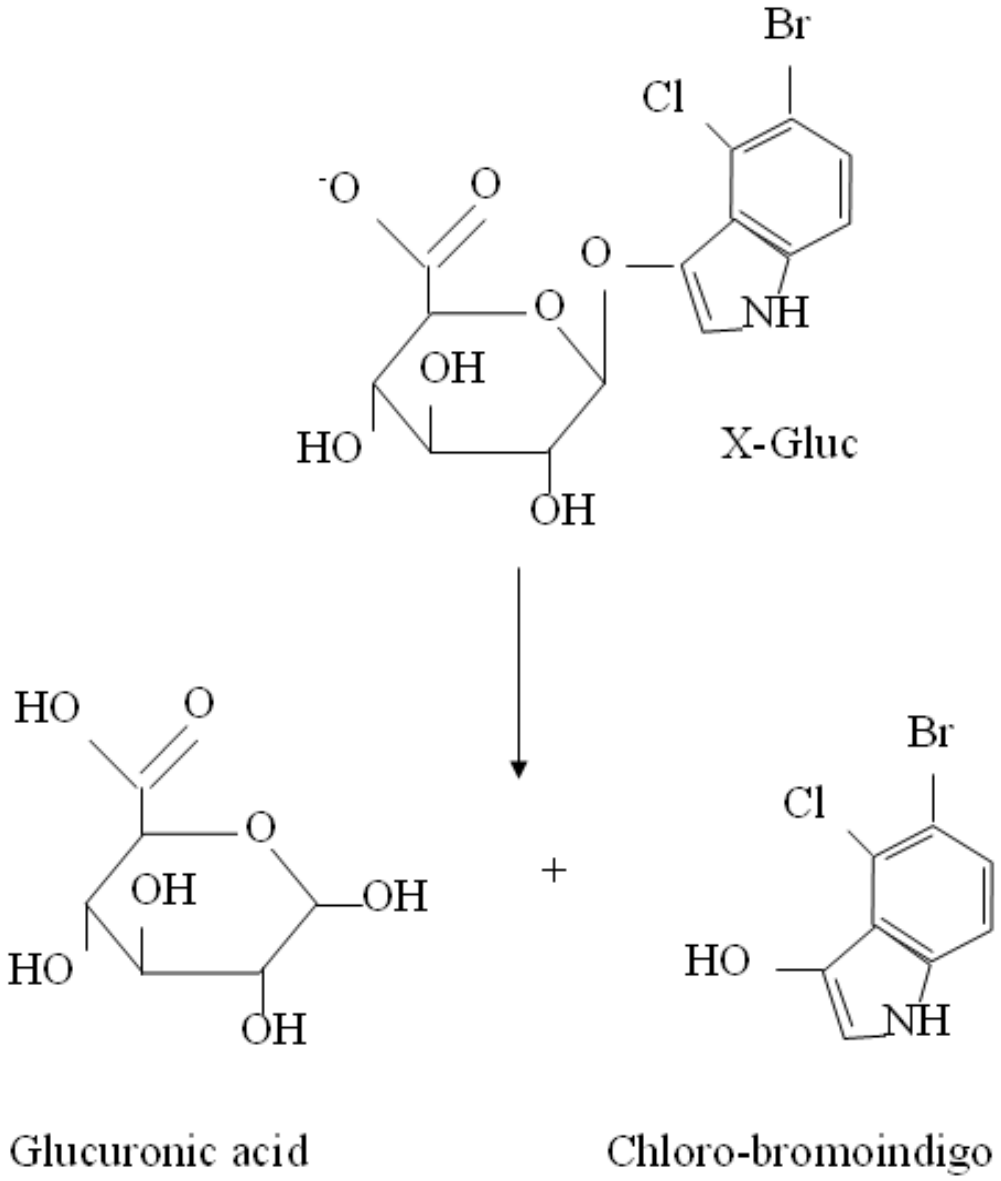


Figure 1. X-Gluc cleavage. The reaction catalyzed by GUS for histochemical staining. Once the glycosidic bond is cleaved, chloro-bromoindigo is able to form a blue precipitate upon oxidation.

showed more efficient recombination repair than control plants [5]. UV-B induction of homologous recombination was also dose-dependent, with both somatic homologous recombination and meiotic recombination increasing two- to five-fold relative to background levels. Heavy metal ions also increased induction of both point mutations and homologous recombination in an uptake- and dose-dependent manner (Cd^{2+}, Pb^{2+}, Ni^{2+}, Zn^{2+}, Cu^{2+}, and As_2O_3 ions). Contaminated soils with complex mixtures of heavy metal ions also caused increases: four to seven-fold increase in homologous recombination and five to ten-fold increase in point mutation [42]. The point-mutation reporters showed a low background reversion rate (three to five out of 100 plants), so 1000 plants were screened for each experiment, for statistical significance. Variance in mutation frequency was observed from line to line. Overall, this set of biomonitors has proven very useful, as it can detect multiple types of damage from various mutagens, and provides an inexpensive and rapid screening assay for damaging agents [2].

Homologous recombination reporter

Initial experiments with the intrachromosomal homologous recombination reporter (**GU**-HG^R-**US)** showed spontaneous recombination rates to be 10^{-6}-10^{-7} per genome. Recombination events were observed in all organs tested and throughout the entire life cycle. Cotyledons and roots showed higher recombination rates than leaves. To confirm that the blue sectors were recombination events, blue and non-blue tissue was isolated and placed on regeneration medium to form calli, which were then tested by X-Gluc staining, HG resistance plating, southern blotting, and progeny testing.

The various staining sector patterns seen indicated when the recombination event occurred. When multiple leaves were stained blue, the event likely occurred in the L2 cell layer of the meristem-shoot apex. An event during early leaf development would show up as a larger section of one leaf. A recombination event in the L3 cell layer would show up as a sector between the mid-rib and inner part of one leaf. A large root sector was indicative of recombination during lateral root development. Finally, a small spot would be due to a late recombination event. Kovalchuk, et al reported their findings as events per genome, as compared to events per cell division, since intramolecular homologous recombination can occur at any time in the cell cycle, and plants undergo a large amount of endoreduplication in leaf tissue. Different cells can also undergo a different number of divisions. Therefore the more relevant number to use for mutation frequency calculations is the total number of genomes as it takes in to account endoreduplication and variable cell divisions. The number of genomes was calculated as total DNA from a plant or specific organ divided by mean DNA per cell [43].

The homologous-recombination reporters were used to study the effect of temperature and day length on recombination. Varying environmental conditions were seen to alter plant genome stability. Homologous recombination could be influenced by vitamin balance, light spectrum, sodium chloride concentration, temperature, and day/night duration. Homologous recombination was increased at 4º C and 32º C compared to the standard 22º C, temperatures which lead to smaller plants with less DNA and possibly less endoreduplication. This effect negatively correlated

with plant metabolic rate and positively correlated with peroxide concentration. Homologous recombination was also shown to decrease with increasing day length, when plants show increased transgene activity and endoreduplication levels. An 8-hour-light, 16-hour-dark schedule led to 15-fold higher recombination than continuous light [44].

Point mutation reporters

The Kovalchuk/Hohn point mutation reporters [40] all encode nonsense mutations at five loci within the 5' end of GUS, and can report mutations from an A:T base pair to G:C, C:G, or T:A. Other possible reversion products were expressed in *Nicotiana plumbaginifolia* protoplasts and tested for activity with the MUG assay. None of the amino acids tested allow activity over 2.2% relative to the wild-type activity. Therefore only the defined reversion pathway which restores the wild-type codon will lead to scorable-active GUS events. To confirm that blue-staining event tissue contained the correct mutation, DNA was extracted from blue tissue, amplified, cloned, and sequenced. The sequences isolated included both the stop codon and the wild-type reversion, which is the expected result from a heterozygous genotype (the reversion only needs to occur in one allele for GUS activity). All sequences produced from control, non-blue tissue contained the defined stop codon. Mutation frequency was defined as numbers of reversion events per plant (stained at the full rosette stage – four to five weeks growth) per *uidA* copy number, as each copy provided an opportunity for reversion. The background reversion frequencies measured in the different lines ranged from 0.0003 to 0.26 events per plant (Table 1). This was

Table 1. Frequency of spontaneous and induced reversions in the Kovalchuk/Hohn point-mutation reporters.

Line	Copy number	Mutant frequency[a]	Mutation ratio (induced/spontaneous)	
			UV-C	MMS
$49_{C\rightarrow T}$-3	1	0.020	1.5	2.0
7	1	0.003	3.3	3.3
1	2	0.005	2.0	2.0
4	4	0.003	3.0	2.0
5	7	0.0003	1.5	1.5
2	8	0.0005	2.5	2.5
$112_{G\rightarrow T}$-2	1	0.010	21.0	3.0
6	1	0.030	11.3	1.7
1	2	0.005	10.0	1.0
4	2	0.020	1.5	1.0
3	3	0.003	56.0	15.0
5	5	0.004	8.0	2.0
$118_{A\rightarrow T}$-4	2	0.005	19.0	6.0
2	3	0.013	3.8	2.8
5	3	0.033	36.1	2.8
6	3	0.007	6.5	2.0
3	5	0.006	3.3	3.0
1	6	0.002	3.0	1.0
$166_{G\rightarrow T}$-2	1	0.020	12.5	4.0
14	1	0.020	5.5	2.0
20	1	0.140	6.6	2.7
4	2	0.030	17.8	4.7
18	2	0.120	2.7	2.4
1	3	0.130	5.0	1.8
$424_{G\rightarrow T}$-1	8	0.003	4.5	1.1
$166_{G\rightarrow A}$-1	1	0.020	29.0	5.5
2	1	0.260	1.2	1.1

[a] Mutant frequency calculated as [(total # reversions)/(total # plants)]/(copy number). Data is from Kovalchuk, et al [40].

estimated to be equivalent to a rate of 10^{-7} – 10^{-8} events per base pair, which agrees with other *Arabidopsis* studies yet is greater than 100-fold higher than other eukaryotes. The point-mutation lines were also tested with UV-C, methyl methane sulfonate (MMS, an alkylating agent), and X-rays. MMS treatment (50 μM) led to a two-fold increase in reversion, and X-rays (absorbed dose equal to 25 Gy) led to a three-fold increase in reversion. UV-C treatment (1000 J/m^2 after two weeks growth) led to various amounts of UV-specific mutation induction (up to 56-fold over background), from a median of 0.0055 induced events per plant (G<u>T</u>A → G<u>C</u>A reporter line 49) to a median of 0.365 induced events per plant (T<u>T</u> → T<u>G</u> reporter line 166). The T → C reporter in the same codon (TTC<u>T</u> → TTC<u>C</u> reporter line 166) also reverted at a high median UV-specific mutation frequency of 0.305 events per plant. These two lines were used for future studies. The UV-specific-mutation frequencies varied among the different constructs as well as between different sublines of the same construct. Therefore both sequence contexts are important factors in mutation frequency. It was anticipated that increased copy numbers would increase the mutation frequency, yet this was not seen, possibly due to silencing effects from the number and location of integration sites. Among the lines with single-copy insertions, the level of transcription negatively correlated with mutation frequency, perhaps through increased levels of transcription that allowed better repair of mutagenic lesions [40].

The homologous-recombination reporters were used in conjunction with point-mutation reporters to analyze the effect of heavy metal ions on mutagenesis.

Kovalchuk, et al tested plants with Cd^{2+}, Pb^{2+}, Ni^{2+}, Zn^{2+}, Cu^{2+}, and As_2O_3 in both media and soil. The plant lines were highly sensitive, showing varying increases in base substitutions and intrachromosomal recombination. A:T → C:G mutations were primarily induced by Pb^{2+} and Zn^{2+}, while A:T → G:C mutations were primarily induced with As_2O_3 and Cd^{2+}. As_2O_3 induced more recombination than base substitutions at low concentrations, and favored base substitutions at high concentrations. Pb^{2+} and Ni^{2+} induction of recombination quickly reached a plateau, but induction of base substitution increased linearly. The variations in response may be due to differences in uptake efficiency as well as the mechanism of mutation or recombination induction. Polluted soils also increased point mutation and recombination. Unlike other tests which are sensitive to only certain metal ions, this test reacted to every metal tested at biologically relevant concentrations [9].

The Depicker C → T reporters

A recent reporter system that partially complements the Kovalchuk/Hohn reporters is a set of lines developed by Van der Auwera, et al. Five missense C → T mutation reporters were introduced within a chimeric *uidA* transgene. The GUS transgene was fused to green fluorescent protein (GFP) with or without hemagglutinin and strep II epitope tags. The point mutations were isolated within the catalytic pocket. The only possible reversion pathways are C → T transitions, except one construct which could also revert by a C → G transversion. The various mutant constructs were expressed in *E. coli* and tested to confirm the mutant GUS was inactive. However, products of amino-acid substitutions resulting from alternative

reversion pathways were not tested for GUS activity. High GFP-signal intensity was observed with high GUS staining across the plant, yet GUS staining was also observed in some lines with low GFP signals. Segregation patterns on selective media were analyzed, however no attempt was made to specifically isolate single locus inserts. 24/81 independent T2 populations of the missense reporters showed spontaneous reversion frequencies that were similar to the Kovalchuk/Hohn point mutation reporters. Spots were seen more frequently than sectors, indicating a preference for mutations later in development.

Spontaneous point mutations, EMS- and UVC-induced mutations were observed. EMS treatment induced large increases in both spots per plant (26-375 fold) and percentage of plants containing spots (4-127 fold). UV-C induction of C → T was low but significant (two- to six-fold). The low induction was likely due to the treatment strategy: plants were grown on synthetic medium, were only exposed to a dose of 80 J/m^2 UV-C, and were immediately placed back in the light. Thus the levels of damage were likely to be low, and any photodimers formed would be quickly repaired by photolyases. This accounts for the UV-induced mutation observed. The lines that were able to form dimers did not show higher induction than lines unable to form dimers at the reversion site. GFP signals were seen in lines with and without spots, indicating that detectable expression did not correlate with observed mutation frequencies. Factors that may influence mutation frequency include characteristics of the reporter system, plant growth and treatment, and differences in the genetic background due to the transformation process, such as disruption of a DNA repair

pathway. As opposed to mutation induction by EMS and UV-C, heavy metals (Pb^{2+} and Cd^{2+}), methyl jasmonate, salicylic acid, and heat and light stress did not induce mutation [27]. This suggests lack of response to reactive oxygen species [10]. The authors stressed the need, therefore, for several different scoring systems, so as to detect all types of damage from many different mutagens [27].

Construction of a set of point mutation reporters

Chimeric GUS

The version of *GUS* used in this study and the Kovalchuk/Hohn studies is a fusion gene with the translational start site of *E. coli uidA* replaced by the first 29 amino acids of ORF V from cauliflower mosaic virus (CaMV). This chimeric gene shows enhanced activity [43]. ORF V encodes the reverse transcriptase of CaMV. In a study of the transcription and start codons of ORF V, the first 87 bases were fused with *GUS* lacking its start codon and expressed in a transient expression system of protoplasts from the host plant *Orychophragmus violaceus*. When compared with wild-type *GUS* expressed from pBI221, the ORF V-fusion GUS showed higher expression and a 23-fold increase in activity. This seems to be due to a better Kozak sequence in the chimeric gene [45].

The Kozak sequence refers to a preferred sequence context surrounding the start codon of a gene that allows optimal translation of the mRNA. The core sequence is 5‘ACCATGG 3’, with the underlined ATG as the start codon and the first A being critical. A purine in the 5' site is important for good translation [46]. The pBI221

plasmid harboring 35S-promoter-driven GUS, which has been used as a standard for GUS assays [47], encodes an unfavorable start sequence: 5‘CTTATGT 3’. Conversely, the ORF V N-terminal fusion provides an optimal start sequence: 5‘ACCATGG 3’. It is unknown whether the rest of the fusion plays any role in the increased expression or activity.

Codon selection for mutation constructs

One of the codons mutated in the Kovalchuk/Hohn study [40] was codon ‘112’, which refers to a defined base substitution in the codon starting at base 112 of the CaMV ORF V-GUS fusion gene. This codon is the 38th codon in the ORF V-GUS fusion gene and the 11th codon in wild-type *E. coli* GUS. This codon was chosen for our mutation study for several reasons. All six transition and transversion pathways were able to be tested within the codon, with ideal sequence contexts for particular mutations of interest (e.g. T → C in a TT context and C → T in a TC context for UV mutagenesis). Additionally, the existing Kovalchuk/Hohn 112 mutant lines showed moderate levels of reversion, the location of the codon was ideal for mutation construction, and amino acid substitutions that were tested for activity indicated the requirement for the wild-type amino acid for GUS activity.

Transformation considerations

Transgene expression and position effect variegation

Up to a certain point, gene copy number correlates with expression of transgenes. Stable, equivalent levels of expression are observed among equivalent

copy number transformants. RNA interference (RNAi) is triggered at a gene-specific threshold. For 35S-promoter-driven GUS in *Arabidopsis*, RNAi is typically triggered by the presence of three or more copies. RNAi was not affected by the location of the multiple copies or by their integration sites, even when in pericentromeric heterochromatin [48]. Another study in *Arabidopsis* suggested that when the transformants harbored a single insert of the transgene, position effects were not observed [49]. This is in contrast to observations with other organisms where variable levels of transgene expression were dependent upon location within the genome and the corresponding chromatin structure. A third *Arabidopsis* study confirmed that a single, complete copy of T-DNA in any of the five chromosomes showed no position effect, but in fact yielded high expression. The homozygous progeny of single-insert transformants retained the high expression and low variability as well, and the RNA levels were also similar. One hypothesis for the lack of position effect is that all five *Arabidopsis* chromosomes are mostly euchromatin, hence regional transcription levels are less variable and less likely to cause transgene expression variance. Tobacco chromosomes, conversely, are more repetitious and gene-dense, thus more prone to transgene position effects [50]. It is therefore important to select transformants with single insertions to avoid transgene expression variation between lines.

Transgene silencing

Another source of expression variation between transformed lines is transgene silencing. Sources of transgene silencing include repeat-induced gene silencing (tandemly-repeated T-DNA inserts) and homology-dependent gene silencing (copy-

number dependent); the latter includes both transgenes and their homologous endogenes. Silencing increases with homozygosity, copy number, and increased transcript level, and can also be influenced by transgene structure (complete sequence versus inserted or deleted fragments) [50]. *Agrobacterium*-mediated transformation [51] can lead to transgene silencing, and was in fact how RNA silencing was discovered. Small interfering RNAs (siRNAs) are important in both antiviral defense and transgene silencing. RNAse III (Dicer) turns the mRNA from the transgene into dsRNA, which is then cleaved into 21- to 24-nucleotide RNAs that are incorporated into RNA induced silencing complexes (RISCs). These complexes lead to cleavage or translational inhibition of the transgene mRNA and/or epigenetic modification of the transgene [52]. Epigenetic changes include methylation and changes in chromatin structure, similar to position effect variegation in Drosophila and X-inactivation/imprinting in mammals. Silencing also affects the selectable resistance gene. In one *Arabidopsis* study, the *nos* promoter-driven resistance gene was silenced more than 35S-GUS, even though both were on the same T-DNA [53].

There are two types of gene silencing. Transcriptional gene silencing (TGS) targets sequences in the promoter, can be triggered by inverted repeat RNA, transmits meiotically and mitotically, and requires chromatin modification or DNA methylation. TGS generally requires multiple copies of the transgene or very high expression levels. Post-transcriptional gene silencing (PTGS) is sequence-specific mRNA (cytosolic) degradation, leading to short RNA fragment accumulation which maintains silencing. The triggers are increased RNA expression or inverted-repeat, double-

stranded, or otherwise aberrant RNA. Gene expression level and dosage are therefore critical, but some genes are more sensitive to PTGS than others. PTGS can also be developmentally controlled [54].

The Gelvin superpromoter system

In an effort to produce a strong promoter for transgene expression, a "superpromoter" [55] was constructed in the Gelvin lab using octopine- and mannopine-synthase promoter elements. The Gelvin superpromoter consists of a trimer of the octopine synthase upstream activating sequence (*Aocs*) and the mannopine synthase promoter/activator region (*AmasPmas*). Superpromoter-driven *uidA* expression induced increased levels of GUS in all tissues examined, both in different plant species and different growth conditions (soil or media). Tobacco leaves expressing GUS with the Gelvin superpromoter showed 156-fold higher expression (measured by activity) than with a 35S promoter, and 26-fold higher expression than with a double 35S promoter. A close correlation was also seen between steady state mRNA levels and GUS activity. Ni, et al [56] examined fully expanded leaf tissue, nearby stem tissue, and young root tissue through MUG assays of protein extracts. The superpromoter induced expression in most cell types, although the different promoter elements themselves are specific for different cell types. For example, the mannopine synthase promoter (*Pmas*) elements are highly induced in the roots and are wound- and auxin-inducible. *Pmas* contains the *As-1* tandem repeat motif, which results in high root expression [56]. A diagram of the Gelvin superpromoter used for this study can be seen in Figure 2. *Aocs* is a cis element which enhances transcription

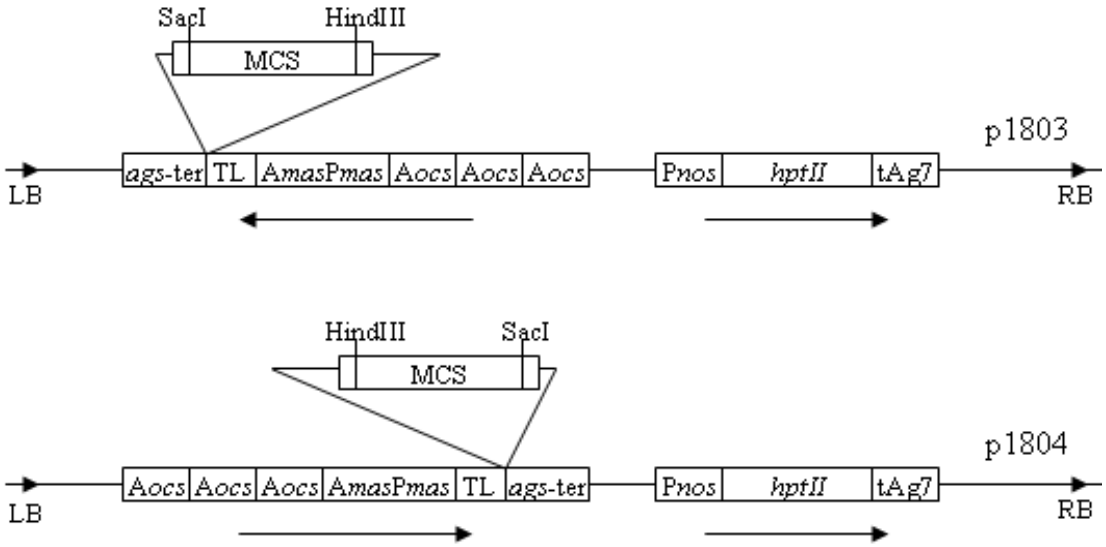


Figure 2. Maps of the T-DNA binary vectors containing the Gelvin superpromoter. A*ocs*, octopine synthase upstream activation element; A*mas*P*mas*, mannopine synthase promoter and upstream activation element; TL, translational leader from TEV 5'NTR; MCS, multiple cloning site; *ags*-ter, poly(A) addition signal from the agropine synthase gene; Pnos, nos promoter; hptII, gene conferring resistance to hygromycin; tAg7, poly(A) addition signal for T-DNA gene 7; LB, left T-DNA border; RB, right T-DNA border. Arrows beneath the map indicate the direction of transcription.

irrespective of orientation [57]. pGPTV-HPT is the binary vector backbone for the Gelvin superpromoter, with the right border near the multiple cloning site and the left border near the resistance gene [58]. The translational leader contains the first 144 nucleotides from the tobacco etch virus (TEV) 5'NTR (nontranslated region). The RNA genome of TEV serves as an mRNA for construction of a large polypeptide, but unlike most mRNAs it does not begin with a 5' cap. The 5'NTR allows cap-independent enhancement of translation, potentially through internal sites which bind proteins or ribosomes. When the 5'NTR was fused to a GUS reporter transgene, translation was increased 8-21 fold relative to wild-type GUS [59].

Expression patterns

Different patterns of histochemical GUS staining have been observed in the past. Expression arising from a clonal L2 sector will extend to the leaf margin, whereas an L3-like sector staining will not spread to the edge [60]. An expression pattern that is specific to vascular tissue has been observed when GUS was expressed using the promoter from the commelina yellow mottle virus (CoYMV). Whereas the 35S promoter from CaMV, another dsDNA virus, can drive transgene expression in all cell types, the CoYMV promoter is specific for vascular expression. Although 35S promoter-driven GUS can be observed in all tissue types, the vasculature stains most intensely, and flowers only show a vascular stain. Therefore a common vascular-specific promoter element exists in both CoYMV and CaMV promoters, and the differences in expression between the promoters are likely due to the influence of the other cis-elements [61].

A study using tobacco focused on GUS expression driven by the mannopine synthase promoter (*Pmas*). In the leaves, stems, and roots, there was no correlation between mRNA and protein activity. Older tissues showed increased GUS activity due to the stability of the GUS protein, which likely accumulates with age. Roots showed higher activity than stems or leaves; *in-vitro*-grown plants showed higher GUS activity than those grown in a greenhouse, although the mRNA levels were equivalent [62]. A second tobacco study produced similar results, with roots, cotyledons, and older leaves showing higher activity than young leaves. *Pmas* was additionally inducible by chemicals important to plant development: auxins, cytokinins, salicylic acid, and methyl jasmonate. In adult plant tissue, activity tended to occur in vascular tissue [63]. A recent study of expression of *GUS* driven by the Gelvin superpromoter used tobacco (dicot) and maize (monocot) plants. Roots showed higher activity than leaves in both plant species (five-fold higher in tobacco), yet only the tobacco showed higher activity in mature versus young leaves. In the maize system, the superpromoter showed activity as high as 35S, Pmas2', and maize ubiquitin promoters, and in tobacco the superpromoter yielded two- to twenty-fold higher activity than double 35S. The addition of the TL enhancer further increased the gene expression two- to three-fold, particularly in dicots. GUS activities within plants were determined by MUG assays of tissue punches [55].

A study in *Arabidopsis* by De Bolle, et al [64] analyzed the role of promoter elements on expression. A 35S-promoter-driven GUS showed equivalent expression in all plant tissues, but displayed a bimodal distribution of variable expression among

independent transformants – roughly two thirds being low (possibly silenced) and one third high. Conversely, a promoter [*(Aocs)$_3$Pmas*] similar to the Gelvin superpromoter showed increased median expression and decreased variability of expression relative to 35S. The roots showed higher activity than stems or leaves, and the variation among independent transformants was normally distributed, suggesting an absence of transgene silencing. The authors suggest the decreased variance shown by this promoter was due to its key cis-acting elements being less sensitive to methylation [64]. Another study of transgene expression in plants deficient in PTGS revealed similar resistance of the *(Aocs)$_3$Pmas* promoter to silencing. 35S-driven GUS in the wild-type lines resulted in 10/13 plants with low expression and 3/13 plants with high expression, and in the PTGS-deficient lines all transformants showed high expression. The *(Aocs)$_3$Pmas*-driven GUS, however, showed moderate expression with low variability that did not change in the PTGS-deficient background [65].

Epitope tags

Epitopes are small peptides that are fused to proteins of interest to simplify immunological detection. The fusion protein is detected using an antibody to the epitope. The transgenic protein can be distinguished from any potential homologous endogenous protein, and detection of an epitope relieves the requirement to use a unique antibody for every protein of interest. C-myc is a commonly used epitope, based on the human c-*myc* oncogene. However, it has been shown to cross-react with plant proteins. Previous studies have also indicated immunological measurement of expression with c-myc fusions yielded quantitatively less protein than expected or

none at all [66]. In one case, a c-myc tag functioned well for immunoprecipitation and immunofluorescence, but not for western blotting [67]. In a second study, recognition efficiency of a c-myc tag varied with its fusion to different proteins, due to their tertiary structure [68]. A more recent epitope, AcV5, consists of nine amino acids from the GP64 envelope fusion protein of *Autographa californica* multiple nucleopolyhedrosis virus. The AcV5 epitope provides a strong signal on western blots without cross-reacting with endogenous plant proteins [66].

Measurement of microsatellite instability

My initial graduate research involved the examination of MMR activity in *Arabidopsis* through measurement of microsatellite instability (MSI) in AtMSH2 deficient lines. Microsatellites are repeats of one to several nucleotides in various lengths, and repeat length polymorphisms are a hallmark for MMR deficiency. Extrahelical loopouts induced by slippage during replication (transient melting followed by out-of-frame reannealing) are one of the major targets for MMR. As the length of the microsatellite increases, so does the potential for loopouts, and consequently repeat-unit insertions and deletions. Identification of cells or cell clusters with altered microsatellite lengths is difficult within a whole organism, so to circumvent this we used two approaches. First, Leonard introduced a mononucleotide run into the N-terminus of GUS (out-of-frame) to visualize independent frameshift events. Single insert, single copy transgene lines were isolated, and a G_7 repeat was selected for further testing. When treated with 1200 J/m^2 UV-C, frameshift events increased five-fold, similar to other studies in *E. coli*. Additionally, background MSI in one line was measured across five generations and remained constant. When introduced into an AtMSH2 deficient background, frameshifting increased five-fold. However, this was less than the observed increases in yeast and mice with comparable mononucleotide runs [69].

In the second approach, I measured endogenous microsatellite allele length in progeny to determine MSI in a parent plant. Studying endogenous microsatellites allows better comparison with other organisms, and avoids the potential interference

of transcription-coupled repair and gene expression effects. Microsatellites are affected by chromosome location and type and length of repeat. Progeny are tested since the majority of the plant must contain the altered length allele to allow detection. The background signal of the assay is fairly high, and altered length alleles which constitute less than 20% of the total template extract are undetectable. Therefore any detectable deviations most likely occur in the germinal tissue, which gametophytic tissue is derived from, or else very early in progeny development. This analysis provides germinal mutation data, but the MSI should be reported as unique versus total shifts per allele per progeny set, as differentiation among an early mutation, multiple late mutations, or an intermediate of these is impossible. In a set of six microsatellites with $\geq$ 25 repeats, at least one shift was seen in every set of progeny from each of three AtMSH2::TDNA lines, in comparison with wild-type Col-0 lines which only showed one unique (two total) shift(s). In addition, some data sets could not be scored, as it appeared that the parent plant was heterozygous and segregating for two different length alleles. Endogenous MSI was also measured in AtMSH2-RNAi lines, which confirmed the expected MSI phenotype (Figure 3). AtMSH2 deficiency led to increases in endogenous MSI that were similar to MSI levels observed in yeast and mice studies. The low levels of endogenous MSI in relation to transgene MSI in somatic tissue indicate that MMR is critical in germinal tissue yet less important for differentiated tissue. Previous measurements indicated MMR protein expression was low in leaves. Accumulation of DNA damage in fully differentiated leaf tissue is not as harmful as accumulation of somatic mutations in

Figure 3. Frequency (%) of repeat length-shifted alleles in endogenous microsatellite loci. Frequencies were calculated by dividing numbers of unique length shifts (A) or total numbers of all alleles showing non-parental lengths (B) in the progeny of two *AtMSH2*$^{+/+}$ plants or of three *AtMSH2*$^{-/-}$ plants or of four individual *AtMSH2*(*RNAi*) plants by the total number of alleles tested in each group. Data represent sums for all nine loci analyzed or for the six loci with $\geq$25 repeat units. No length shifts were detected in progeny of two plants transformed with the empty binary vector FGC5941. Figure and legend from Leonard, et al [69].

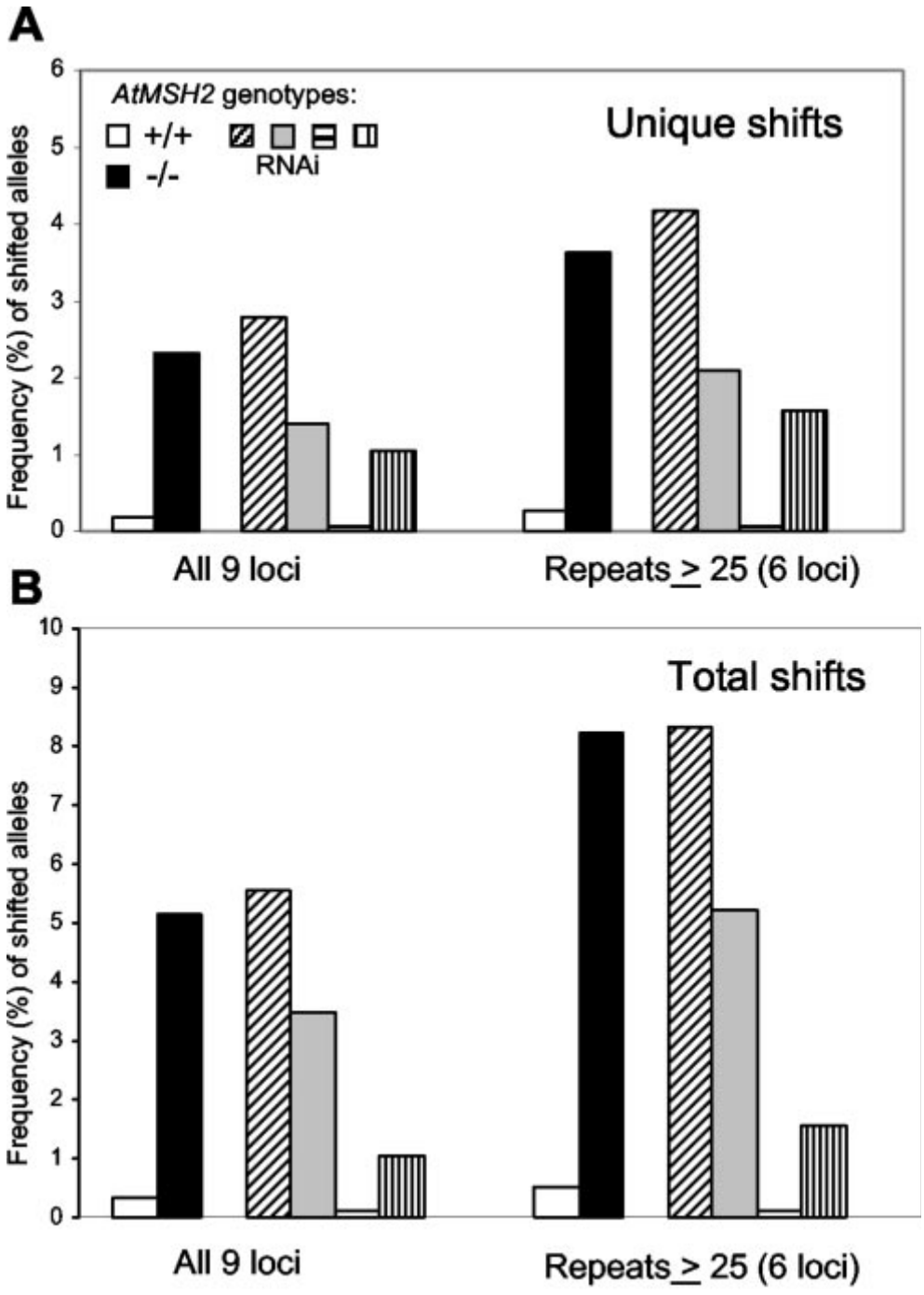


Figure 3. Frequency (%) of repeat length-shifted alleles in endogenous microsatellite loci.

mammals, which can develop cancer [69].

A second study in our laboratory monitored genome stability in an MMR-deficient line by growing AtMSH2::TDNA lines for several generations in parallel to wild-type controls. A progeny plant was randomly selected for propagation in each line to minimize bias. Lines were measured at the fifth generation (the first visible mutations occurred in the second generation), at which point 34/36 lines showed a phenotypic change, with four lines already extinct due to deleterious mutations. In comparison, all 36 wild-type lines remained normal. I measured MSI in these various lines. The G5 progeny were tested, so G4 MSI was being measured. Wild-type plants showed no alterations in microsatellite length. In comparison, one plant from each AtMSH2-deficient line was tested, and each showed from one to ten length shifts in the six loci (twelve alleles) tested. Six lines were then chosen at random and used for more extensive progeny testing. Five out of the six lines showed one to three new "baselines" for shifting: a previous generation accrued a microsatellite length change which was stably inherited as the new "wild-type" repeat length (Figure 4). Correcting for these baseline changes allowed estimation of G4 → G5 MSI rates, which were equivalent to G0 → G1 rates (Figure 5). Therefore the observed phenotypic changes in the G5 lines were most likely due to mutation accumulation and not additional repair deficiencies [70]. These two studies can be viewed in more detail at: http://www.plantphysiol.org/cgi/content/full/133/1/328, http://www.genesdev.org/cgi/content/full/18/21/2676.

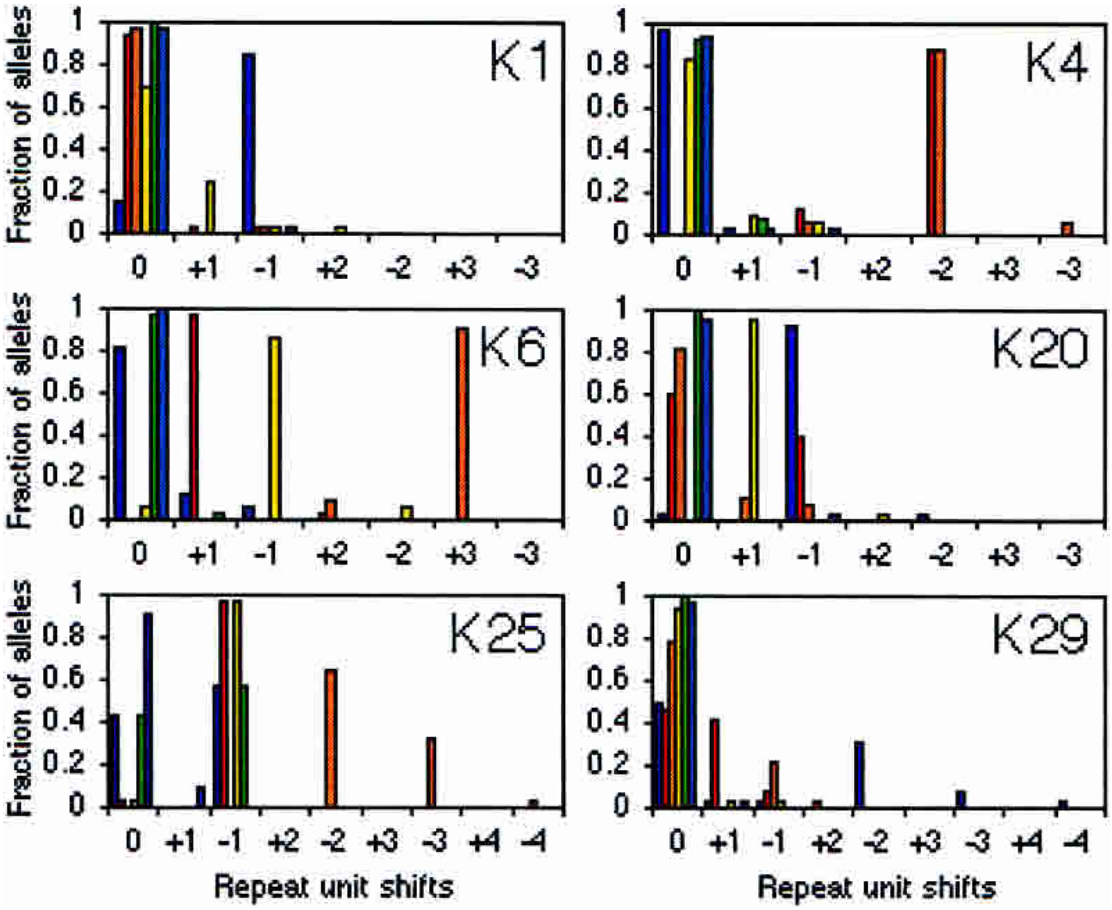


Figure 4. Microsatellite analysis of G5 *Atmsh2-1* plants. Microsatellite "fingerprints" for indicated *Atmsh2-1* lines: fractions of the 32 alleles (16 plants) that are shifted by the indicated numbers of dinucleotide repeat units relative to G0 *Atmsh2-1* plants (purple, NGA6; red, NGA8; orange, NGA139; yellow, NGA151; green, NGA172; blue, NGA1107). Figure and legend from Hoffman, et al [70].

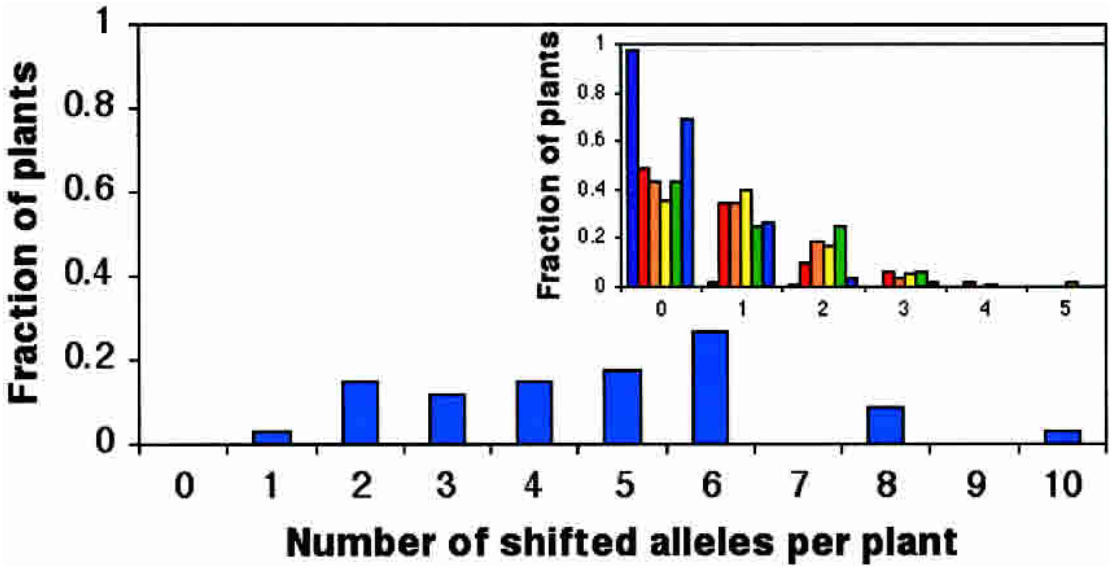


Figure 5. Representative microsatellite genotypes for all *Atmsh2-1* (*AtMSH2*::*TDNA*) lines. One G5 plant from each line was analyzed for MSI at six loci (12 alleles) and the fraction of these plants showing the indicated numbers of cumulative shifts was determined (dark blue). (*Inset*) Similar analyses for G0 → G1 shifts of wild type (purple), *AtMSH2-1*::*TDNA* (red), *AtMSH2-RNAi* (orange), *AtMLH1*::*TDNA* (green), and *AtPMS2*::*TDNA* (blue), and for *Atmsh2-1* G4 → G5 (yellow). Figure and legend from Hoffman, et al [70].

Materials and Methods

Construction of plasmids

Standard techniques were used for all DNA manipulations [71], and reactions were carried out in accordance with manufacturer instructions. Qiagen kits were used for bacterial DNA preparations, polymerase chain reaction (PCR) clean-up, and extractions of DNA from agarose gels. Oligonucleotides were purchased from MWG (Huntsville, AL), and enzymes were purchased from NEB (Ipswich, MA) and Fermentas (Glen Burnie, MD). Plasmids for GUS construct cloning (including pGUS23 [72]) were generously donated by Igor Kovalchuk (University of Lethbridge, Canada), and the Gelvin superpromoter plasmid set was licensed through Stan Gelvin (Purdue University).

Polymerase Chain Reaction

Standard PCR conditions were as follows: 1 μL DNA template (various concentrations depending on reaction), 2 μL 10X PCR buffer (500 mM KCl, 100 mM Tris HCl pH 9.0, 1% Triton-X), 2 μL 2.5 mM dNTP mix, 1 μL 5 μM forward primer, 1 μL 5 μM reverse primer, 0.8 μL 50 mM $MgCl_2$, 0.2 μL Taq polymerase, and 12 μL distilled water in a total volume of 20 μL. A standard reaction consisted of an initial 94º C incubation for 3 minutes, followed by 30 rounds of amplification (94º C for 30 seconds, 54º C for 30 seconds, and 72º C for an estimated 1 minute per 1 kb), and a

final 72º C incubation for 2 minutes. All PCR amplifications were incubated in a Robocycler Gradient 96 thermocycler (Stratagene, La Jolla, CA).

Bacterial transformation

For the transformation of *E. coli* by a chemical method, 1-5 μL of vector DNA was added to 100 μL of thawed chemically competent cells. This mixture was placed on ice for 30 minutes, then incubated at 42º C for 30 seconds to heat shock the cells. Two hundred fifty microliters of SOC medium (20 g/L tryptone, 5 g/L yeast extract, 0.585 g/L NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 0.36% glucose) was added and the cells were incubated at 37º C for one hour, then plated on selective medium (ampicillan, AMP, or kanamycin, KAN).

For *Agrobacterium* transformation, strain GV3101 [73] was chosen, since it works well for floral dip transformation [74]. This strain was generously donated by Walt Ream (Oregon State University). To perform chemical transformation, 2 μL of vector DNA was added to 100 μL of frozen chemical-transformation competent cells. The cells were thawed by hand and mixed, followed by incubation in liquid nitrogen for five minutes. Cells were allowed to thaw at room temperature for five to ten minutes, 1 mL TBY medium (10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl, pH 7.5) was added, and cells were incubated overnight at 28º C. Cells were plated on kanamycin selective medium 16 hours later.

For transformation of bacteria by electroporation, 1 μL of vector DNA was added to 50 μL of thawed electroporation competent cells. This mixture was placed in

a cuvette and exposed to 2.5 kV of electricity. The cells were resuspended in 1 mL SOC medium and transferred to a culture tube for an hour incubation (37º C for *E. coli*, 28º C for *Agrobacterium*), then plated on kanamycin-selective medium.

DNA preparation from Agrobacterium

A 48-hour culture of transformed *Agrobacterium* was centrifuged to form a pellet, which was resuspended in 100 μL TE (25 mM Tris, pH7.5, 10 mM EDTA), 200 μL lysis buffer (200 mM NaOH, 1% SDS), and 150 μL 3M NaOAc. Suspensions were centrifuged and the liquid was transferred to a new microfuge tube. One milliliter of 95% ethanol was added and tubes were incubated at 4º C or 22º C for ten minutes (or -20º C overnight), then centrifuged and resuspended in 100 μL TE. After addition of 50 μL 7.5 M ammonium acetate and 300 μL 95% ethanol, tubes were incubated on ice for ten minutes (or -20º C overnight). Tubes were centrifuged and washed with 70% ethanol. Air-dried pellets were resuspended in 50 μL TE, 1 μL RNase and incubated at 37º C for 30 minutes. Finally, 25 μL 3M NaOAc and 150 μL 95% ethanol were added, and DNA was centrifuged and resuspended in 30 μL 10 mM Tris [75].

In vitro *transcription/translation assay*

The Promega (Madison, WI) TNT T7 Quick for PCR DNA system was utilized for protein expression from PCR templates. I constructed the DNA template to contain a T7 promoter and Kozak sequence, which allow efficient expression in the

master mix. Template, master mix, and [^{35}S]methionine were incubated at 30º C for 90 minutes, and the resulting [^{35}S]methionine-labeled-protein was analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and autoradiography, as well as a MUG activity assay.

SDS-PAGE and autoradiography

SDS-PAGE was used to resolve proteins produced in the *in vitro* transcription/translation assays. SDS is a detergent that breaks down protein secondary structure and coats the protein with negative charges, limiting separation on the gel to be due to size of the protein. Ten percent acrylamide gels were used with a 4% stacking gel. The GUS construct products were roughly 70 kDa in size, and were easily identified by exposure of X-ray film to the radioactive proteins in the gel.

MUG assay

GUS activity of *in vitro* synthesized protein extracts was measured as cleavage of MUG, a fluorogenic substrate, into glucuronic acid and 4-methylumbelliferone, the fluorescent product [37]. Five microliters of an *in vitro* reaction was added to 495 μL of MUG assay solution: 50 mM $NaHPO_4$, pH 7.0, 10 mM β-mercaptoethanol, 10 mM Na_2EDTA, 0.1% sodium lauryl sarcosine, 0.1% Triton X-100, and 1 mg/mL MUG substrate. The assay solution was incubated at 37º C, and 5 μL aliquots were taken at various timepoints and added to 195 μL of carbonate stop buffer (0.2 M sodium carbonate) on ice to stop the reaction and adjust the pH to an optimum for fluorescence. Fluorescence was measured on a SpectraMAX Gemini Dual-Scanning

Microplate Spectrofluorometer, with excitation at 355 nm, emission at 460 nm, and cut-off at 435 nm. Endogenous GUS from a DH5α extract was used for the positive control for the assay.

Seed preparation

To prepare seeds for plating, they were counted out into microfuge tubes and treated with 95% ethanol for one minute. The seeds were briefly centrifuged and the ethanol was removed, then seeds were treated with 50% bleach plus 0.05% Tween-20 (one drop per 50 mLs) for ten minutes. Seeds were briefly centrifuged, the bleach solution was removed, and the seeds were washed four times with sterile water, with a final suspension in sterile water. The seeds were vernalized at 4ºC for 72 hours before plating. When preparing seeds for planting on soil, the seeds were counted out into microfuge tubes and suspended in 0.1% agarose, then vernalized as above.

Floral dip

For transformation of *Arabidopsis*, the floral dip method was used [74]. *Agrobacterium* overnight cultures carrying the binary vector constructs were added to 500 mL TBY + KAN media and grown in a 4 L flask at 28º C overnight. Cells (OD_{600} ~ 0.8) were centrifuged for twenty minutes at 5500g at room temperature. The bacteria were resuspended in 1 L infiltration medium (5% sucrose, 0.05% Silwet L-77) and placed in a 4 L beaker. The aboveground tissues of thoroughly-watered wild-type Col-0 plants with secondary bolts (primary bolts were clipped) were dipped in the bacterial solution for three to five seconds, then the pots were set on their sides under

domes overnight. Plants were dipped a total of three times, with one dip every seven days. The dipped plants matured and were harvested.

Isolation and selection of independent lines

T_0 seeds bulked from *Agrobacterium*-transformed plants were sterilized and plated on hygromycin-selective medium at a density of 500 seeds per plate. Resistant plants were transplanted to soil after one week, roughly one to five resistant plants per plate, and allowed to self-pollinate and mature. Individual plants were then harvested, and random progeny sets of 50 T_1 seeds were plated onto hygromycin-selective plates for segregation analysis after seven and eleven days-growth. Those lines whose progeny phenotypes indicated single site insertion (segregating 3:1 for resistant: sensitive) were selected for further testing. Resistant plants (which contained the T-DNA insert) were healthy and produced roots, or in rare instances appeared unhealthy yet produced a root, or appeared healthy and did not produce a significant root. The majority of sensitive plants, especially after 11 days, contained no root, did not grow beyond the cotyledon stage, and usually turned brown. Occasionally, sensitive plants barely germinated, producing either part of a root or a small amount of green tissue. Seeds that did not germinate had nothing visible protruding from the seed coat, and were not included in the segregation-analysis calculations. The cut-off for the probability of a line containing a single locus of transgene insertion was $p < 0.05$, as determined by the Chi-square test: [(observed resistant – expected resistant)^2 / (expected resistant)] + [(observed sensitive – expected sensitive)^2 / (expected sensitive)]. Progeny sets with Chi-square values over 3.84 were rejected.

Preparation of positive control protein extracts for western blots

BL21 Codon+ cells that contained pET21a vector carrying AcVS or c-myc-tagged wild-type *GUS* were grown in TBY + AMP liquid culture overnight, diluted 1/20 into fresh TBY + AMP (two cultures), and incubated at 37º C for 90 minutes. One of the cultures was induced with 0.5 mM IPTG, and both induced and uninduced cultures were held at 37˚ C. Samples were taken from each culture at various timepoints, from 30 minutes to four hours; cells were pelleted and stored at -20˚C. Pelleted cells were resuspended in 2X protein loading buffer (50µL for 30 minute to two hour timepoints, 100µL for three and four hour timepoints) and boiled for three minutes. Aliquots were then separated on a polyacrylamide gel and stained with Coomassie blue (50% methanol, 10% acetic acid, 0.05% w/v Coomassie brilliant blue R-250) to confirm transgene induction.

Plant protein extracts

Two different protein-extraction protocols were used. The first used a Pierce (Rockford, IL) P-PER® Plant Protein Extraction kit, with the addition of Roche (Switzerland) Complete Protease Inhibitor Cocktail. The Pierce BCA™ Protein Assay kit (Reducing Agent Compatible) was used to quantify the Pierce protein extracts. Neither of these kits were reliable in our hands. I ultimately used a plant-protein-extraction protocol similar to that used for extracting the positive controls from *E. coli*. Roughly 50 mg of tissue (three 10 day-old seedlings of four to six leaves each) was frozen in liquid nitrogen and ground up by hand in a microfuge tube. Fifty microliters

of 2X protein loading buffer (0.5 M Tris-HCl pH 6.8, 4.4% w/v SDS, 20% v/v glycerol, 2% v/v 2-mercaptoethanol, and Bromophenol Blue) were added and the frozen tubes were placed on ice. Tubes were boiled for three minutes, vortexed for thirty seconds, sonicated for three minutes, boiled for three minutes, and vortexed for thirty seconds. Homogenized samples were centrifuged for fifteen minutes at 4˚C, then liquid was transferred to a new tube and centrifuged again to minimize cellular debris. Fifteen-microliter aliquots were loaded onto SDS-PAGE gels without any further manipulation.

SDS-PAGE and western blotting

The majority of western blots were performed with the Bio-Rad (Hercules, CA) Criterion XT system. Criterion XT polyacrylamide gels (4-12% Bis-Tris) were used with XT MOPS running buffer for SDS-PAGE at 200 V for one hour. Protein was transferred to PVDF membrane in transfer buffer (3.03g Trizma base, 14.4g Glycine, 20% methanol in 1L, cold) at 100 V for one hour. Gels were stained with Coomassie blue to detect untransferred protein. Membranes were stained with Ponceau S solution [0.1% Ponceau S (w/v), 5.0% acetic acid (v/v)] for five minutes, then rinsed with distilled water to minimize background stain. An image was obtained with a Chemi Genius Bioimaging System (Syngene, Frederick, MD) to quantify relative amounts of protein transferred. Protein amounts were used to normalize samples. GeneSnap and GeneTools software (Syngene) were used for quantification of the predominant band (~73 kDa) in the extracts. When quantifying the images, peak detection was set to “lowest slope” for Ponceau S staining (due to the high

background) and "rolling disk" for the western blots (due to the low background). The membranes were destained in 0.1 M NaOH for one minute, followed by thorough rinsing in distilled water and TBS-T (12.1g Trizma base, 40g NaCl in 5L, pH 7.6 with HCl, with Tween-20 added to 0.1%). The membranes were blocked for one hour with TBS-T plus 5% dried milk and incubated for one hour in TBS-T-milk plus primary antibody (primarily 0.1μg/mL anti-AcV5 in 20mL TBS-T-milk). Membranes were washed five times with TBS-T, incubated for one hour in TBS-T-milk plus secondary antibody (primarily 0.08 μg/mL anti-Mouse in 20mL TBS-T-milk), and washed with TBS-T as before. SuperSignal West Pico chemiluminescent substrate (Pierce) was prepared and poured over the membranes while within the Syngene imaging apparatus. After a five-minute incubation, images were obtained from increasing exposures, in most cases from one minute to one hour, although exposures up to five hours were used for western blots with lower signals. The fusion-GUS product was roughly 70 kDa. For the western blot shown in Figure 11, we used a Kodak (Rochester, NY) digital science Image Station 440 imaging system instead of our standard system. Protein standards used included prestained SDS-PAGE standard (broad-range, 15 μL) and unstained Precision Plus protein standard (2 μL, with 1.2 μL streptactin-HRP added for chemiluminescent detection) from Bio-Rad. For initial western blots, I used the Bio-Rad Ready gel 7.5% Tris-HCl 15-well gels, run at 225 V in SDS Running buffer (15.1g Trizma base, 72.0g Glycine, 5.0g SDS in 5L), and transferred protein to nitrocellulose membrane. Dot-blots were performed by pipeting drops of antibody or protein extract directly onto equilibrated PVDF membrane,

allowing the drop to dry, then continuing with TBS-T-antibody incubations as previously described. Antibodies used in this study included Monoclonal Anti-c-myc (clone 9E10, prepared in mouse, Roche, 0.4 mg/mL), Monoclonal Anti-AcV5 (prepared in mouse, Sigma-Aldrich, St. Louis, MO, 1.4 mg/mL), Anti-β-Glucuronidase (N-terminal, prepared in rabbit, Sigma-Aldrich, 1.7 mg/mL), Monoclonal Anti-Actin (plant) Clone 16-B6 (Mab13a, prepared in mouse, Sigma-Aldrich, 1.0 μg/mL), and ImmunoPure Goat Anti-Mouse IgG (H+L, Peroxidase Conjugated, Pierce, 0.8 mg/mL).

Isolation of homozygous sublines for mutation analysis

Lines which showed moderate to high protein expression and apparent single-transgene insertion sites were selected for further analysis. Eight progeny plants were transplanted to soil and allowed to mature. Each subline plant was harvested and T_2 progeny seed sets of 50 were plated on hygromycin-selective media for segregation analysis. Sublines that showed a homozygous pattern of segregation in the progeny (most or all resistant, one or none sensitive) were selected for testing. After repeated hygromycin-resistance plating to confirm homozygosity, one subline was tested for each selected independent transformant. Four progeny plants were also transplanted to soil and allowed to go to seed for bulk harvesting of the T_3 generation. For some sublines, too few seeds were recovered to permit further testing; bulked T_3 seeds were used instead.

Histochemical staining

Histochemical GUS staining was performed as described [76]. Vegetative plant tissue was collected and immersed in GUS stain buffer [0.1 M sodium phosphate pH 7.0, 0.1% Triton X-100, 0.05% NaN_3, and 0.63 mM X-Gluc substrate (RPI, Mt. Prospect, IL) dissolved in DMF]. Plants were vacuum-infiltrated for 10 minutes and incubated for 48 hours at 37º C. Vacuum infiltration allows for a continuous liquid-leaf surface interface [77]. Plants were treated with 70% ethanol to remove chlorophyll and allow easier detection of GUS stain. The majority of X-Gluc substrate came in the form of CHA salt (f.w. 521.8 g/mol; 33 mg / 100 mL GUS stain buffer). A few of the tests were performed using a sodium salt of X-Gluc (f.w. 498.7 g/mol; 31.86 mg / 100 mL GUS stain buffer). It is not anticipated that the form of salt would cause any significant difference in staining.

UV-C treatment

Each tube of vernalized seeds (suspended in 0.1% agarose) was pipeted randomly over a pot of prepared soil (50 seeds per pot). I divided the pots at the time of planting into 'No UV' and 'UV-C' treatments to avoid bias. An average of 100-150 plants were typically assayed for each data point. Two weeks after planting, UV-C pots were exposed to a single dose of 1000 J/m^2 UV-C (average exposure 6 minutes and 40 seconds of 2.5 J/m^2/sec), then both the UV-C & No UV pots were placed in the dark for 24 hours before returning them to the growth chamber. This was to prevent removal of UV-induced dimers by photolyases, thus increasing time for mutation

fixation. Three weeks after planting, untreated plants were histochemically stained, and UVC-irradiated plants were stained four weeks after planting, due to their growth being stunted by the UV-C treatment. In UV-B experiments, plants were grown as before, except three treatments were spaced at the emergence of the first (7 days), second (12 days), and third to fourth true-leaf pairs (18 days). The UV-B dosage was 5000 J/m^2, with an average exposure of 17 minutes of 4.9 J/m^2/sec.

Heavy metal ion treatment

Sterile, vernalized seeds (an average of 100-150 plants were assayed for most data points) were plated 50 seeds to a plate on standard MS media with or without various concentrations of heavy metal ions (0.4 mg/L or 1 mg/L $CdCl_2$, 6 mg/L or 18 mg/L $ZnCl_2$). The majority of heavy-metal-ion plates were prepared by adding an aqueous solution of heavy metal salt before autoclaving. Some of the 1 mg/L $CdCl_2$ plates were prepared by adding a filter-sterilized aqueous solution of cadmium chloride after autoclaving the medium (noted in figures as 1 mg/L Cd^{2+} post). Seedlings were allowed to grow for three weeks before their vegetative tissues were histochemically stained.

Statistical analysis

Differences in mutation frequency between controls and mutagen treatments were tested for significance with the Mann-Whitney (rank-sum) test. A difference was considered significant if the p-value was less than 0.05. Selected data were presented in histograms and compared to Poisson distributions to confirm that the data was

Poisson distributed and to calculate the best fitted Poisson-derived average. Computations were performed using VassarStats: Statistical Computation Web Site.

Preparation of DNA extracts from plants

DNA was isolated from plants in order to perform southern blots and for PCR amplification for sequence confirmation of constructs. The method I used for DNA extraction was a bead-shaker DNA preparation [78]. Thirty to fifty milligrams of leaf material (equivalent to a three-week-old 10-12 leaf plant), one metal bead, and 400 μL of lysis buffer (10 mM Tris-HCl pH 9.5, 10 mM EDTA, 100 mM KCl, 0.5 M sucrose, 4 mM spermidine, 1.0 mM spermine, 0.1% v/v β-mercaptoethanol, 2% w/v sarkosyl; final pH 9.5) were added to a microfuge tube and mixed with the aid of a reciprocating saw for 30 seconds. Homogenized samples were centrifuged for five minutes and the liquid was transferred to a new tube. DNA was extracted with 300 μL of phenol:chloroform, then precipitated with ethanol (1/10 volume 3M sodium acetate pH 5 and 2.5 volumes 95% ethanol) and washed with 80% ethanol. Air-dried pellets were resuspended in 100 μL TE pH 8.0, 10 μg/mL RNase.

Southern blotting

Plant genomic DNA extracts were digested with *Eco*RV or *Bcl*I restriction endonucleases and separated by agarose-gel electrophoresis. DNA was transferred onto a positively-charged nylon membrane with an alkaline buffer. Transgene DNA was detected using the digoxigenin (DIG) system from Roche Applied Science (Mannheim, Germany). Briefly, the membrane was equilibrated and incubated in DIG

Easy Hyb with DIG-labeled probes at 50º C. The probe-bound membrane was washed, blocked, and treated with anti-digoxigenin-AP. Antibody-bound bands were detected with CPD-Star chemiluminescent substrate and the Chemi Genius Bioimaging System.

Results

Testing and production of mutation constructs

Construction and testing of codon-11 amino acid substitutions

The Kovalchuk/Hohn "codon 112 reporter" refers to a defined base substitution in the codon starting at base 112 of the CaMV ORF V – GUS fusion gene. This codon is the 38th codon in the ORF V – GUS fusion gene and the 11th codon in wild-type *E. coli* GUS. In order to confirm the necessity of the wild-type amino acid glutamate in *E. coli* GUS codon 11, the intended base substitutions were tested for activity, as well as any other amino acids that were potential mutation endpoints. To simplify construction, mutations were made in wild-type *E. coli* GUS lacking the CaMV ORF V fusion.

The first set of codon-11 substitutions constructed included known [40] and presumed positive and negative controls: the wild-type amino acid glutamate (known positive), aspartate substitution (presumed positive), glutamine substitution (known negative), and histidine substitution (presumed negative). These amino acid substitutions were encoded in the mutagenic forward primer. The two reverse primers included a sequencing primer several hundred bases into GUS (*OGUS1* primer), as well as a primer just outside of the GUS coding region suitable for making full length GUS amplifications (*CHK2* primer). pGUS23, generously provided by Igor Kovalchuk, was the template for synthesis [40]. Since the mutation-coding primers

did not extend to the 5' end of GUS, a second round of amplification was performed using the first PCR product as a template. The second forward primer included sequence from the 5' end of GUS as well as the T7 promoter and Kozak sequence, which are essential to the *in vitro* transcription/translation assay. With *CHK2* primer as the reverse primer, the new forward primer allowed efficient amplification of the substitution constructs. The PCR product was utilized in the TNT T7 Quick for PCR (Promega) DNA transcription/translation system. The resulting protein, which was labeled with ^{35}S-methionine, was analyzed for size and concentration with SDS-PAGE and autoradiography; protein activity was measured with the MUG assay.

The initial set of codons tested yielded mixed results. As shown in Figure 6, the GUS construct with glutamate in codon 11 (termed the glutamate TNT construct) yielded an appropriately high signal, and the aspartate and histidine TNT constructs yielded high, partial activity (45%) and basal activity (< 1%), respectively, as expected. However the glutamine TNT construct yielded 27% activity in our hands, as compared to a previous measurement of < 1% activity of a glutamine substitution in *Nicotiana plumbaginifolia* protoplasts by Kovalchuk, et al [40]. We assumed that the discrepancy was due to the difference in expression systems. The previous experiment was presumed more likely to reflect how the constructs would behave in *Arabidopsis* as a plant cell expression system was utilized. Therefore the rest of the amino acid substitutions were generated, and the observed glutamine TNT construct activity of 27% was used as a cutoff. Any amino acid substitution whose product showed less than 27% activity was considered inactive.

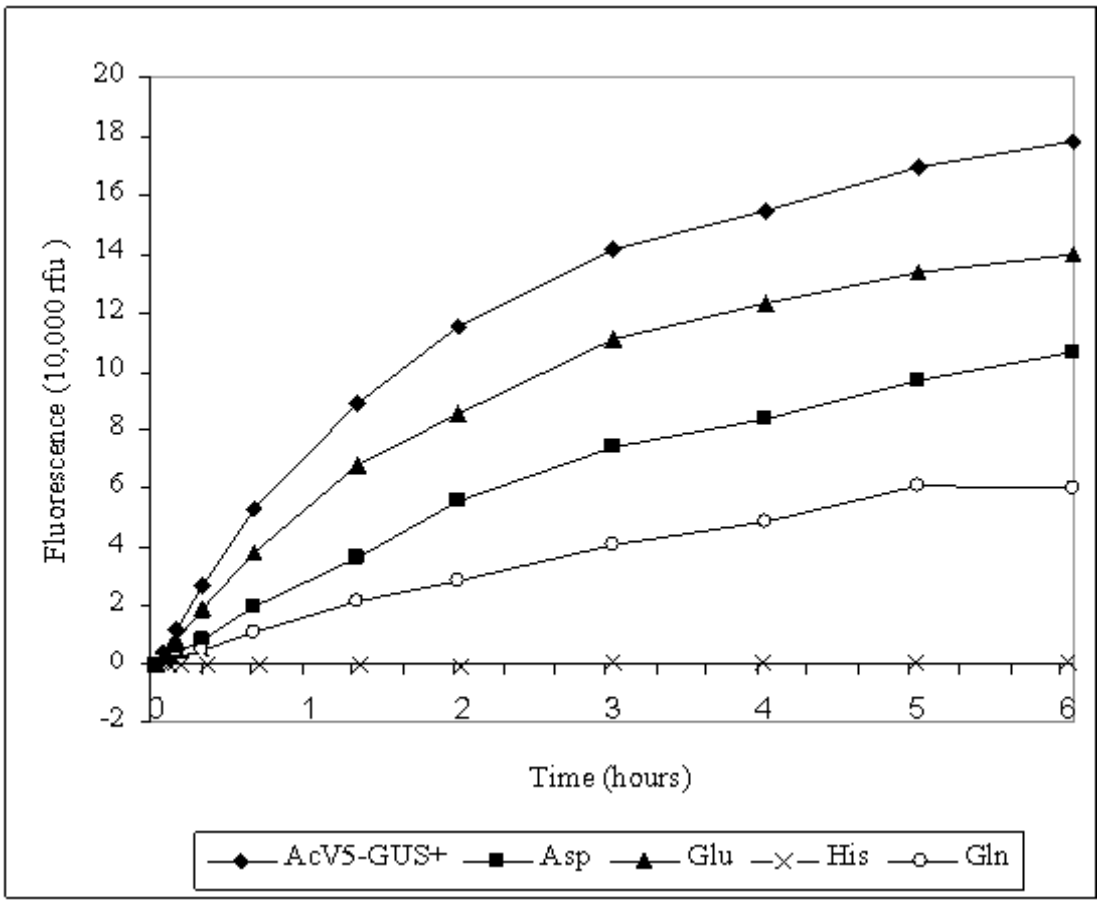


Figure 6. Activity of mutant GUS proteins synthesized by *in vitro* transcription/ translation. Activity was measured by the MUG assay, as described in Materials and Methods, page 54. Fluorescence of MUG after cleavage by GUS was measured in rfu. All data point values were corrected for background signal. Aspartate, glutamate, histidine, and glutamine construct samples were normalized for relative protein level. The positive control consisted of a protein preparation from an *E. coli* culture which contained endogenous GUS.

As shown in Table 2, the 27% activity seen with the glutamine TNT construct was the highest observed among the different substitutions of interest, with the majority showing ≤ 1% activity. The aspartate TNT construct showed higher activity than the other substitutions, but this high activity should not pose a problem for the mutagenesis study. More than one base substitution would be required for any mutant codon to revert to aspartate (for example, G**C**A→G**A**T). All substitution constructs were confirmed by sequence analysis.

Construction of mutation reporters

Table 3 details the altered codons in the mutation reporter constructs. To produce the different GUS substitution constructs for plant transformation, the first step involved PCR amplification with a mutagenic 5' megaprimer. The megaprimer was amplified using a constant forward primer and a mutation-encoding reverse primer. The forward primer contained sequence complementary to the 5' end of the ORF V – GUS fusion gene with an additional HindIII site for cloning. The purified megaprimer was used with a reverse primer that contained sequence complementary to the 3' end of GUS with addition of *Avr*II and *Pst*I restriction endonuclease (AvrII and PstI; other restriction endonucleases similarly abbreviated) sites for cloning. The different PCR products were cloned into TOPO vectors and sequenced. The sequence data indicated additional mutations had been introduced into the different base substitution constructs. Recovery of efficient PCR amplifications for cloning was not optimal, so rather than trying to repeat the process with a more error-free polymerase, the best sequences from different plasmids were combined. There were four critical

Table 2. Determination of activity of mutant GUS proteins synthesized *in vitro*.

Codon 11 amino acid[a]	Codon sequence	Specific activity (%)[b]	Codon 11 amino acid[a]	Codon sequence	Specific activity (%)[b]
Glu (wt)	GAA	(100.0)	His	CAT	0.1
Asp	GAT	45.4	Ile	ATA	1.4
Ala	GCA	7.9	Leu	CTA	1.0
Lys	AAA	1.1	Asn	AAT	1.5
Gly	GGA	8.7	Pro	CCA	0.4
Gln	CAA	27.0	Arg	AGA	0.2
Val	GTA	0.8	Ser	TCA	1.2
Stop	TAA	0.7	Thr	ACA	0.5
Frame-shift[c]	$(A)_8$	0.9	Tyr	TAT	0.04

[a] Proteins containing indicated amino acid substitutions at codon 11.
[b] GUS activities were measured by the MUG assay. Values were normalized for protein level as determined by SDS-PAGE and autoradiography (as described in Materials and Methods, page 54).
[c] Frameshift mutation constructed immediately downstream of codon 11: 5' GAA ATA AAA AAA A 3'

Table 3. Mutation Reporter Constructs.

Line	Amino Acid Substitution	Sequence[a]	Reversion to wild-type	Possible Premutagenic Mismatch
(wt)	(Glutamate)	CGT **GAA** ATC GCA **CTT** TAG		
M1	Stop	CGT **T**AA ATC GCA **A**TT TAG	T → G	T[CPD]**T** A – **C**
M2	Glycine	CGT G**G**A ATC GCA C**C**T TAG	C → T	**A** – A **C**[CPD]T
M3	Glutamine	CGT **C**AA ATC GCA **G**TT TAG	G → C	G:(8-oxoG)
M4	Alanine	CGT G**C**A ATC GCA C**G**T TAG	G → T	A:(8-oxoG)
M5	Valine	CGT G**T**A ATC GCA C**A**T TAG	A → T	A:(2-OH-A)
M6	Lysine	CGT **A**AA ATC GCA **T**TT TAG	T → C	**G** – A **T**[CPD]T

[a] Mutant base-pair bolded in lines M1-M6.

sections: the 5' end with proper cloning sites, the interior 5' end coding for the different mutations, the 3' end with proper cloning sites, and the middle of the GUS coding sequence. The TOPO vector containing the ORF V – GUS fusion gene with alanine at codon 38 (termed the alanine construct) had the correct sequence for the 5' end and a 3' end which was clonable, but not error-free. The stop codon construct had accurate sequence in the mutation-encoding 5' interior, with additional mutations in the other key regions.

The 3' end of GUS was PCR amplified, cloned, sequenced, and inserted into the alanine-construct plasmid using BsgI and AvrII. The accurate 3' end of GUS was also inserted into the stop codon construct plasmid with BsgI and NotI. Both the alanine and stop-codon-construct plasmids were digested with BbsI and BsgI, and purified fragments (TOPO vector sequence and 5' end of ORF V – GUS fusion gene from the alanine construct, majority GUS insert from the stop codon construct) were ligated. The middle constant GUS coding sequence was isolated from pGUS23 with MfeI and BsgI. The product of this cloning strategy contained accurate GUS sequence with the stop codon substitution at codon 38, and this plasmid was used as the intermediate vector for further mutation cloning.

The next step in the cloning strategy was insertion of the coding sequences for epitope tags c-myc and AcV5 at the C-terminus of the intermediate vector GUS gene. The epitope coding sequences consisted of double-stranded oligonucleotides which contained ends that mimicked XbaI- and SacI-restricted ends. To add the epitope sequences to the intermediate vector, the epitope oligonucleotide was ligated into an

XbaI, SacI-digested pUC19 vector, then into plasmid pBlu2SK- with SalI and EcoRI, and finally into the intermediate vector at the 3' end of GUS with AvrII and NotI. This resulted in two intermediate vectors, one with a C-terminal AcV5 epitope tag, one with a c-myc epitope tag.

Synthesis of the other substitution constructs utilized PCR amplification of mutation-encoding inserts. A constant forward primer with proper 5' ORF V – GUS sequence was used with mutagenic reverse primers for mutation insert construction. HindIII and BsmI treatment was followed by insertion of the mutation-encoding inserts into the intermediate vector, as well as a wild-type-insert construct to be used as a positive control for transgene expression. The epitope-tagged mutation-reporter constructs were then inserted into the Gelvin-superpromoter-containing binary vector using restriction endonucleases HindIII and SacI. A diagram of the final constructs is shown in Figure 7.

All novel PCR products and ligation junctions were sequenced. Additionally, the completed binary vectors were fully sequenced from the vector region upstream of the 5' insertion site to the region downstream of the 3' end of the insert. Binary vectors are low-copy-number plasmids, making direct sequencing problematic; therefore our strategy involved PCR amplification to make four overlapping products. Both strands of these products were sequenced, aligned and confirmed (Figure 8). If necessary, additional sequencing runs were performed after initial testing to cover any regions of poor sequence. For example, downstream primer *GUS3.5b*, which was located between primers *GUS3b* and *GUS4b*, was used with *bvup* primer to amplify

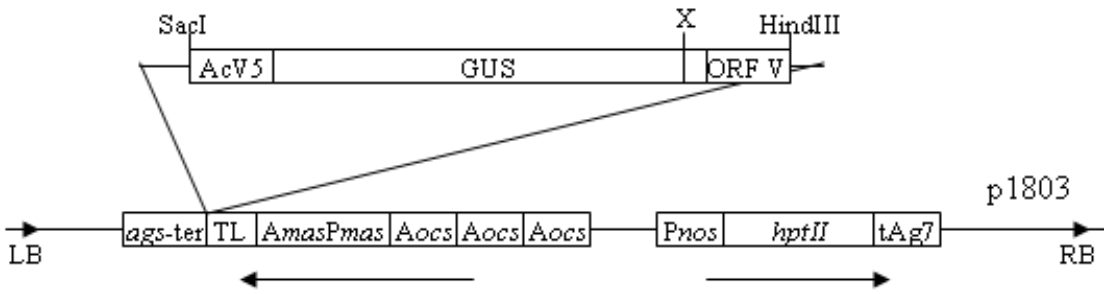


Figure 7. Maps of the T-DNA binary vectors containing the mutant GUS constructs. ORF V, first 29 amino acids from CaMV ORF V; GUS, *uidA* coding sequence from E. coli; AcV5, AcV5 C-terminal epitope; A*ocs*, octopine synthase upstream activation element; A*mas*P*mas*, mannopine synthase promoter and upstream activation element; TL, translational leader from TEV 5'NTR; MCS, multiple cloning site; *ags*-ter, poly(A) addition signal from the agropine synthase gene; Pnos, nos promoter; hptII, gene conferring resistance to hygromycin; tAg7, poly(A) addition signal for T-DNA gene 7; LB, left T-DNA border; RB, right T-DNA border. Arrows beneath the map indicate the direction of transcription. X indicates the site of defined base substitutions.

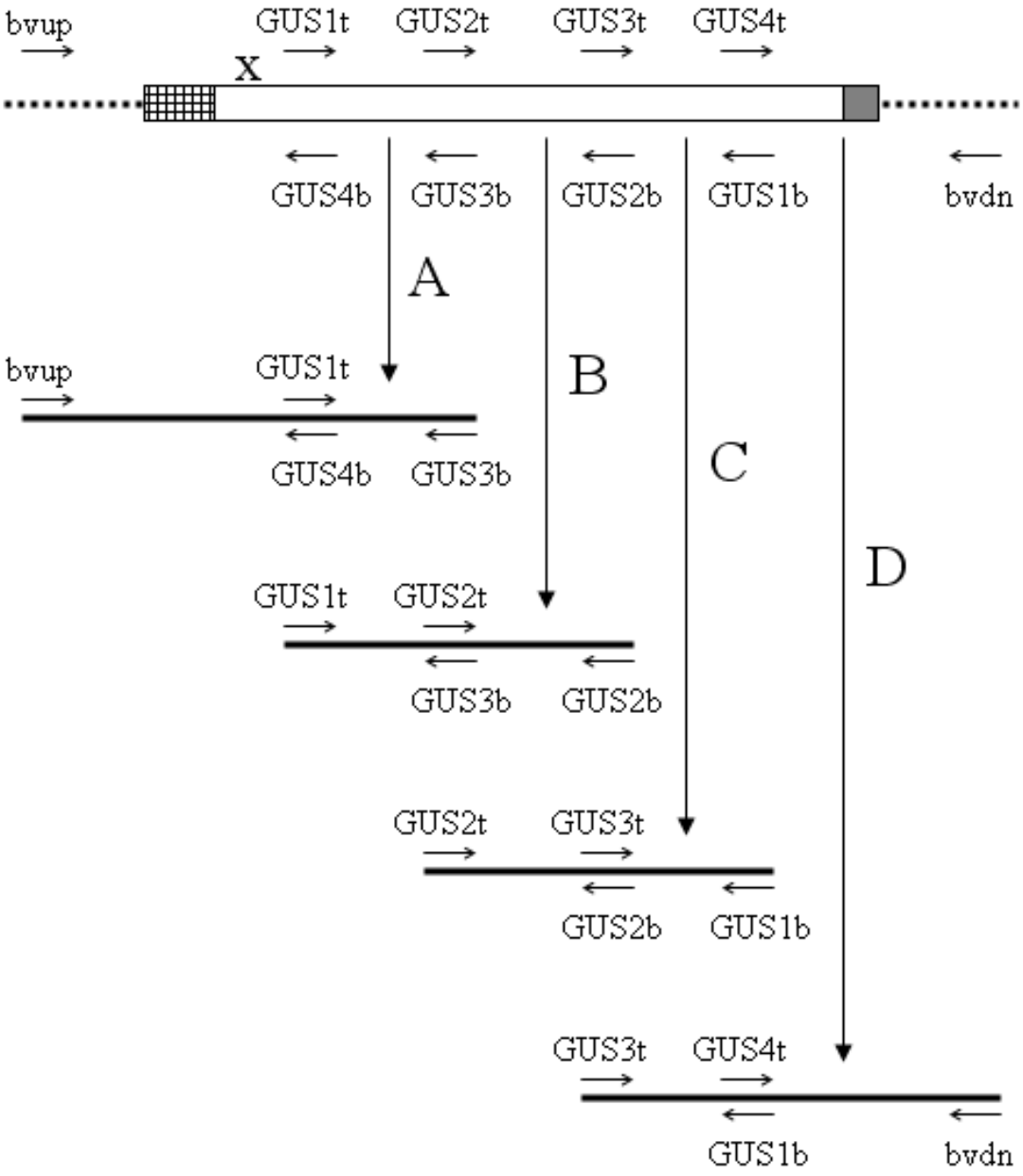


Figure 8. Strategy for sequencing binary-vector constructs. The top diagram refers to the Gelvin superpromoter binary vector with a mutated *GUS* construct inserted into the multiple cloning site. The grid pattern box represents the CaMV ORF V fusion sequence, the white box represents the *E. coli GUS* sequence, the grey box represents the epitope tag (AcV5 or c-myc), the X marks the site of generated mutations, and the relative locations of the various primers used for sequencing and PCR amplification are shown. PCR amplification of smaller fragments with primers (A) *bvup* and *GUS3b*, (B) *GUS1t* and *GUS2b*, (C) *GUS2t* and *GUS1b*, and (D) *GUS3t* and *bvdn* was followed by sequence analysis using indicated primers.

the upstream region of GUS, and *GUS3.5b* primer was used as the sequencing primer. All constructs contained proper sequence with a single silent mutation located 170 base pairs upstream of the SacI site. Not only was the mutation silent, the new codon was in fact better for expression in *Arabidopsis.* The original CGC codon is used 7% vs. 23% for the mutated CGT codon for arginine amino acid coding [79]. An additional site roughly 600 base pairs downstream of the HindIII site showed a mix of two different bases in sequencing reactions. Most encoded the wild-type GGA codon, while a portion favored a silent mutation to GGG; both codons are fully functional in *Arabidopsis* (37% and 12%, respectively).

Four different versions were constructed for each reversion reporter. Two different epitopes (AcV5 and c-myc) were added C-terminally. C-myc is a standard epitope, but its immunoassay has been problematic in plants due to high background signals. AcV5 is reportedly more specific [66], but is a newer epitope and therefore relatively untested. In addition to two epitopes, two binary vectors were used. Both contain the same version of the Gelvin superpromoter (with the additional TEV translational leader) and hygromycin-resistance gene. They differ as to the transcription direction of the superpromoter relative to the resistance gene. p1803 transcribes the transgene and the resistance gene in opposing directions, p1804 transcribes them in the same direction (Figure 2). Our goal, therefore, was the construction of a total of 34 different plant lines: p1803 and p1804 as empty vector controls, and four versions of wild type, alanine, lysine, glycine, glutamine, valine,

stop codon, and frameshift constructs. The p1803-AcV5 set was selected for further testing, and construction of all but clone p1804-alanine-AcV5 was completed.

A total of 33 constructs were confirmed by sequencing and transformed into *Agrobacterium tumefaciens*. Binary-vector-containing *Agrobacterium* strains were used for transformation of Columbia ecotype plants with a floral dip method [74]. Six pots of four plants each were dipped for each construct, and each set of plants was dipped a total of three times. The first dip occurred when secondary bolts were a few inches high, and plants were allowed to recover for a week before the next treatment. Plants were allowed to go to seed, and bulk seeds were harvested to allow for random progeny sampling. Transformed seedlings (T_0) were isolated by growth on hygromycin-selective medium. Transformation efficiencies were typically 0.2 - 0.6%. One-week-old transformed seedlings were transferred to soil and allowed to self-fertilize and go to seed, with each transformant harvested independently. T_1 seeds were plated on hygromycin-selective medium (50 per plate per independent transformant) to observe segregation patterns.

Characterization of plant lines

Isolation of plant protein extracts for western blots

Independent transformants often show varied levels of transgene expression due to multiple insertion or position effects. To ensure sufficient expression for histochemical staining, western blots were used to identify lines with higher protein

expression. Four wild-type constructs were initially tested, and western-blot data was compared to expression as visualized by GUS histochemical staining.

The wild-type ORF V – GUS fusions with either c-myc or AcV5 epitope tags were expressed in *E. coli* to provide protein for positive controls for western blotting. The wild-type constructs were cloned into pET21a expression vectors using HindIII and NotI. These plasmids were transformed into *E. coli* and confirmed by sequence analysis. Expression vectors were transformed into BL21 Codon+ cells and GUS expression was induced with isopropyl β-D-1-thiogalactopyranoside. Protein extractions showing high specific induction were selected for further analysis. Various dilutions of primary and secondary antibodies were tested for activity in western blots. Using antibody conditions that were experimentally defined, various dilutions of positive control extracts were tested with a western blot to determine an appropriate dilution for further studies. Wild-type GUS plant protein samples were then analyzed. Leaves from primary transformants were isolated for protein extraction with a Pierce P-PER® Plant Protein Extraction Kit. The protein extracts yielded no signal on western blots. Control reactions consisting of a mix of positive control extract from *E. coli* and non-transgenic plant extract showed reduced western-blot signal compared to the positive control extract alone (Figure 9), indicating the presence of inhibitory agents in the plant extracts.

Progeny (T_1) of wild-type GUS independent transformants were used for a second method of protein extraction in parallel with histochemical staining. Potential concerns regarding proteases were addressed by freeze-grinding tissue, immediately

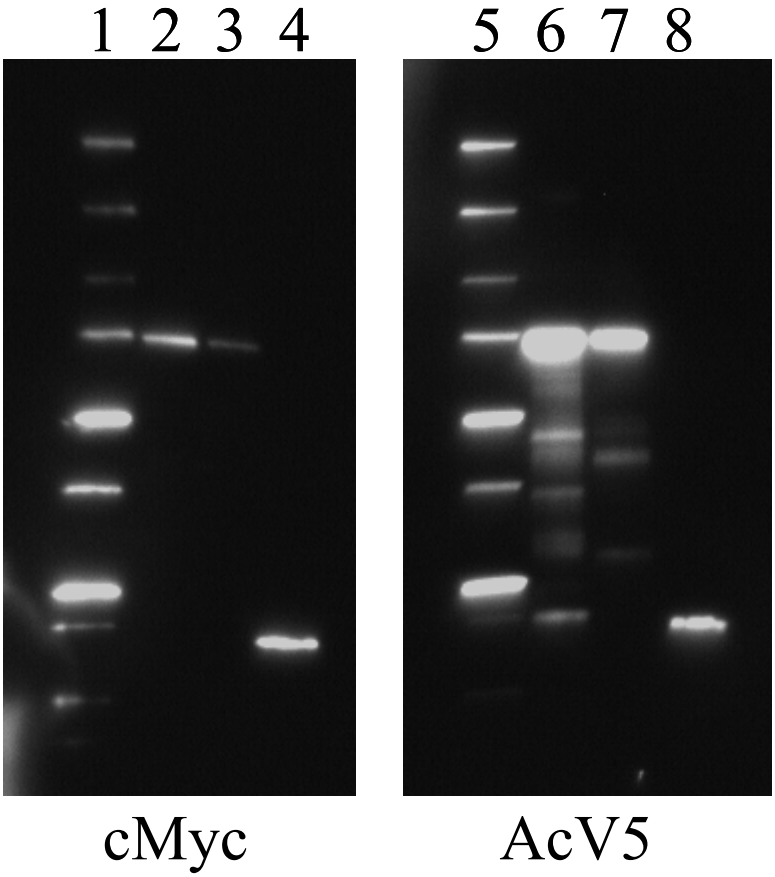


Figure 9. Western-blot signals for AcV5-GUS^{+} protein diluted in plant protein extracts. Lysates of *E. coli* expressing AcV5-GUS^{+} were diluted in either 2X loading buffer (lanes 2 and 6) or a protein extract from Columbia-ecotype plants (lanes 3 and 7). Lanes 1 and 5 contain the chemiluminescent ladder (Precision Plus protein standard from Biorad), and lanes 4 and 8 contain extracts from untransformed DH5α cells, showing a background-staining band. Western blotting with antibodies against indicator epitopes was performed as described in Materials and Methods (page 58).

placing it in 2X protein loading buffer, and boiling it before further processing. This method yielded protein extracts that showed detectable signals on western blots and did not inhibit positive-control extract signals. Levels of GUS protein in extracts from wild-type GUS controls correlated well with activity seen with histochemical staining (Figure 10). The c-myc system proved unsuitable for this work due to limited signal and high background. Attempted detection of expressed protein with a GUS-specific antibody also yielded very low signals. Thus, the AcV5 epitope was selected for detection of the expressed protein.

Quantification of protein levels on the western blots required normalization of the protein extracts for total protein loading. An antibody to plant actin was utilized for an internal control; however this antibody yielded inconsistent results with low signal. These barely-detectable actin signals were used to analyze the effectiveness of Ponceau S staining of proteins blotted onto PVDF membranes for protein quantification. Protein concentrations were calculated from measurements of the same blot. Actin signals from two repetitions of four plant extracts were quantified and normalized to the highest value. The most intensely staining band from each of the three repetitions of Ponceau S stains of the same plant extracts was quantified and normalized to the highest value. The actin and Ponceau S protein quantifications were comparable (Figure 11), and Ponceau S staining was utilized for further western blot analyses. Ponceau S staining provides a simpler, more reliable measure of total protein levels. To confirm the accuracy of the Ponceau S stain, a wild-type plant extract (p1804-glutamate-AcV5) was used to produce a standard curve. As seen in

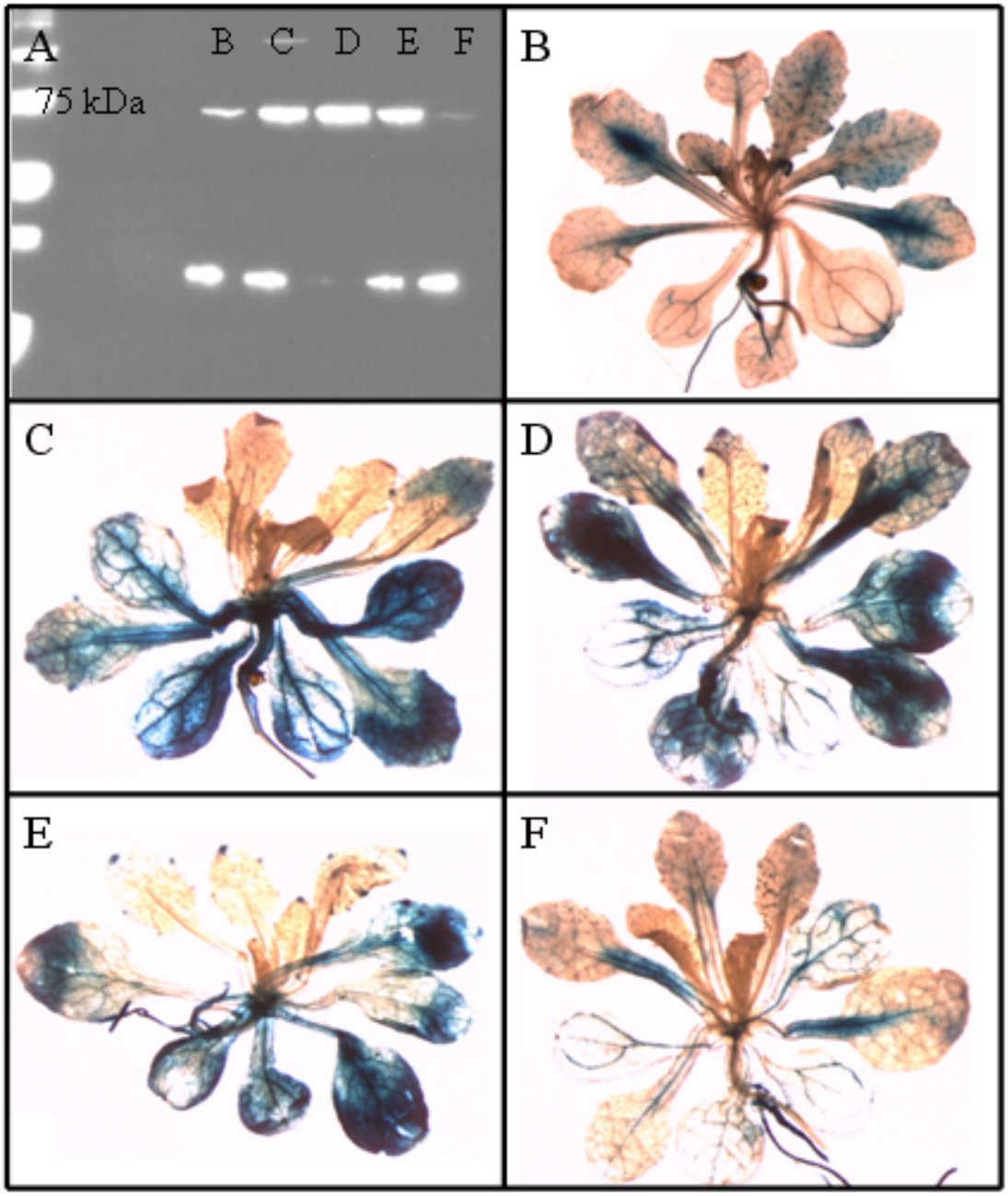
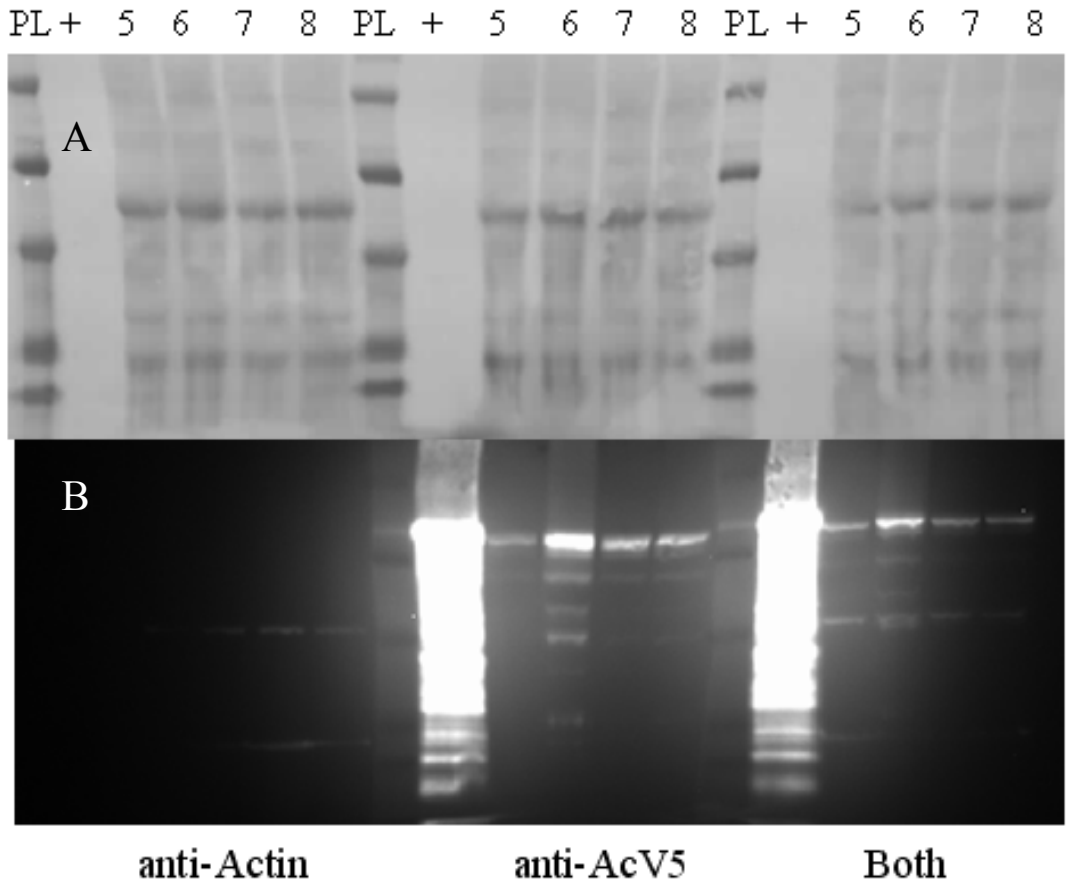


Figure 10. Analysis of AcV5-GUS^{+} expression by western blotting and histochemical staining. (A) Representative western blot of protein extract from plants expressing AcV5-GUS^{+}. The upper band is AcV5-tagged GUS, and the lower band is a nonspecific band presumably bound by the secondary antibody. (B-F) Representative histochemical staining patterns. Lines were (B) W1803-AcV5- 2, (C-F) W1804-AcV5-1, 2, 3, and 4, respectively. Younger leaves tend to be in the top of the picture, and more mature leaves are at the bottom.

Figure 11. Comparison of staining of total protein with actin antibody and Ponceau S. (A) Ponceau S stain image used for quantification. (B) Western blot image used for quantification. Wild-type GUS samples were analyzed in triplicate to test three indicated primary-antibody solutions: anti-actin only, anti-AcV5 only, and a mixture of anti-actin and anti-AcV5. Antibodies were tested alone and together to test for interference. PL: protein ladder (Biorad Broad-Range Prestained standard); +: lysate of *E. coli* expressing AcV5-GUS^{+}; 5-8: protein extracts from W1804-AcV5 plant lines 5-8. Faint bands in the middle of the blot correspond to Actin. (C) Final quantification of relative total protein levels based on either Actin signals or Ponceau S stain. Ponceau S stain and western blot performed as described in Materials and Methods (page 58). Relative total protein levels calculated by normalizing peak height of the Actin protein band or the predominant Ponceau S staining band to the largest signal observed for each (sample W1804-AcV5-6). Bands were quantified using GeneSnap and GeneTools software.



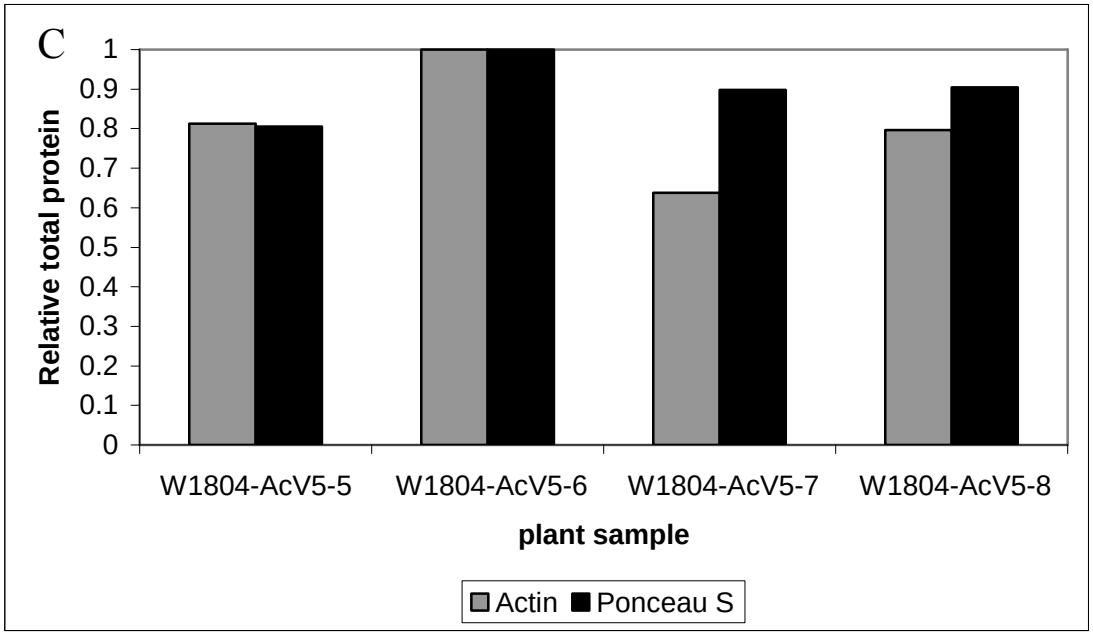


Figure 11. Comparison of staining of total protein with actin antibody and Ponceau S.

Figure 12, normalization by Ponceau S stain appeared to be reliable and reproducible, and the majority of normalized sample measurements fell within the linear range.

Isolation of plant mutation-reporter lines

Two selection criteria were used for selecting transformed plant lines for further testing: high protein expression, measured by western blots, and single-locus insertions as indicated by segregation analysis. Seeds (T_1) from independent T_0 transformants were plated on hygromycin-selective medium and plants were grown ten days. Numbers of resistant, sensitive, and non-germinating seedlings defined phenotypic segregations for the transgene. Progeny showing a ratio of three resistant plants to one sensitive plant indicate a hemizygous parent with a single site of transgene insertion (although this does not rule out multiple copies at that locus). Progeny sets with a higher ratio indicate multiple sites of insertion; for example two independent sites would show a segregation ratio of ~15:1. Segregation data was analyzed by the Chi-square test with a cut-off value of $p > 0.05$.

The ten day-old seedlings were tested for protein-expression levels. Three seedlings (four- to six-leaf stage) were chosen at random and used for protein extraction and western blotting as described (page 57, 58). A representative set is shown in Figure 13. For a given blot, extracts were normalized with Ponceau S stain as described (page 58) for total protein level. AcV5 antibody signals on the western blot were expressed relative to the strongest signal, unless the highest signal was an extreme outlier, in which case the second highest signal was used for normalization. Relative western-blot signals were normalized for total protein levels. Lines with

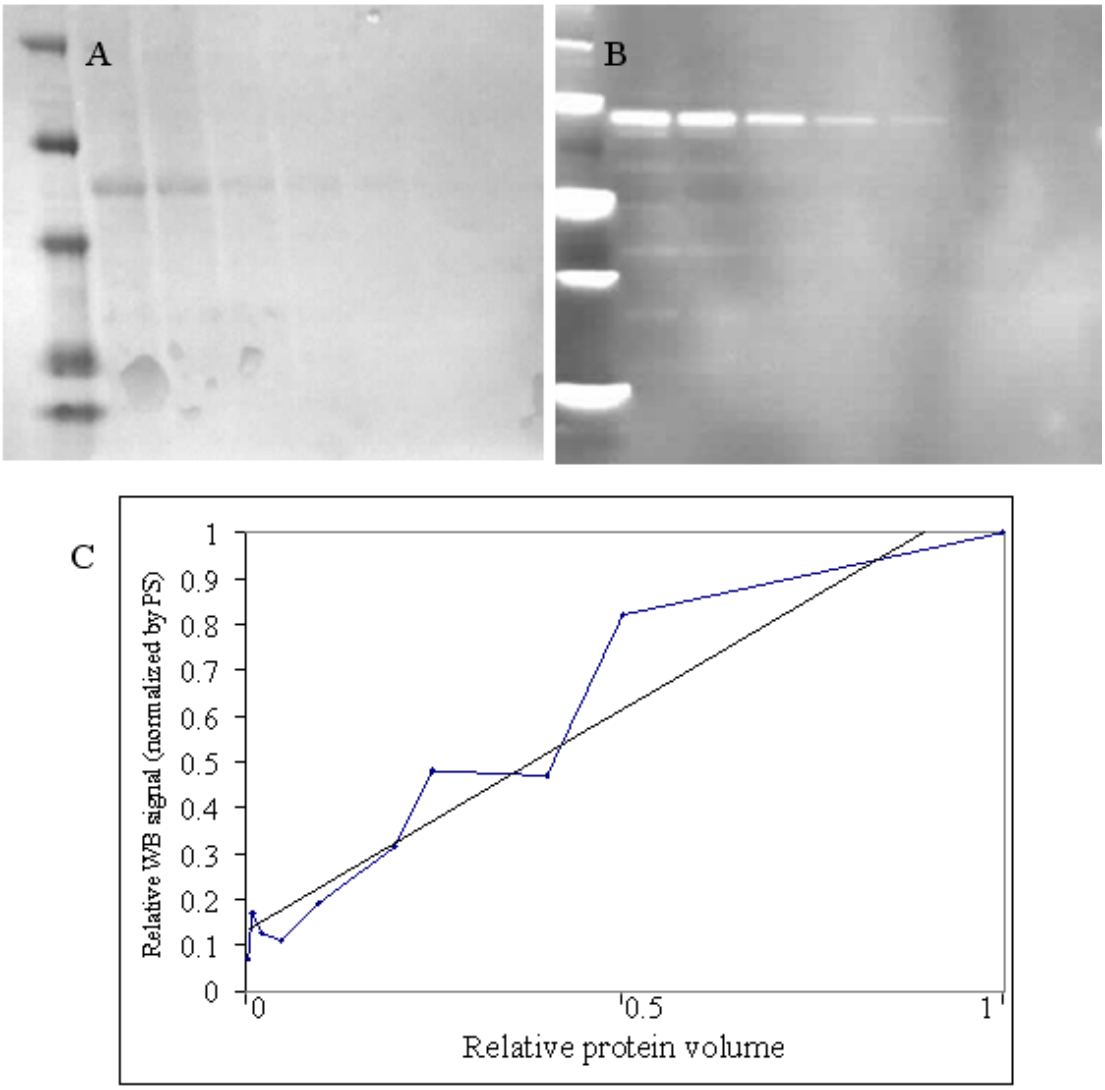


Figure 12. Standard curve for Ponceau S staining. Successive dilutions (1:1 to 1:320, right to left in panel C) of an extract of plants expressing W1804-AcV5 protein were analyzed by Ponceau S staining (PS) and compared to anti-AcV5 western-blot (WB) signals. (A) Ponceau S stain image, (B) western blot image, (C) relative height of AcV5 western-blot chemiluminescent signal normalized by Ponceau S stain compared against relative protein volume. Ponceau S stain and western blot were performed as described in Materials and Methods (page 58). Relative western blot signal was determined by dividing the signal intensities of the AcV5-GUS bands (relative to the highest signal) by the total protein level calculated by Ponceau S staining (relative to the highest signal). Peak heights in this experiment were normalized to the initial dilution.

Figure 13. Western-blot analysis of GUS expression in plants. (A) Representative western blot of plant protein extracts from independent sublines. Lane 1 (CL) contains the chemiluminescent ladder, lane 2 (PS) contains the AcV5-GUS^{+} protein standard, and lanes 3-18 contain samples listed in the same order in (B). (B) Final quantification of transgene expression for tested lines. The dashed grey line indicates the cut-off chosen for selecting sublines for further analyses. X indicates a subline which segregation analysis suggested to contain multiple transgene insertion loci, and was therefore not used for further testing. Ponceau S staining and western blot were performed and quantified as described in Figure 12 legend.

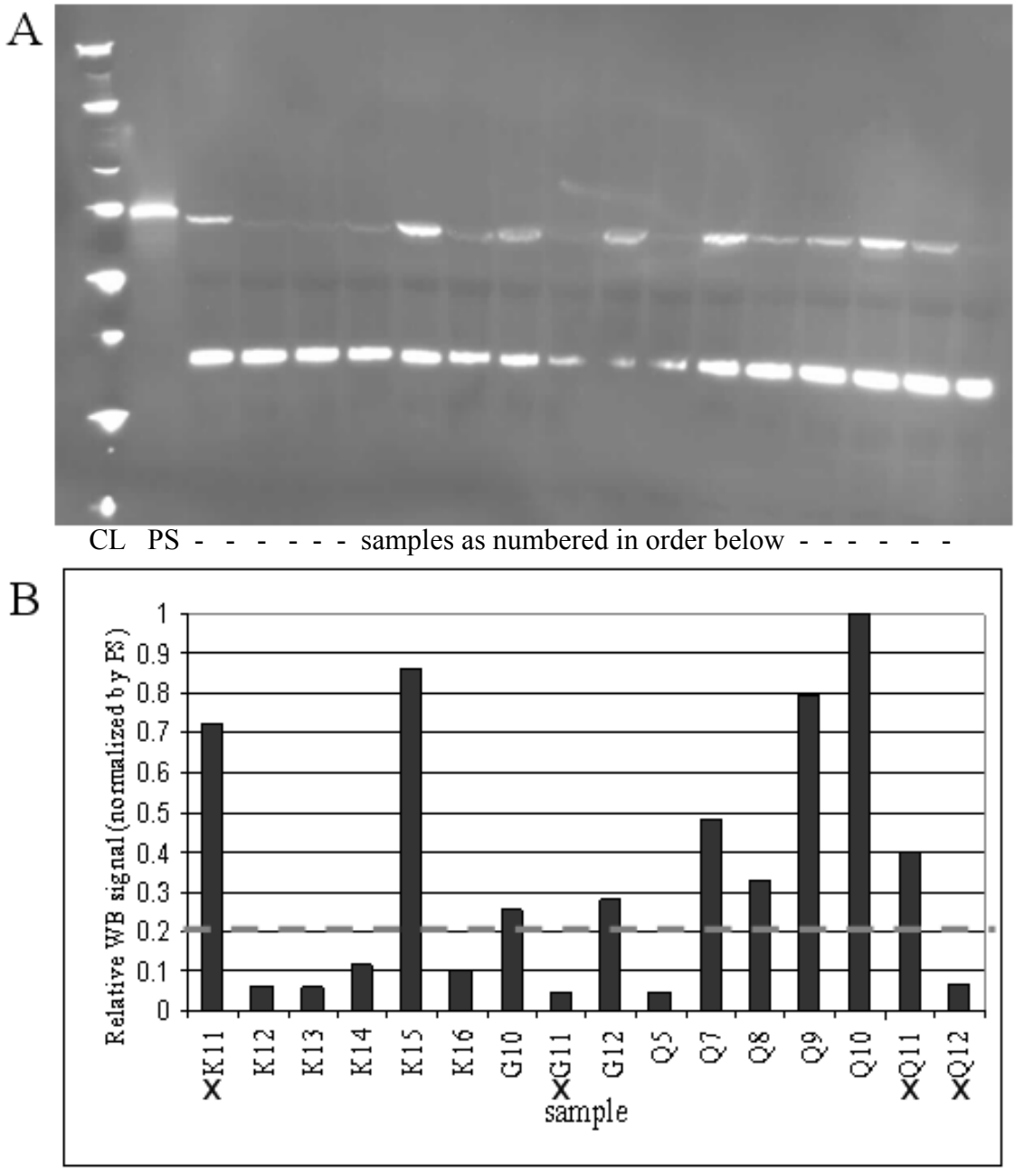


Figure 13. Western-blot analysis of GUS expression in plants.

levels of expression higher than 20% relative to the maximum line, as well as an apparent single locus of transgene insertion, were selected for further testing. Eight progeny plants from each selected line were transplanted to soil for independent seed collection and isolation of homozygous sublines.

Although the transgene is able to function in a hemizygous state, isolation of homozygous sublines simplifies mutation measurements and removes the need for continuous selection. Single transgene locus lines were isolated for this purpose. Southern blots remain to be performed, in order to determine copy number. To identify homozygous sublines, T_2 seeds were harvested separately from each subline and plated onto hygromycin-selective medium. Progeny sets which contained only resistant seedlings were transplanted to soil for bulk harvesting. In the case where more than one homozygous subline was isolated for a given line, the subline with the highest seed yield and best resistant: sensitive seedling ratio was chosen for further testing. These T_2 seeds were used for initial characterization of spontaneous, UVC-induced, and heavy-metal-ion-induced reversion.

Expression patterns of wild-type constructs

To study the levels and patterns of expression of transgenes under the control of the Gelvin superpromoter, a wild-type *GUS* gene fused to DNA sequence encoding a C-terminal AcV5 epitope tag was cloned into binary vectors p1803 and p1804. These vectors include three copies of the octopine synthase activator, followed by the mannopine synthase activator and promoter as well as the TEV translational leader 5' to the inserted gene and a terminator 3' to the gene (Figure 2). The GUS reporter

allowed easy visualization of expression patterns, and the epitope tag aided in quantitative measurements of expression levels among independent lines. Protein lysates were isolated from ten day-old seedlings (three per sample) grown on MS plates, and western blots were performed to compare overall expression between independent lines. Relative protein levels among the various lines correlated well with histochemical staining.

Based on the original characterization of the Gelvin superpromoter and follow-up experiments from the Gelvin lab, we anticipated either preferential expression in older tissue, as in tobacco, or high expression levels throughout the entire plant, as in maize [56]. High expression was previously indicated by mRNA presence in different plant tissues; mRNA was observed at levels higher than in a 35S-promoter-driven GUS transformant [55]. Here, three plants each were stained at several developmental stages: two leaves (cotyledons only), four leaves, six leaves, eight leaves, one week past eight leaves (short bolts), two weeks past eight leaves (several inch-tall bolts with flowers), and three weeks past eight leaves (full-size bolts with siliques). Similarly to tobacco, almost every independent line studied showed variable expression among different plant tissues, the most predominant pattern being expression in the most mature tissues (e.g. cotyledons). Staining patterns observed at later times revealed tissues in younger plants that showed little or no GUS stain developed increased expression in older plants (Figure 14). This developmental pattern of expression minimized concerns regarding transgene silencing during plant growth, an effect that has been observed when using a strong promoter. Other expression patterns were

Figure 14. Staining of wild-type GUS protein expressed during development. Plants from line W1804-AcV5-1 were stained as described in Materials and Methods (page 61) at various stages (from 2-leaf stage, upper left to 8- to 10-leaf stage, bottom) to view GUS expression. c: cotyledons; 1: 1st true leaves; 2: 2nd true leaves; 3: 3rd true leaves.

observed in the wild-type lines, but we speculate that variations in this pattern were due to position effects (insertion into the *Arabidopsis* genome near to and influenced by tissue-specific enhancers). Although previous studies indicated no position effects in *Arabidopsis* containing single inserts, these previous studies analyzed whole plants or leaf samples, not specific tissues. Thus there may be variation among tissues even though a high level of expression is maintained. To confirm that the expression pattern was due to the superpromoter and not an artifact of our staining technique, we stained seedlings containing a 35S-promoter-driven GUS transgene and observed high levels of blue stain throughout the entire plant. Therefore the promoter elements in the Gelvin superpromoter direct expression to certain tissues during development. The promoter region induced high expression, but not ubiquitously. We ruled out endogenous GUS activity as a factor, since histochemical stain with a buffer of pH 4 (as opposed to pH 7) was required to view endogenous GUS activity (light blue all-over stain with darker stain in stems). Additionally, staining patterns varied not only among leaves, but also within a leaf (Figure 15). A common leaf-expression pattern was staining primarily within vascular tissue or staining limited to a subsection of the leaf.

Testing of mutant plant lines with or without mutagens

Kovalchuk/Hohn point-mutation reporters

To verify our ability to detect and quantify reversions, we repeated a set of spontaneous and UV-induced mutation studies with the Kovalchuk/Hohn point

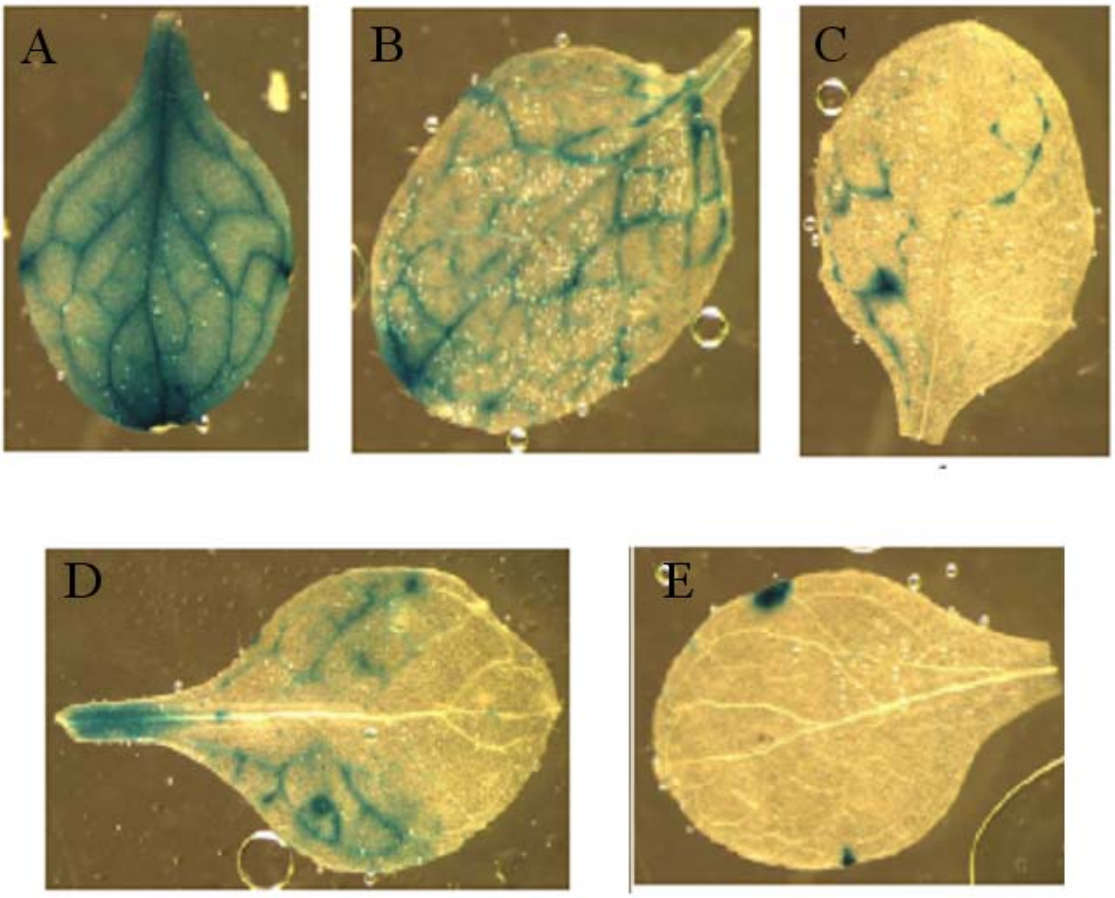


Figure 15. Examples of GUS staining patterns in wild-type-construct lines. Leaves from independent transformants were histochemically stained to observe transgene expression and activity. (A) Full leaf staining pattern. Vascular tissue is more intensely stained. (B) Vascular specific staining pattern, (C) partial vascular staining pattern, (D) paired, sectored vascular staining pattern, (E) paired spots. (A-C) W1803-c-myc lines, (D) W1804-AcV5 line, (E) W1804-c-myc line.

mutation lines 166 (C<u>T</u> → C<u>C</u>) subline 1 and various sublines of line 166 (T<u>T</u> → T<u>G</u>) [40]. We detected reversion event frequencies similar to previous measurements (Table 1, 4), but some of the reporter lines showed generational instability in mutation frequency. Examples include spontaneous mutation frequencies in line 166 (T → G) subline 20 on MS medium (0.01 observed events per plant versus 0.14 reported events per plant) and line 166 (T → G) subline 4 on soil (0.24 observed events per plant versus 0.03 reported events per plant) (Table 4, rows 5 and 14). The observed mutation frequencies were obtained from seeds one to two generations removed from the generation reported by Kovalchuk, et al [40]. It is unknown whether the instability we observed is due to differences among the line generations or differences in testing conditions between the two laboratories. Generational instability was not systematically addressed by Kovalchuk, et al [40].

The most dramatic result was the large discrepancy in mutation frequency between soil and MS media growth. Spontaneous mutation and UV-specific-mutation frequencies were both higher in soil. This difference between soil and MS media growth was corroborated by later results with our T→G lines, which only showed UV-induced reversion when grown on soil (Figure 16 and data not shown). Due to these results, further UV mutagenesis studies were performed on soil-grown plants with 1000 J/m^2 UV-C. Another difference was observed in line 166 (C<u>T</u> → C<u>C</u>) subline 1 grown on heavy-metal-ion-supplemented plates. Although UV-specific reversion was similar to reversion frequencies previously reported (Table 4, experiment 5, row 2) [40], there was a reduction in response to heavy metal ions compared to previously

Table 4. Reversion of Kovalchuk/Hohn mutation-reporter lines.

Exper-iment	Line	Growth medium	Mutagen treatment	Mutation frequency[e]	Mutagen-specific MF	Mutation ratio[f]
1	$166_{T\rightarrow G}$ 2-5	MS		0.02 (0.02)[a]		
	$166_{T\rightarrow G}$ 4	MS		0.02 (0.03)[a]		
	$166_{T\rightarrow G}$ 18	MS		0.05 (0.12)[a]		
	$166_{T\rightarrow G}$ 20	MS		0.01 (0.14)[a]		
	$166_{T\rightarrow G}$ 22	MS		0.06		
	$166_{T\rightarrow C}$ 1	MS		0.19 (0.02)[a]		
2[b]	$166_{T\rightarrow C}$ 1	MS		0.12 (0.02)[a]		
	$166_{T\rightarrow C}$ 1	MS	1 mg/L $CdCl_2$[c]	0.05	-	-
	$166_{T\rightarrow C}$ 1	MS	1 mg/L $CdCl_2$[d]	0.15 (0.7)[a]	0.03 (0.68)[a]	1.2 (35)[a]
3	$166_{T\rightarrow G}$ 4	MS		0.04		
	$166_{T\rightarrow G}$ 4	MS	500 J /m^2 UV-C	0.08	0.04	2
	$166_{T\rightarrow G}$ 4	MS	1000 J/m^2 UV-C	0.13	0.09	3
4	$166_{T\rightarrow G}$ 4	Soil		0.24 (0.03)[a]		
	$166_{T\rightarrow G}$ 4	Soil	500 J/m^2 UV-C	1.20	0.96	5
	$166_{T\rightarrow G}$ 4	Soil	1000 J/m^2 UV-C	1.10 (0.53)[a]	0.86 (0.50)[a]	4.5 (17.8)[a]
5[b]	$166_{T\rightarrow C}$ 1	Soil		0.04 (0.02)[a]		
	$166_{T\rightarrow C}$ 1	Soil	1000 J/m^2 UV-C	0.30 (0.58)[a]	0.26 (0.56)[a]	6.9 (29)[a]

[a] Values in parentheses were reported by Kovalchuk, et al [10, 37].
[b] Bulked seeds from the generation following initial analysis in Experiment 1.
[c] Cadmium chloride added before autoclaving.
[d] Cadmium chloride added after autoclaving.
[e] Mutation frequency (MF) measured as events per plant.
[f] Mutagen-specific mutation frequency divided by spontaneous frequency.

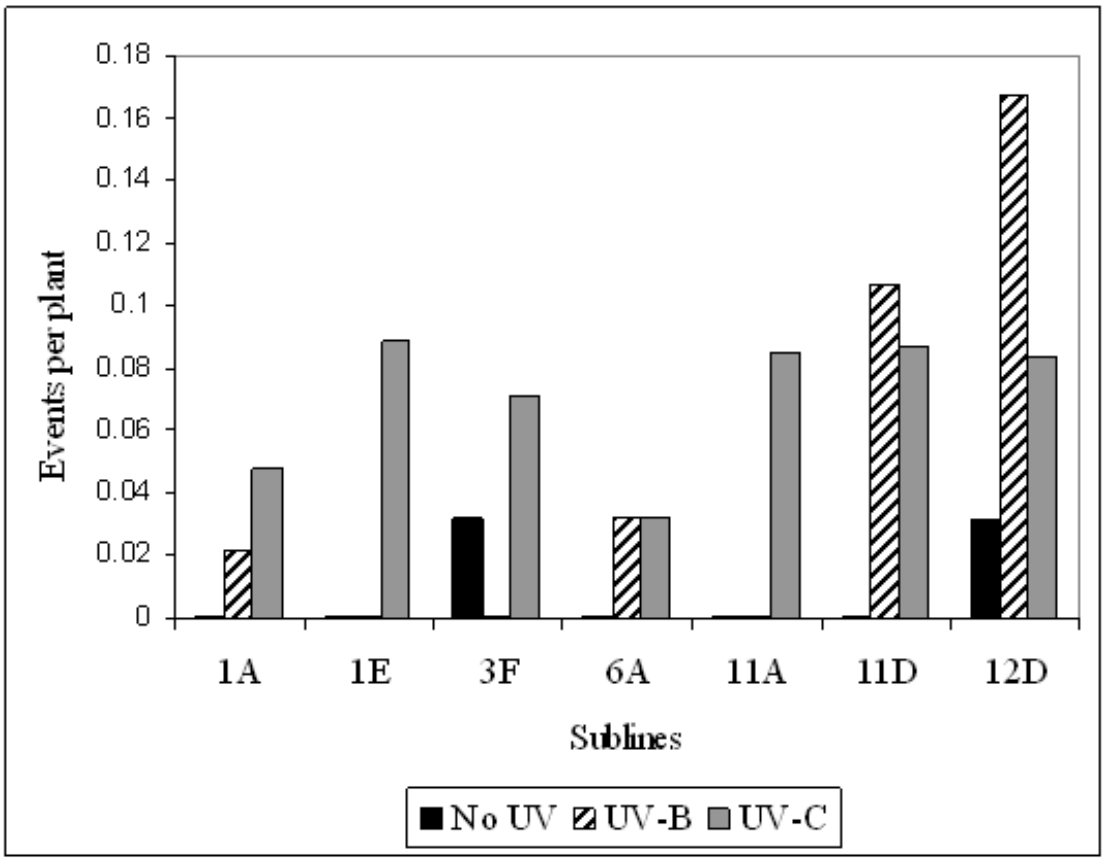


Figure 16. UV mutagenesis in M1 (TT → TG) sublines. Plants were grown on soil, under continuous light. UV-B treatments of 5000 J/m^2 were repeated three times, at the emergence of the first, second, and third pairs of true leaves. Alternatively, plants were exposed to a single dose of 1000 J/m^2 UV-C at two weeks after planting. Plants were placed in the dark for 24 hours after treatment. UV-induction of mutation was not statistically significant ($p < 0.05$), because low numbers of plants were tested in this initial screen.

reported frequencies (Table 4, experiment 2, row 3) [9]. The observed reversion frequency on MS medium supplemented with 1 mg/L $CdCl_2$ was 0.15 events per plant (1.2-fold increase over spontaneous) compared to the reported reversion frequency of 0.70 events per plant (35-fold increase over spontaneous). Assay conditions were recreated as closely as possible to reported conditions, and we are unable to explain the discrepancy.

Analysis of mutation events

Accurate interpretation of mutational events took into consideration transgene expression patterns and plant development. In wild-type GUS control plants, protein expression was not uniform, but instead tended to occur in more mature tissue, starting at the leaf tip and spreading toward the stem (Figure 14). However, other patterns of transgene expression were also observed. Plant development was another factor in interpretation of mutation events. If a mutation occurred in a leaf that had already stopped cell division and was growing only by expansion and endoreduplication, one would expect the mutation event to be confined to a single spot or small sector. If a mutation occurred in a leaf that had not finished cell division, perhaps while emerging from the meristem, one would expect to see different patterns, including a large sector or entire leaf, two sectors with bilateral symmetry, or even multiple "random" spots across a leaf or within the leaf veins, dependent on non-uniform expression across the leaf. We observed all of these patterns. An early mutation in the meristem could also result in novel patterns, such as identical patterns on multiple leaf stems, or identically patterned events on all newly developing leaves. To make conservative estimates of

mutation frequencies, we defined all of the following to be single events: a single spot or sector in either a leaf or stem; a bilaterally symmetrical pair of spots or sectors on a single leaf; closely adjacent multiple spots or sectors on a single leaf; an identical pattern in all new leaves emerging from the meristem; staining that is within the sub-apical meristem. Representative events are shown in Figure 17.

For calculating mutation frequencies, two different methods were employed. First, mutation frequencies were measured as total events observed per total plants tested. The importance of measuring events per plant rather than per cell division stems from the variability of plant size in the different tests, especially in the UVC-treated plants. When these plants were isolated for GUS staining, there was a large amount of variation in plant sizes, the largest being eight to ten leaves, and the smallest containing only the cotyledons. This is due to occasional crowding in the pot and growth inhibition caused by UV-C treatment. Plant growth was stunted about one week in UVC-irradiated plants and the most exposed sections of leaves appeared necrotic. Therefore, estimates of cell division through quantification of total DNA would be too variable to be useful. Less variation was seen in the plants tested on plates with or without heavy-metal-ion treatment. Control conditions (soil or MS plates) were analyzed in parallel with UV-C or heavy-metal-ion treated plants. The second method for calculating mutation frequencies used the frequency of plants with zero events to calculate the Poisson-distribution-derived average. The average number of events (x) was calculated as $x = -\ln[(\#\text{ plants with zero events})/(\text{total }\#\text{ plants tested})]$. This method provided results similar to those from calculation of events per

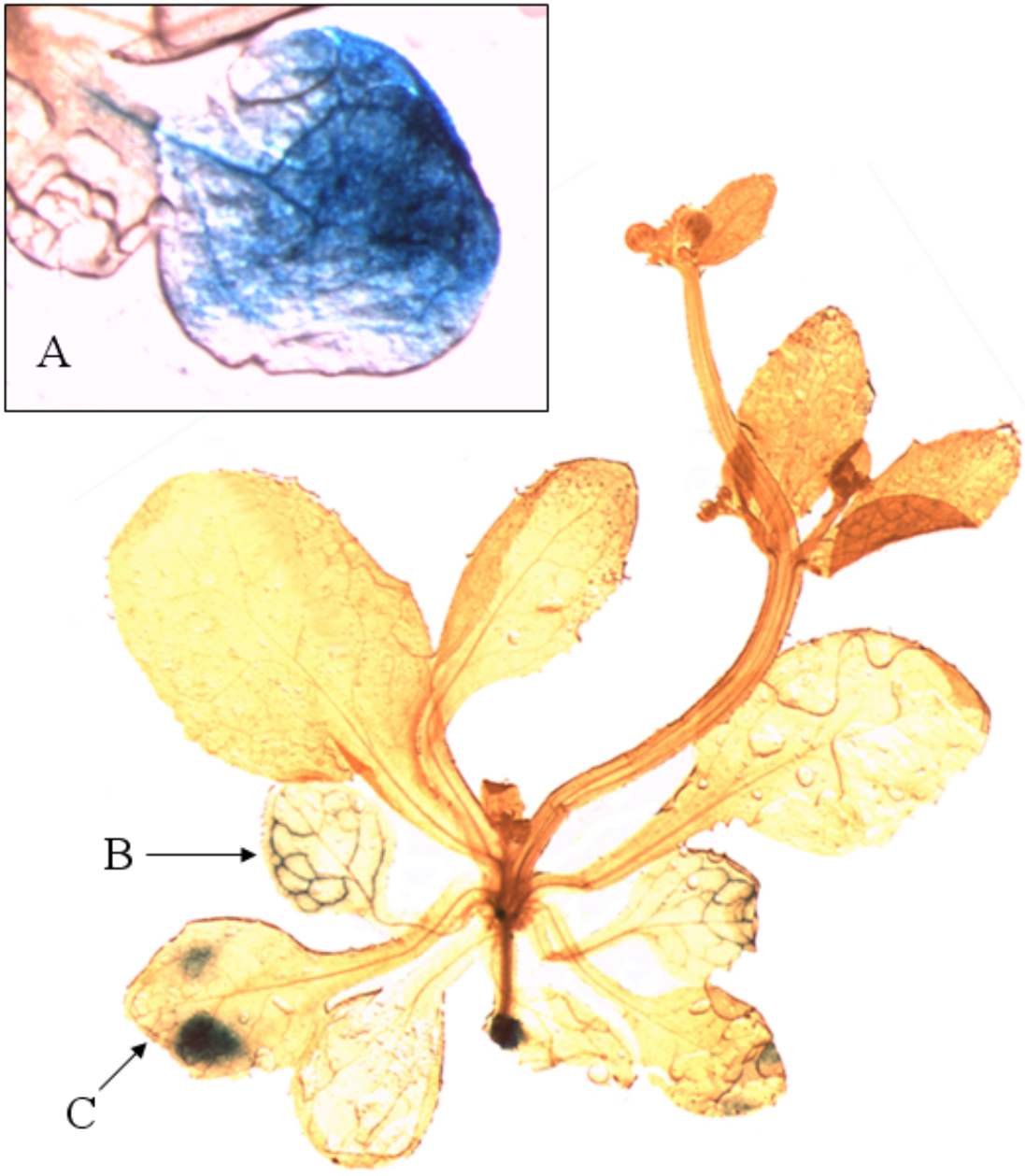


Figure 17. Staining in cells and tissue containing GUS reversion. Selected examples of reversion- event patterns. (A) Large isolated spot; (B) vascular-specific sector; (C) paired spots.

plant (Table 5), which provides further evidence of the data being Poisson distributed. Mutation frequencies were also calculated from the best-fit Poisson distribution to the data for representative data sets. Figure 18 shows representative histograms of data sets compared with fitted-Poisson distributions. Frequencies calculated using either the zero-term or fitted-Poisson distributions were comparable, due to the majority of the data points being zeroes. Both Poisson calculations were the same or slightly less than mutation frequencies calculated by total events per plant (Table 6), and Poisson-derived values were never less than 38% of the event-calculated mean. The Poisson calculation is useful for lines that show a high frequency of reversion. With the Poisson method, there is no need to make a determination whether a complex pattern is one or multiple events, because any event would not allow a plant to be counted as a "zero". The discrimination between an event and no event also controls for outliers and "jackpot" effects, which are caused by early-mutations, since a plant with multiple events would count the same as a plant with a single event.

Spontaneous Mutation

Plants grown on standard MS plates or soil were assayed for spontaneous levels of mutation to provide a background for comparison to mutagen treatments. Untreated plants also indicated levels of endogenous damage, such as ROS-induced mutation. p1803 construct lines that showed reversion of G:C base pairs showed higher spontaneous mutation frequencies than constructs showing A:T base pair reversion (Table 5, Figure 19). G:C constructs showed an average of 0.45 events per plant or 0.04 events per plant when line M4 (G → T) was not included. In

Table 5A. Spontaneous and UV-induced mutation frequencies.

	Mutant Frequency (events/plant)[b]			Mutant Frequency (Poisson average)[c]		
Subline[a]	Spontaneous frequency	UV-specific frequency	Ratio (UV-specific/ Spontaneous)	Spontaneous frequency	UV-specific frequency	Ratio (UV-specific/ Spontaneous)
M1(T→G)						
11D	<0.02	0.09	>4.1	<0.02	0.09	>4.2
11A	<0.02	0.09	>3.9	<0.02	0.09	>4.0
1E	<0.02	0.09	>3.8	<0.02	0.07	>2.9
12D	0.03	0.05	2.7	0.03	0.06	2.7
1A	<0.02	0.05	>2.3	<0.02	0.05	>2.4
3F	0.03	0.04	2.2	0.03	0.04	2.3
6A[i]	<0.03	0.03	>1.1	<0.03	0.03	>1.1
M2(C→T)[g]						
10C	0.007[f]	0.42	61	0.007[f]	0.36	53
20B	0.05	0.20	5.1	0.05	0.18	4.6
12G T_2	0.008	0.15	20	0.008	0.13	19
12G T_3	0.12	0.06	1.5	0.11	0.02	1.2
	0.02[f]	0.39[f]	18[f]	0.02[f]	0.24[f]	12[f]
8D[f]	0.04	0.15	5.2	0.04	0.14	5.1
15G T_3[i]	0.02	0.14	8.2	0.01	0.17	18
p1804-3G[f]	0.17	0.51	4.1	0.12	0.49	5.0
M5(A→T)[g]						
12D T_2	0.05	0.09	2.7	0.05	0.10	2.8
12D T_3[i]	0.06	0.09	2.4	0.07	0.07	2.0
M6(T→C)[f]						
18F T_2	<0.01	1.04	>127	<0.01	0.92	>111
18F T_3	<0.01	0.05	>5.2 [g]	<0.01	0.05	>5.3
	<0.01[g]	0.16[g]	>16	<0.01[g]	0.12[g]	>12[g]
18G T_2	<0.03	0.24	>9.5	<0.03	0.16	>6.1
20C T_2	<0.01	0.60	>76	<0.01	0.51	>64
20C T_3	0.01	0.08	8.2	0.01	0.07	7.4
22B	<0.02	0.33	>16	<0.02	0.36	>17
15F T_2	0.005	0.11	23	0.005	0.11	20
		0.09[gh]			0.10[gh]	
15F T_3	<0.01	0.16	>15	<0.01	0.16	>15
24B	<0.01	0.11	>7.7	<0.01	0.10	>6.7
21C	<0.01	0.05	>6.0	<0.01	0.05	>6.2
5C	<0.01	0.03	>3.9	<0.01	0.04	>4.0
6B	0.007	0.01	3.1	0.007	0.01	3.2
M4(G→T)[e]						
5B T_2[g]	0.91			0.62		
5B T_3 [g]	0.46			0.45		
	0.13[f]			0.13[f]		

Table 5B. Spontaneous and cadmium-specific mutation frequencies.

	Mutant Frequency (events/plant)[b]			Mutant Frequency (Poisson average)[c]		
Subline[an]	Spontaneous frequency	Cd-specific frequency[j]	Ratio (Cd-specific/ Spontaneous)	Spontaneous frequency	Cd-specific frequency[j]	Ratio (Cd-specific/ Spontaneous)
M3(G→C)[f]						
Q1G	0.04	0.02	1.5	0.05	0.02	1.5
Q8B	0.03	0.02	1.5	0.03	0.01	1.3
M4(G→T)[f]						
5B T_2	1.54	0.62	1.4	0.98	0.98	2.0
		1.02[lm]			0.69[lm]	
		1.03[km]			0.64[km]	
20C	1.28	0.07	1.1	1.09	0.21	1.2
9F[d]	2.08	ND	ND	2.21	ND	ND
14A[d]	0.39	ND	ND	0.35	ND	ND
7D[d]	0.18	ND	ND	0.15	ND	ND
M6(T→C)[f]						
18F T_3[ik]		0.01[h]			0.01[h]	

[a] Independent transformants with constructs encoding *GUS*⁻ alleles reporting indicated pathways. Unless otherwise stated, the T_2 generation was assayed.
[b] 70-150 plants were scored, except K22B, K24B, A9F and T → G lines (n < 50).
[c] Mutation frequency calculated using Poisson distribution: x = -*ln*(zero fraction) as described in Results section (page 96).
[d] Lines showed possible germinal mutation (germinal events were not included in mutant frequency calculations).
[e] Spontaneous mutation frequency on soil.
[f] Plants grown under continuous light.
[g] Plants grown under photoperiod light.
[h] Total mutation frequency (spontaneous frequency was not determined).
[i] T_3 line tested due to low seed recovery in T_2.
[j] Cadmium chloride added to final 1 mg/L before autoclaving.
[k] Cadmium chloride added to final 1 mg/L after autoclaving.
[l] Cadmium chloride added to final 0.4 mg/L before autoclaving.
[m] Mutation frequency (value was less than spontaneous).
[n] Plants grown on MS media.

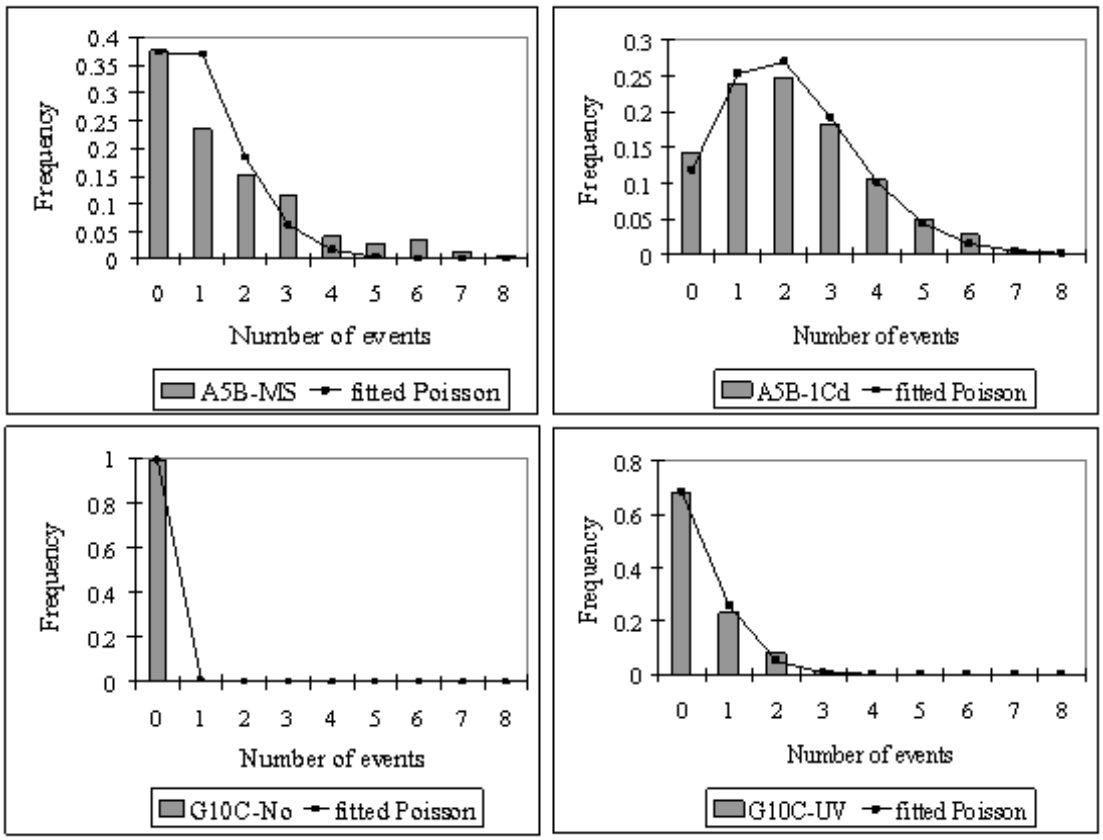


Figure 18. Comparison of data to Poisson distributions. Representative data are plotted as histograms (bars) compared to the fitted Poisson distribution (lines). Sample data is plotted as the frequency of plants with a given number of events. The fitted Poisson distribution was calculated using VassarStats.

Table 6. Comparison of mutation-frequency calculations.

Line	Treatment	Photoper-iodicity[a]	Mutation-event mean[b]	Zero-term[c]	Fitted-Poisson mean[d]
M1-11D T_2	Soil	24:0	<0.01	<0.01	0.01
	Soil+UV-B	24:0	0.11	0.11	0.12
	Soil+UV-C	24:0	0.09	0.09	0.09
M2-10C T_2	Soil	16:8	<0.01	<0.01	0.01
	Soil+UV-C	16:8	0.43	0.37	0.38
M2-12G T_3	Soil	16:8	0.12	0.11	0.11
	Soil+UV-C	16:8	0.18	0.13	0.12
	Soil	24:0	0.02	0.02	0.02
	Soil+UV-C	24:0	0.41	0.26	0.26
M3-8B T_2	MS	24:0	0.03	0.03	0.03
	MS+1Cd[e]	24:0	0.05	0.04	0.04
M4-5B T_2	Soil	16:8	0.91	0.62	0.60
	MS	24:0	1.54	0.98	0.99
	MS+1Cd[e]	24:0	2.17	1.96	2.13
	MS+1Cd[f]	24:0	1.03	0.64	0.71
M4-5B T_3	Soil	16:8	0.46	0.45	0.48
	Soil	24:0	0.13	0.13	0.13
M5-12D T_3	Soil	16:8	0.06	0.07	0.07
	Soil+UV-C	16:8	0.15	0.14	0.14
M6-15F T_2	Soil	24:0	<0.01	<0.01	0.01
	Soil+UV-C	24:0	0.11	0.10	0.09
M6-15F T_3	Soil	24:0	<0.01	<0.01	0.01
	Soil+UV-C	24:0	0.16	0.16	0.16
M6-18F T_2	Soil	24:0	<0.01	<0.01	0.01
	Soil+UV-C	24:0	1.04	0.92	0.96
M6-18F T_3	Soil	16:8	<0.01	<0.01	0.01
	Soil+UV-C	16:8	0.16	0.12	0.11
M6-18F T_3	Soil	24:0	<0.01	<0.01	0.01
	Soil+UV-C	24:0	0.05	0.05	0.05

[a] Plants were either grown in continuous light (24:0) or photoperiod light (16:8).
[b] Total number of events observed divided by the total number of plants assayed.
[c] Mutant frequency (x) calculated using Poisson distribution: x = -*ln*(zero fraction), as described under "Results" (page 96).
[d] Mutant frequency calculated from the best-fit Poisson distribution to the data.
[e] Cadmium chloride added to final 1 mg/L before autoclaving.
[f] Cadmium chloride added to final 1 mg/L after autoclaving.

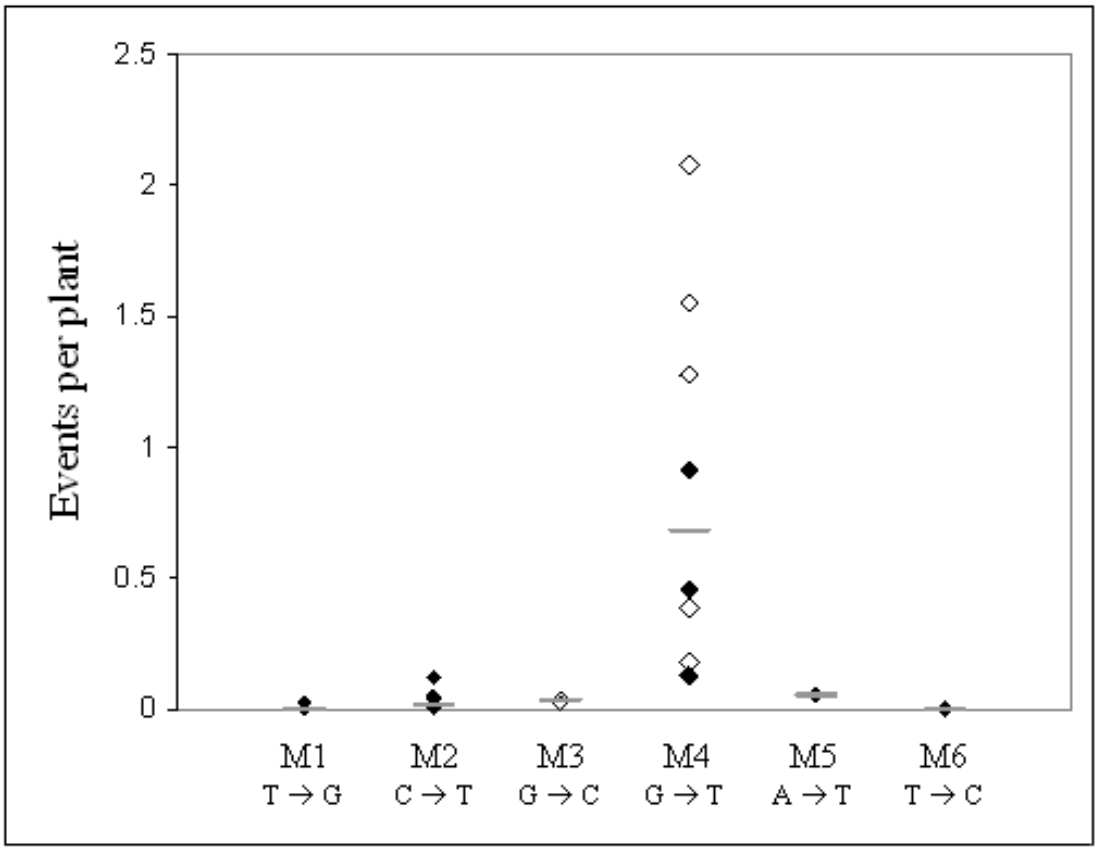


Figure 19. Spontaneous mutation frequencies. Data is from Table 5. Solid symbols refer to plants grown on soil; open symbols refer to plants grown on MS medium; grey bar indicates median value. M1 (T → G) sublines (top to bottom): 12D, 3F; 5 other sublines. M2 (C → T) sublines (top to bottom): 12G (T_3, photoperiod light); 20B; 8D; 15G; 12G (T_3, continuous light); 12G (T_2); 10C. M3 (G → C) sublines (top to bottom): 1G; 8B. M4 (G → T) sublines (top to bottom): 9F; 5B (T_2, MS); 20C; 5B (T_2, soil); 5B (T_3, soil, photoperiod light); 14A; 7D; 5B (T_3, soil, continuous light). M5 (A → T) sublines (top to bottom): 12D (T_3); 12D (T_2). M6 (T → C) sublines (top to bottom): 20C (T_3); 6B; 15F (T_2); 10 other sublines.

comparison, A:T constructs showed an average of 0.008 events per plant. Most striking was spontaneous G → T mutation. This might reflect endogenously-induced 8-oxoG because 8-oxoG in the template strand mispairs with adenine, leading to G → T transversions [8]. The M4 (G → T) sublines showed spontaneous mutation levels as high or higher than UV-C-induced mutation levels in M6 (T<u>T</u> → T<u>C</u>) and M2 (T<u>C</u>C → T<u>T</u>C) lines (Figure 20A versus 20B). Out of six M4 (G → T) sublines tested, three showed an apparent germinal mutation: staining pattern similar to that of wild-type constructs (Figure 21). G → T mutation levels were high on both soil and MS plates, so growth medium did not have a large impact here.

UV mutagenesis

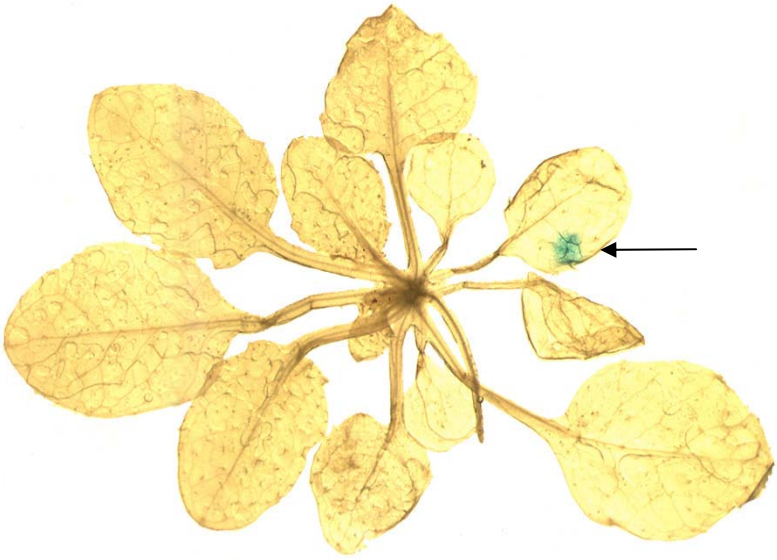
Initial experiments using UV radiation were performed on line M1 (T<u>T</u> → T<u>G</u>). The first set of plants (a total of 461) was grown on MS plates, but no reversions were seen with UV-C treatment. The second set of plants was grown on soil, and both UV-B and UV-C treatment led to increased reversion (Figure 16). We cannot explain why there was a different response between plate- and soil-grown plants.

Plants were given three different treatments: a single fluence of 1000 J/m^2 UV-C after two weeks growth (second leaves present but not expanded and third leaves barely noticeable), three doses of 5000 J/m^2 UV-B at the emergence of first (7 days), second (12 days), and third to fourth true leaves (18 days), or mock treatments performed in parallel with the UV-B treatments. After each treatment, plants were placed in the dark for 24 hours to prevent removal of UV-induced dimers by photolyases, thus increasing time for mutation fixation. Photolyase repair has been

A

T → C, no treatment (soil)

B



T → C, UV-C (soil)

Figure 20. Spontaneous and UVC-induced mutation.

C

G → T, spontaneous (MS medium)

Figure 20 (Continued). Spontaneous and UVC-induced mutation. Examples of spontaneous and UV-C-induced reversion events. (A) Subline M6 (TT → TC)-18F grown on soil. No mutation events were observed. (B) Subline M6 (TT → TC)-18F grown on soil; acute dose of 1000 J/m^2 UV-C. (C) Subline M4 (G → T)-5B grown on MS medium. Arrows indicate reversion events.

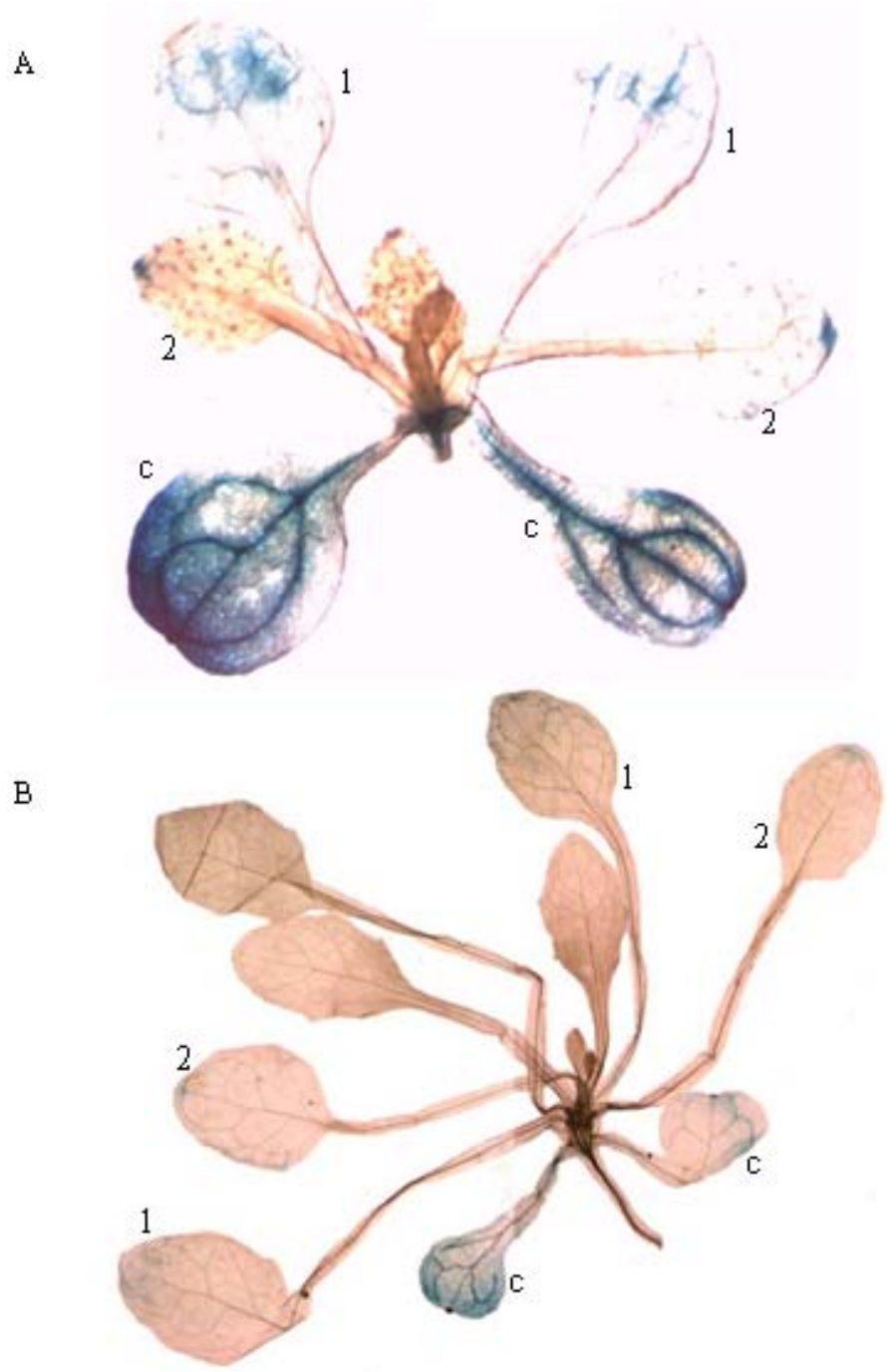


Figure 21. Staining of apparent line M4 (G → T) germinal-revertant plants. (A) Line W1804-AcV5-1 expressing GUS^{+}; (B) putative germline-reverted subline M4 (G → T)-14A progeny plant.

shown to be very rapid, with 30% of CPDs and [6-4]s repaired after only two hours and 50% of photodimers repaired in 24 hours. Conversely, dark repair is less efficient, with 13% of CPDs and 23% of [6-4]s repaired after 24 hours [25]. Therefore dark incubation leads to longer persistence of UV-induced lesions and a better chance for mutation. As shown in Figure 22, UV-B treatment did not effect plant growth, yet UV-C treatment led to stunted growth (a delay of roughly one week) and necrosis of the most exposed tissues. Our analysis indicates the two doses yielded similar levels of reversion, but the UV-C mutagenesis was more consistent. UV-C radiation is more efficient than UV-B in dimer formation and provides a more specific damage spectrum (UV-B induces dimers and oxidative damage). Treatment of the plants with UV-C radiation also allowed for comparison to previous studies, such as the Kovalchuk/Hohn reporter study [40].

The point-mutation constructs were anticipated to show different levels of response to UV-C treatment. Line M6 (T<u>T</u> → T<u>C</u>) and line M2 (T<u>C</u>C → T<u>T</u>C) were expected to show the highest reversion, with a smaller response expected in line M1 (T<u>T</u> → T<u>G</u>). Line M5 (A → T) was expected to show no response to UV treatment. Observed mutation frequencies are listed in Table 5. Line M2 (T<u>C</u>C → T<u>T</u>C) sublines showed high UV-specific mutation frequencies, as expected. The median frequency of UV-specific-mutation induction was 0.15 events per plant (0.17 Poisson-derived average events per plant), with a range from 0.06 – 0.42 events per plant (0.02 – 0.36 Poisson-derived average). In addition, a subline which contained the M2 (T<u>C</u>C → T<u>T</u>C) construct in the alternative binary vector p1804 (4–3G) showed mutation

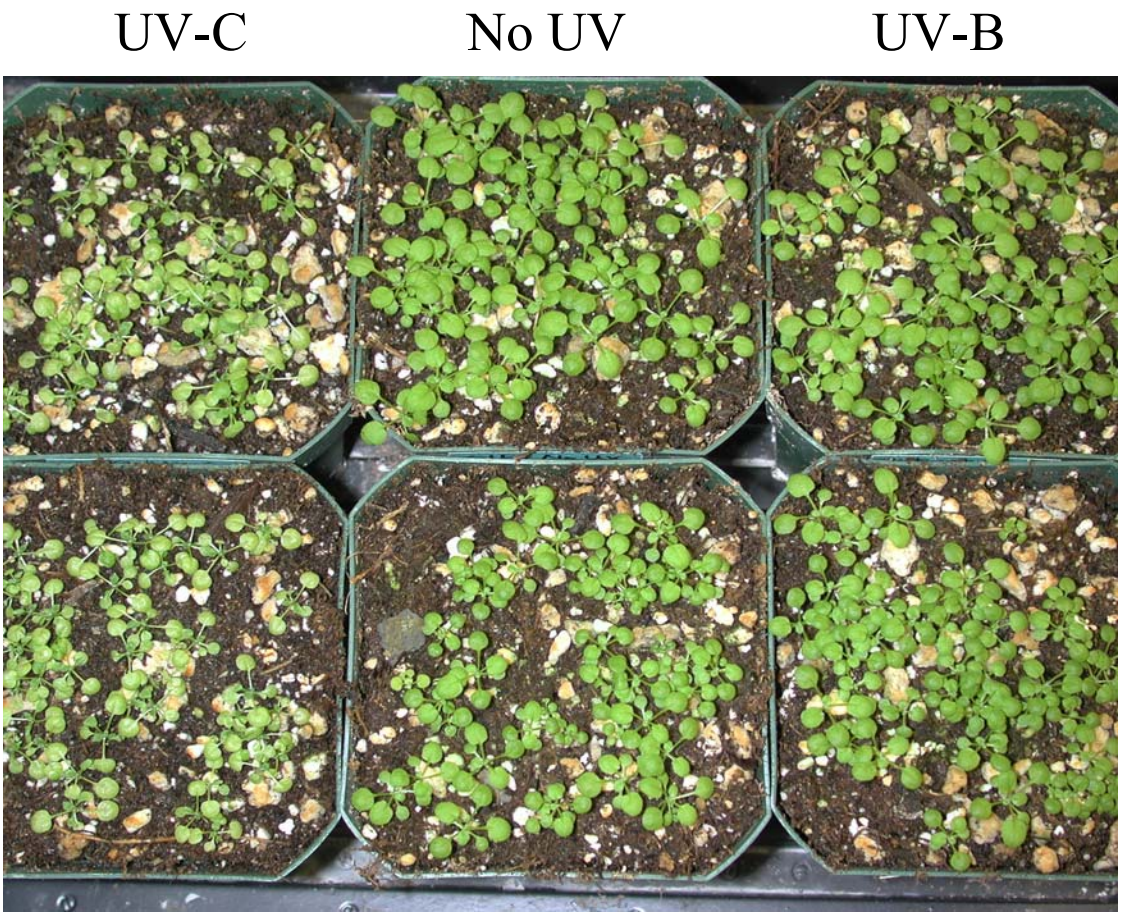


Figure 22. Effect of UV treatment on plant growth. Repeated doses of 5000 J/m^2 UV-B caused no observable change in plant growth, whereas a single dose of 1000 J/m^2 UV-C caused growth inhibition, and necrosis of the most exposed leaves. UV treatments are as described in Figure 16 legend.

induction similar to p1803-binary-vector-derived lines. Two C → T experiments were repeated. Plants from sublines 10C and 12G were exposed to UV-C as before, and showed mutation induction. However, the mutation frequencies were three-fold lower than the previous measurements (Appendix). This decrease could be due to variability between batches of soil or other growth condition factors. Therefore, it is critical to maintain equivalent growth conditions between control and treatment groups; testing of larger sample sizes may also be required. Although the overall C → T mutation induction was high, we observed similar induction in the M6 (T<u>T</u> → T<u>C</u>) lines, which conflicts with previous reports [20, 22]. One report that parallels our results was the Cupples/Miller study, which showed increased T → C reversion relative to C → T reversion [29]. The median UV-specific-mutation induction for line M6 (T<u>T</u> → T<u>C</u>) sublines was 0.11 events per plant (0.105 Poisson-derived average), with a range from 0.01 – 1.04 events per plant (0.01 – 0.92 Poisson-derived average).

Unexpectedly, lower levels of UV-specific-mutation induction were seen with line M1 (T<u>T</u> → T<u>G</u>) as compared to Kovalchuk/Hohn line 112 (T<u>T</u> → T<u>G</u>). The median UV-specific-mutation induction for M1 (T<u>T</u> → T<u>G</u>) sublines was 0.05 events per plant (0.06 Poisson-derived average), with a range from 0.03 – 0.09 events per plant (0.03 – 0.09 Poisson-derived average). In comparison, the Kovalchuk/Hohn line 112 (T<u>T</u> → T<u>G</u>) sublines showed a median UV-specific-mutation induction of 0.105 events per plant, with a range from 0.01 – 0.31 events per plant. Only one M5 (A → T) subline was tested, and it showed low UV-specific-mutation induction of 0.09

events per plant (0.07 – 0.10 Poisson-derived average). Figure 23 shows a plot of UV-specific-mutation induction in all reversion lines tested.

Mutagenesis by heavy metal ions

Our lines were tested for reversion on media supplemented with heavy metal ions for comparison to previous work by Kovalchuk, et al [9]. Cadmium and zinc were selected for analysis. The previously tested point-mutation constructs responded differently to cadmium and zinc [9]. It was expected that our full set of base substitution reporters would show even larger variation between the two metal-ion treatments. Based on results from the Kovalchuk study we chose 0.4 mg/L and 1 mg/L cadmium chloride concentrations for testing: they both lie within the reported linear phase of cadmium-dose response and a 2.5-fold difference in concentration can be analyzed. Several of our constructs were anticipated to show a response to heavy metal ions due to oxidative damage. The M4 (G → T) construct was expected to show the most response due to 8-oxoG mutagenesis [8, 9]. Indeed, the highest cadmium-specific mutation frequencies were observed in M4 (G → T) sublines grown on 1 mg/L $CdCl_2$ media. Increases over spontaneous mutation of 0.07 and 0.62 events per plant (0.21 and 0.98 Poisson-derived average) was observed for sublines 20C and 5B (Table 5). Lower cadmium-specific induction of mutation (0.02 events per plant) was observed in M3 (G → C) sublines 1G and 8B. However, heavy-metal-ion responses were not consistent. M4 (G → T) subline 5B (T_2 generation) mutation frequencies on the different control or heavy-metal-ion-supplemented plates were as follows: 1.54 events per plant (MS), 1.02 events per plant (MS + 0.4 mg/L $CdCl_2$), 2.17 events per

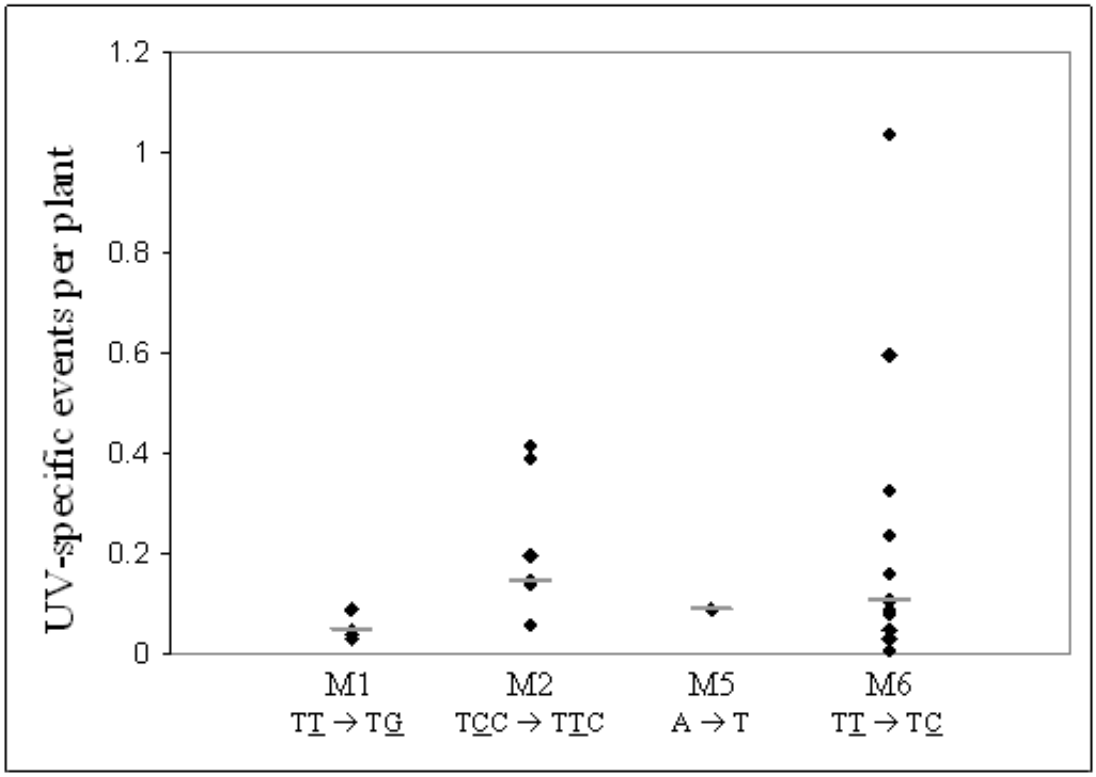


Figure 23. UV-specific mutations. All plants were grown on soil. Unless otherwise stated, lines M2 and M5 were grown under photoperiod-light conditions, and line M6 was grown under continuous-light conditions. Frequencies were corrected for spontaneous events. Grey bar indicates median value. M1 (TT → TG) sublines (top to bottom): 1E, 11A, 11D; 1A, 12D; 3F, 6A. M2 (TCC → TTC) sublines (top to bottom): 10C; 12G (generation T_3, continuous light); 20B; 12G (generation T_2), 8D; 15G; 12G (T_3, photoperiod light). M5 (A → T) sublines (equivalent values): 12D (T_2), 12D (T_3). M6 (TT → TC) sublines (top to bottom): 18F (T_2); 20C (T_2); 22B; 18G; 18F (T_3, photoperiod light), 15F (T_3); 24B, 15F (T_2, continuous light); 15F (T_2, photoperiod light); 20C (T_3); 21C, 18F (T_3, continuous light); 5C; 6B.

plant (MS + 1 mg/L $CdCl_2$), 0.61 events per plant (MS + 6 mg/L $ZnCl_2$), and 0.39 events per plant (MS + 18 mg/L $ZnCl_2$). A previously reported Kovalchuk/Hohn line that we tested also showed lower heavy-metal-ion response than expected. Line 166 (T → C) subline 1 showed 0.12 spontaneous events per plant, 0.05 events per plant on 1 mg/L $CdCl_2$ medium in which $CdCl_2$ was added before autoclaving, and 0.15 events per plant on 1 mg/L $CdCl_2$ medium in which $CdCl_2$ was added after autoclaving. Conversely, reported mutation frequencies were 0.02 spontaneous events per plant and 0.70 events per plant on 1 mg/L $CdCl_2$ medium in which $CdCl_2$ was added after autoclaving [9]. Growth and media conditions were as similar to those in the previous work as possible. We cannot explain the apparent lack of heavy-metal-ion response.

Photoperiodicity effects on mutagenesis

Homologous-recombination reporters have been used to study the effect of photoperiodicity (light to dark ratio) on mutation rates, but no studies with point-mutation-reversion reporters have been published. In addition, we observed differences in growth patterns between two different conditions. Therefore, we treated an M6 (T<u>T</u> → T<u>C</u>) subline and an M2 (T<u>C</u>C → T<u>T</u>C) subline with 1000 J/m^2 UV-C radiation. Plants were divided into two groups. One set was grown in continuous light, and the other set was grown in standard photoperiod light (16 hours light: 8 hours dark).

Among all the different lines treated with UV-C radiation, the majority showed growth stunting that required roughly one week for recovery. This occurred under both growth conditions. The observed growth inhibition was consistent with previous

observations by Curtis and Hays [80]. No UV control plants were stained three weeks after planting, whereas UVC-treated plants were allowed an extra week of growth to achieve the same size (eight to ten leaf stage) before staining. A few lines, however, showed a differential response dependent on growth conditions. Whereas continuous-light-grown plants were delayed for one week, photoperiodic-light-grown plants showed reduced growth stunting, requiring fewer additional days growth to reach appropriate staining size (Figure 24). It is not clear why some of the lines respond differently in photoperiodic light. I speculate this may be due to disruption of key genes by T-DNA integration. The two lines analyzed in detail for photoperiodicity effects showed differential growth responses: M6 (T<u>T</u> → T<u>C</u>) subline 18F (T_3 generation) showed growth stunting irrespective of day length while M2 (T<u>C</u>C → T<u>T</u>C) subline 12G (T_3 generation) showed reduced growth inhibition in photoperiod light.

Plants grown under continuous versus photoperiodic light conditions not only showed differences in growth response, but also showed variation in observed mutation frequencies (Table 5). The M6 (T<u>T</u> → T<u>C</u>) reporter in the T_2 generation (tested in continuous light) showed no spontaneous mutation (< 0.01 events per plant) and UV-specific-mutation induction of 1.04 events per plant (0.92 Poisson-derived average), which is a statistically-significant increase. Conversely, the T_3 generation showed a 20-fold reduction in mutation induction by UV-C when grown in continuous light after 24-hour-dark incubation, as well as low spontaneous mutation background (< 0.01 events per plant). There was also a difference in observed UV-specific-

A

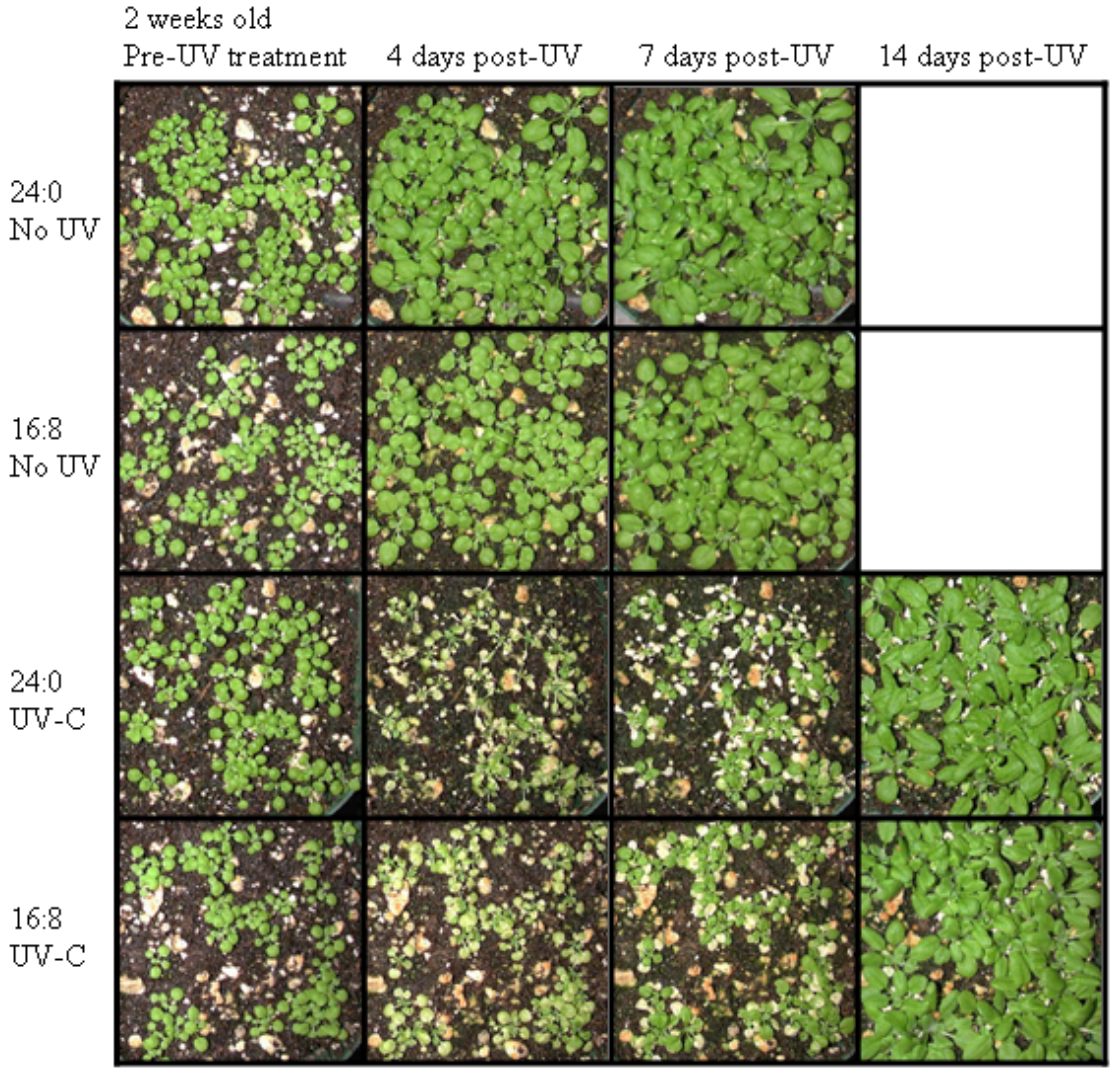


Figure 24. Effect of photoperiodicity and UV-C treatment on plant growth. (A) Subline M6 (T<u>T</u> → T<u>C</u>)-18F, generation T_3. The majority of plants showed growth stunting when treated with UV-C radiation. (B) Subline M2 (T<u>C</u>C → T<u>T</u>C)-12G, generation T_3. Some lines showed reduced UV-induced growth stunting when grown in photoperiod light. 24:0 indicates continuous light; 16:8 indicates photoperiod light.

B

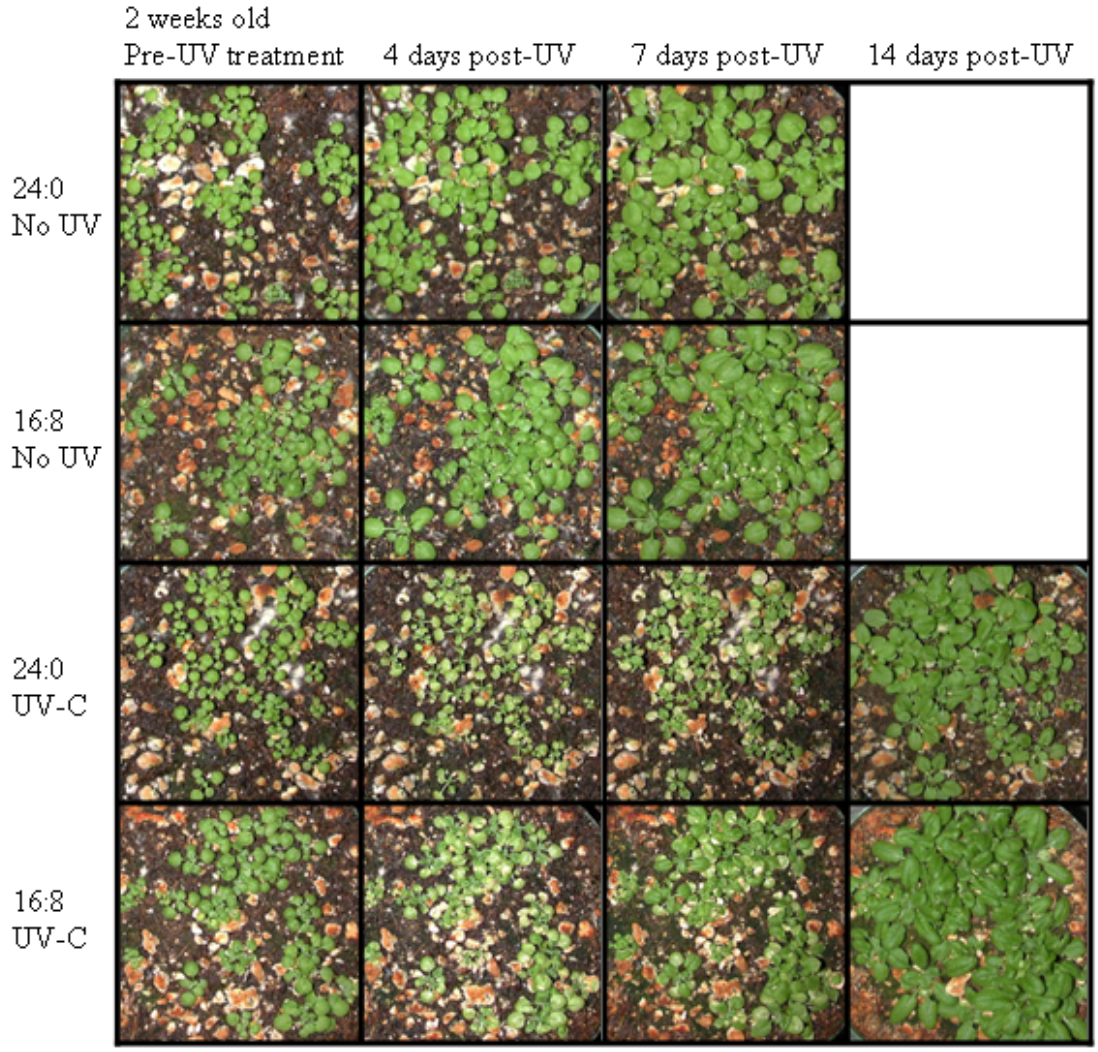


Figure 24 (Continued). Effect of photoperiodicity and UV-C treatment on plant growth.

mutation frequency under the two growth conditions: plants grown in photoperiodic light showed 0.16 UV-specific events per plant (0.12 Poisson-derived average) and plants grown in continuous light showed 0.05 UV-specific events per plant (0.05 Poisson-derived average). None of these differences were statistically significant. A second UV-C treated M6 (T<u>T</u> → T<u>C</u>) subline (15F, T_2 generation) grown under continuous or photoperiod light conditions also showed comparable mutation frequencies: 0.11 events per plant (continuous light) and 0.09 events per plant (photoperiod light).

The M2 (T<u>C</u>C → T<u>T</u>C) subline 12G showed a much different response than the M6 (T<u>T</u> → T<u>C</u>) sublines. The T_2 generation (grown in photoperiod light) yielded mutation frequencies of 0.008 spontaneous events per plant (0.008 Poisson-derived average) and 0.15 UV-specific events per plant (0.13 Poisson-derived average), which was a statistically-significant increase. The T_3 generation differed not only from the T_2 generation but also showed growth condition effects. Plants grown in photoperiodic light showed increased levels of spontaneous mutation (0.12 events per plant, 0.11 Poisson-derived average) and decreased UV-specific mutation induction (0.06 events per plant, 0.02 Poisson-derived average) that was not statistically significant. This subline shows the highest spontaneous mutation frequency measured among the M2 sublines as well as the lowest UV-specific increase in mutation frequency. We cannot explain why the spontaneous mutation was so high for this subline and growth condition. On the other hand, plants grown in continuous light showed low spontaneous mutation (0.02 events per plant, 0.02 Poisson-derived

average) and increased UV-specific-mutation induction (0.39 events per plant, 0.24 Poisson-derived average). The UV-induced mutation observed with the continuous-light treatment was statistically significant. Taken together, these data suggest mutation frequency can be dramatically affected by light photoperiodicity, and this variation can be dependent on the line or construct tested.

Variation in mutation with generation and subline

To determine if our lines would show the same variation due to generation as we observed with the Kovalchuk/Hohn lines, we repeated a selection of experiments using bulked T_3 seeds (Table 5, Figures 25, 26). The T_3 generation of M2 (T<u>C</u>C → T<u>T</u>C) subline 12G showed higher spontaneous mutation (0.12 events per plant) and lower UV-specific-mutation (0.06 events per plant) than the T_2 generation (0.008 spontaneous events per plant and 0.15 UV-specific events per plant). However, when I included the data from T_3 generation plants grown in continuous light conditions (0.02 spontaneous events per plant and 0.39 UV-specific events per plant), mutation frequencies between the T_2 and T_3 generations were similar. Differences in mutation frequency between the T_2 and T_3 generations were larger in M6 (T<u>T</u> → T<u>C</u>) sublines. Three lines were chosen to cover the range of mutation frequencies observed: 18F showed high reversion, 20C showed moderate reversion, and 15F showed low reversion. The two higher reverting sublines showed reduction of UV-specific mutation in the T_3 generation. Subline 18F showed reduction from 1.04 events per plant in the T_2 generation to 0.05 events per plant in the T_3 generation, and subline 20C showed reduction from 0.60 events per plant in the T_2 generation to 0.08 events

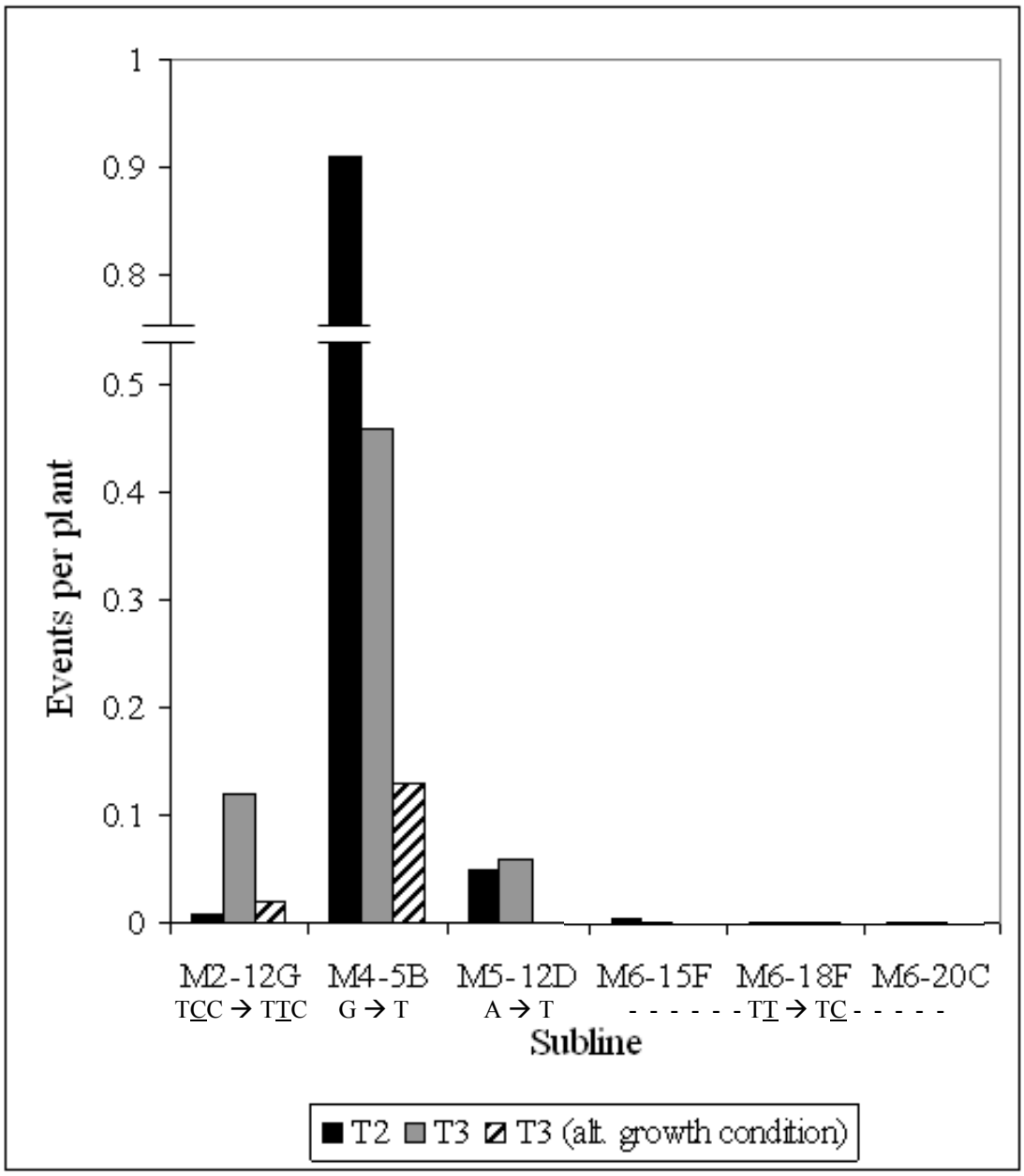


Figure 25. Comparison of spontaneous mutation frequencies in successive generations. Except where indicated, M2, M4, and M5 lines were grown under photoperiodic-light conditions, and M6 lines were grown under continuous-light conditions. Spontaneous mutation was measured in indicated sublines. (■) T_2 generation plants; (▩) T_3 generation plants; (▨) T_3 generation plants, grown under the alternate photoperiod condition.

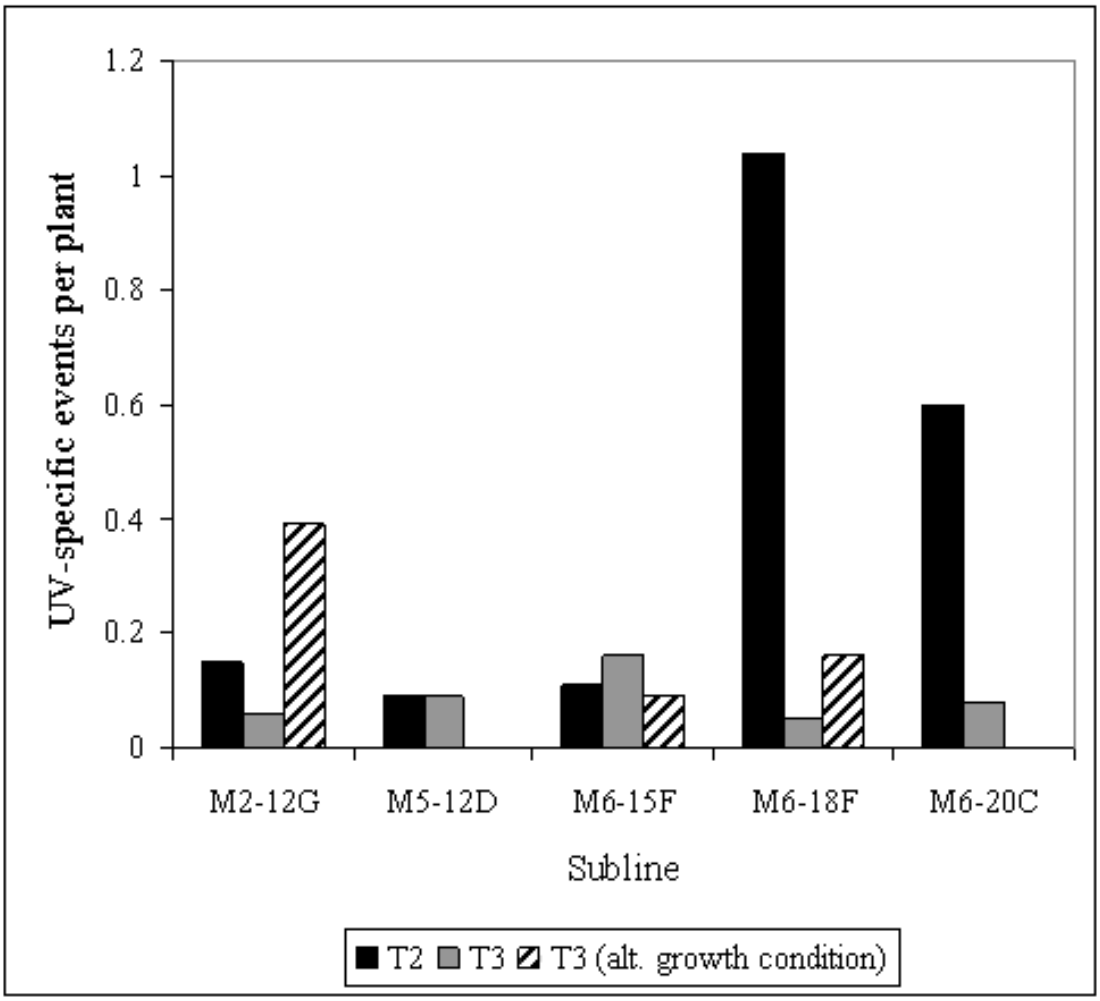


Figure 26. Comparison of UV-specific mutation induction between successive generations. Except where indicated, M2 and M5 lines were grown under photoperiodic-light conditions, and M6 lines were grown under continuous-light conditions. UV-specific mutation (corrected for spontaneous mutation) was measured in indicated sublines. (■) T_2 generation plants; (▒) T_3 generation plants; (▨) T_3 generation plants, grown under the alternate photoperiod condition.

per plant in the T_3 generation. The lower-reverting subline showed stability in UV-specific mutation frequencies between generations: 0.11 events per plant in the T_2 generation and 0.16 events per plant in the T_3 generation. In addition, a second T_2 subline (18G) of the high UV-specific reversion line was tested, and showed 0.24 UV-specific events per plant, four-fold lower than the 18F subline. The final subline tested for differences in UV-specific mutation across generations was M5 (A → T) subline 12D. Results were comparable between generations: 0.09 spontaneous events per plant and 0.10 UV-specific events per plant in the T_2 generation; 0.09 spontaneous events per plant and 0.07 UV-specific events per plant in the T_3 generation. Spontaneous frequencies of mutation in T_2 and T_3 generation M4 (G → T) subline 5B (grown on soil) were also measured. The T_2 generation showed 0.91 events per plant (0.62 Poisson-derived average), and the T_3 generation showed 0.46 events per plant (0.45 Poisson-derived average). This difference was almost statistically significant ($p = 0.067$). Overall, the results indicate there is a large amount of variability among different sublines and generations within the same line. Therefore assays of mutation induction always need to include a spontaneous-mutation control of the same seed pool for comparison.

Problematic reporter lines

The M3 (G → C) reporter line

As previously mentioned, the product of the glutamine TNT construct, corresponding to line M3 (G → C), showed 27% activity relative to the wild-type-

glutamate-TNT-construct product in the *in vitro* transcription/translation assay. This is not very surprising, since the amino acids are similar except for charge. However, the same amino acid substitution was tested in the Hohn lab with *Nicotiana plumbaginifolia* protoplasts and yielded < 1% activity [40]. The difference was assumed to be due to the type of activity assay employed. Our test results appear to be more accurate, however, since light background GUS stain is observed in the M3 (G → C) lines. Prior incubation of plants at various temperatures before staining was used to test the stability of glutamine-substitution GUS products at higher temperatures. The wild-type GUS controls showed consistent activity except when incubated for a long exposure at 65º C. The M3 (G → C) plant GUS activities were more variable, yet they showed the same temperature stability.

In order to circumvent the issues inherent with a background level of activity, lines that were shown on a western blot to contain lower levels of protein expression were isolated in an attempt to find a background stain that was low enough to detect reversion. Indeed, one line (M3-8B) displayed lighter background staining that was limited to cotyledons and stems. This allowed detection of reversion events on true leaves. When plants were tested on MS or heavy-metal-ion-supplemented plates, I was able to detect reversion events whose frequency increased with mutagen treatment. The reversion frequencies observed were similar to another, more-highly-staining background line (M3-1G). The agreement in the data suggests reversion detection in this reporter is possible despite the presence of background staining.

To confirm the presence of the correct transgene construct, DNA was isolated from differentially stained tissues for sequencing. DNA from lighter staining background tissue and unstained tissue contained the expected glutamine codon. These low-background lines can be used to detect mutation, with the caveat that researchers analyzing them need to ignore the background stain and only count isolated events. This presumably leads to a more conservative measure of reversion than achieved in other lines, due to the inability to detect reversions within the background-stained cotyledons and stems.

The M5 (A → T) reporter line

When the various amino acid substitution products were tested *in vitro*, the product of the valine TNT construct allowed < 1% relative activity, similar to the majority of constructs tested. Surprisingly, when independent M5 (A → T) lines were isolated and tested, substantial background staining was observed. The pattern is similar to that seen in wild-type constructs, although staining was more splotchy, and covered a significant portion of the plant. Only one subline, 12D, showed low enough background staining to allow further analysis. Even 12F, another subline derived from the same independent transformant, showed high background stain. An additional complication was the effect of growth conditions on background staining. 12D plants grown on soil showed light background staining, but when grown on plates their background staining was darker, similar to that of other M5 (A → T) sublines. Thus only UV-C mutagenesis was analyzed in this line. It is unclear why these construct lines contain such high background staining. DNA isolated from both background

staining tissue and unstained tissue confirmed the expected valine codon. Two generations were tested, and the ubiquitous but low background stain and the calculated mutation frequencies were similar.

Analyses of mutation reporters

Sequence confirmation and southern blotting of reporter lines

DNA was isolated from plants of various sublines to be used for sequence analysis. PCR amplification of the genomic DNA extracts was used for sequencing to confirm the expected base subsitutions (Appendix). Additionally, DNA was extracted from reversion-event tissue in histochemically GUS-stained plants to confirm wild-type reversion. A sample of genomic DNA extracts was also used for a southern blot. Preliminary data suggests different lines contain one to several copies of the transgene, as expected (data not shown). The rest of the lines will also be tested to determine transgene copy number.

Isolation of tissue for comparison of protein levels

Tissues from representative lines were isolated for extraction of total protein. Roughly 20-25 seedlings (eight- to ten-leaf stage) were collected for each extract. Tissues were isolated from roots, cotyledons, first true leaves, second leaves, and third leaves. Protein extracts were prepared as before (page 57). The western blot data yielded two interesting results. First, a doublet of full-length bands was observed in the blots, with the lower band more intense. The lower band mostly appears in root

and cotyledon extracts (Figure 27A). When we isolated whole seedlings for initial protein characterization, the lower band was used for quantification because the upper band was not seen reproducibly. The positive control that was expressed in *E. coli*, however, appeared to be the same size as the upper band. I speculate the two bands might reflect post-translational modification or proteolytic processing of GUS. When calculating transgene-expression levels, the signals of both full-length bands were added together. The second observation of interest is the pattern of expression. As expected, roots showed the highest levels of expression, with high levels also seen in cotyledons. Expression in true leaves was relatively low (Figure 27B). Thus quantified protein-expression levels confirm the observed staining pattern in the wild-type construct lines.

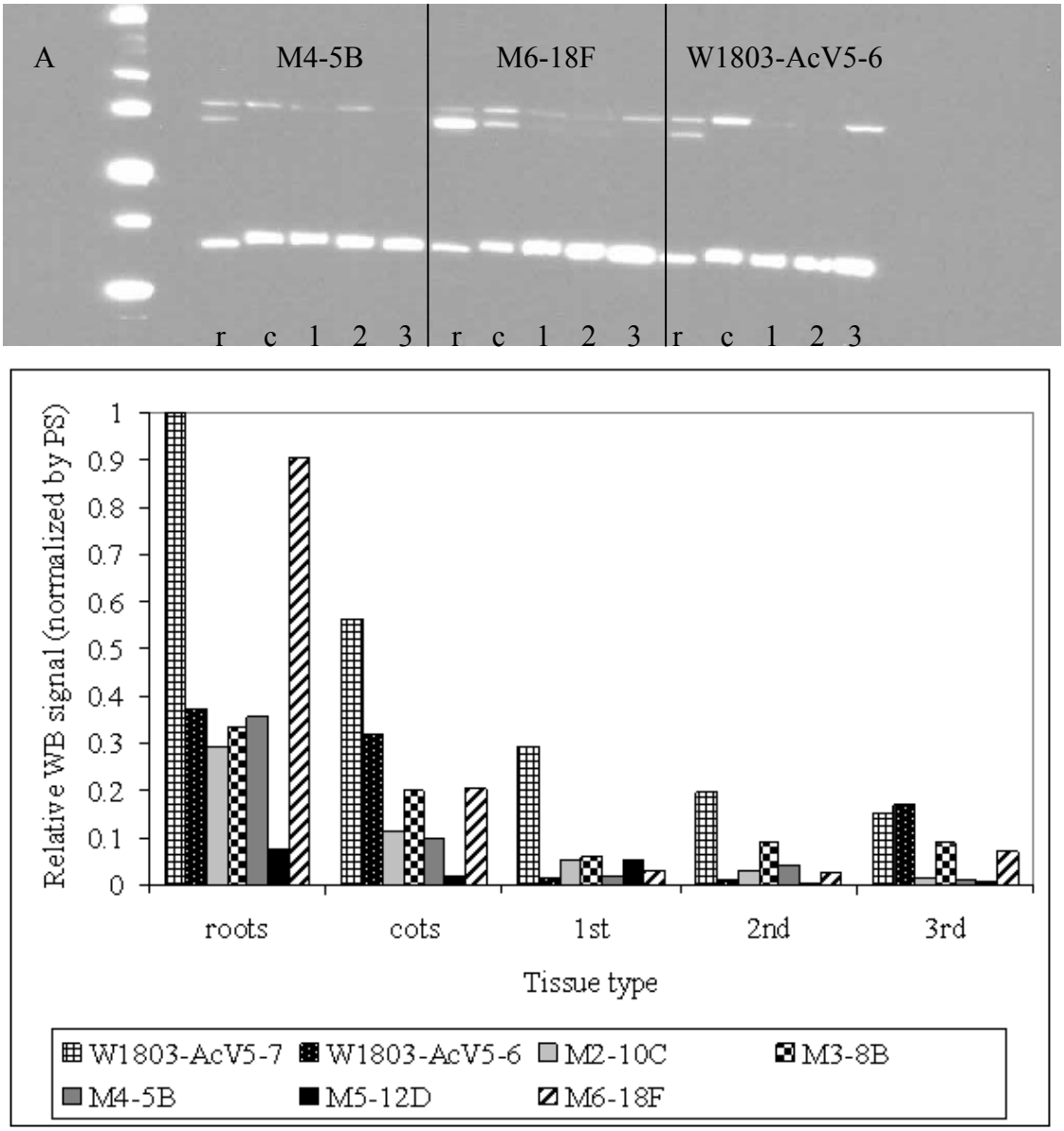


Figure 27. Protein expression in different plant tissues. Tissue was isolated from roots, cotyledons, first, second, and third pairs of true leaves and protein was extracted as described in Materials and Methods (page 57). (A) Representative western blot of M4 (G → T)-5B, M6 (TT → TC)-18F, and W1803-AcV5-6. (B) GUS protein levels in tissue samples were normalized by Ponceau S staining. Ponceau S staining and western blot were performed and quantified as described in Figure 12 legend.

Discussion and Conclusions

Discussion will begin with an overview of the limitations of the previous plant point-mutation reporters and how my reporter lines correct those deficiencies. Initial observations using my reporter lines will then be discussed, followed by a review of the mutation data and its significance. Finally, suggested future experiments will be presented.

Limitations of previous transgenic mutation reporters

The Kovalchuk/Hohn reporters have proven very useful for analysis of *in planta* mutation, but there are some limitations on their utility. First, the point mutation reporters show response to UV-C treatment, but the target sequences do not include 5'TC 3', a pyrimidine pair that has a high propensity for dimerization as well as the highest mutagenicity relative to other dipyrimidines. Also, the predominant mutations induced by UV-B and UV-C treatment are C → T and T → C, but the Kovalchuk/Hohn lines only include a T → C reporter, and it is within an unfavorable context (5'CT 3') [40]. T → C mutations are typically observed as TT → TC [22], and dimers are formed more often within a TT context than a CT context [20]. Likewise, the C → T reporters assayed in the Depicker laboratory [27] were designed more for studies of cytosine deamination than photodimer formation. Due to this different emphasis, only two of the five constructs can be used for analysis of UV-induced photoproducts and mutation. The first context used by the Depicker group measures CC → CT transitions. 5'CC3' has been shown to be a mutagenic context.

However, CC → TT is also likely to be induced, which would not allow wild-type activity. Indeed, the 5'CC3' construct shows UV-induced mutation levels similar to the non-dimer contexts. The lack of a difference between mutation levels of dimer and non-dimer constructs is also likely due to the treatment protocol used. In the Depicker experiment, the UV-C dose was low and plants were immediately placed back in the light, allowing rapid reversal of UV-induced photoproducts by photolyases. The second dimer-specific reversion that they observed was TC → TT, which should show high UV-induction. However, the authors were unable to measure the spontaneous reversion rate, and the UV-induced mutation frequencies were lower than the other constructs [27]. In comparison, the lines that I used in this study included a TT → TC reporter and a TC → TT reporter, whose transgenes allowed analysis of UV-induced mutation at sites that are more favorable for dimer formation and are more mutagenic. The other advantage of these lines over the previous point mutation reporters is that all reversions are localized within the same codon. Mutation induction can vary with sequence and chromatin context, so constructing original reversion reporters all within the same codon minimizes variation due to local sequence context. This should enable better comparison among my different reporter lines by neutralizing a source of variation.

Initial observations of properties of reporter lines

The Gelvin superpromoter

Although the Gelvin superpromoter provided strong expression levels, the expression tended to be limited to mature tissues, instead of being universally expressed. This tissue-specific expression was demonstrated by both the histochemical staining of the wild-type constructs and the tissue-specific western blotting for protein expression. GUS-expressing lines predominantly showed GUS expression preferentially in more mature tissues, and western blotting of separated tissues revealed high GUS expression in roots and cotyledons, with decreasing expression in the true leaves. More experiments need to be performed to ascertain the cause of this variegated expression, and mRNA levels in the different tissues should also be quantified to determine if the transgene is being transcriptionally or translationally regulated. The critical lesson is the importance of studying transgene expression at the protein level, as opposed to the mRNA level, especially when utilizing a new promoter cassette. Reporter genes such as GUS work well for this purpose, as they provide an easily visible display of where and when expression of a transgene occurs. This information can then be used to determine whether or not a given promoter fits the expression requirements for the experiment of interest. In this case, the superpromoter cassette does provide high protein expression, yet may not be reliable for studying transgene expression in certain tissues, such as newly emerging leaves and early floral tissue. Root tissue, which showed the highest levels of

expression, was not tested due to the inability to reproducibly dissect out and stain intact tissue.

Although the Gelvin superpromoter shows limitations with respect to the location and timing of transgene expression, it may be resistant to transgene silencing. Overexpressed transgenes are often silenced in the recipient plant. However, promoters similar to the Gelvin superpromoter have shown resistance to silencing [64, 65]. Perhaps promoters with more plant-like enhancer and promoter elements, as opposed to viral promoters, may be able to escape silencing where other strong promoters cannot. Plants have developed antiviral defenses, such as siRNA [52], which may silence viral-based promoters more efficiently than other promoters. The Gelvin superpromoter is an effective, strong promoter that limits studies of transgene expression by this promoter to older tissues.

In vitro transcription/translation assay

Protein activities quantified using *in vitro* or transient *in vivo* systems are not always reliable. This is evidenced by two inconsistencies observed in this study. The M3 (G → C) lines showed a low but ubiquitous level of GUS activity upon histochemical staining that confirmed the *in vitro* transcription/translation quantification of 27% relative activity. However, the same construct showed < 1% GUS activity in a transient transfection expression system [40]. Secondly, the M5 (A → T) construct product showed < 1% activity in the *in vitro* assay, a very high level of ubiquitous histochemical staining in the plant lines. This suggests that simplified assays of protein activity may not always reflect the true nature of protein activity *in*

vivo. The *in vitro* transcription/translation assay was performed using mutant GUS DNA constructs lacking the CaMV ORF V fusion. It was not anticipated that lack of the N-terminal fusion would influence the activity of the GUS enzyme. However, this cannot be ruled out, especially as one of the *in vitro* product activities did not match *in planta* activity.

Reversion frequency analyses

Mutation frequency was measured as events per plant, similarly to Kovalchuk, et al [40], due to variability in plant growth among the different treatments. Plants were histochemically stained at three weeks after germination, at which point most plants had eight to ten leaves but some had as few as two. In addition, the soil-grown plant leaves were much broader than plate-grown plant leaves. Therefore the soil-grown plants may have undergone more cell division and endoreduplication. Photoperiodicity also induced variation, as witnessed by the enhanced growth rate of continuous light-grown plants over photoperiod-grown plants, particularly when grown in soil. Despite the fact that seeds were imbibed for 72 hours before sowing, plant growth was not fully synchronous, even on plates. Lastly, the acute dose of UV-C presumably affected only the cell divisions that occurred before UV-lesion removal, as opposed to the chronic exposure to heavy metal ions that affected all cell divisions. Therefore the rate of DNA divisions was not the same for the different treatments. This makes it more appropriate to measure events per plant and to include spontaneous mutation controls grown in comparable conditions (i.e. soil or plate growth). However, measuring mutation as events per plant, as opposed to events per division,

makes measurements of reversion only semi-quantitative. Also, we can directly compare our results only to previous plant studies with similar reports. The preliminary measurements of mutation frequencies of the Kovalchuk/Hohn point mutation reporters were also reported as events per plant [40], so direct comparisons could be made with my work.

Transgene copy number

A preliminary southern blot suggested that even though lines were selected for apparent single sites of insertion, tandem insertions occurred. The rest of the lines will undergo southern blot analysis to identify copy number. Transgene copy number determination is critical since the number of insertions can influence mutation frequency. For example, if two copies of the GUS transgene are present, the plant line may show two-fold higher mutation, dependent on the expression of each transgene. Plants with multiple transgene copies add complexity to mutation frequency calculation, because each copy has the potential for mutation, but in reality may have different mutation induction rates or levels of expression. Ideally, lines with a single transgene copy should be used for further mutation analysis to minimize unknown variables and simplify calculations.

Spontaneous mutation

Previous studies of the mutational effects of endogenous oxidative damage show G→ A, G → T, and CC → TT mutations, although cytosine deamination was suggested to be the main cause of C → T (G → A) transitions [8]. Conversely, a high

spontaneous frequency of G → T transversions was observed in our lines, with lower levels of G → A transitions. Two possible explanations for this difference in reversion levels are lower relative levels of cytosine deamination in *Arabidopsis* or differences in growth conditions inducing different ROS production as compared to previous reporters. In addition, as our C → T reporter was designed for response to UV-irradiation rather than endogenous damage, the cytosine does not lie within a methylation context (CpG or CpNpG). Methylcytosine is more mutagenic than cytosine during deamination since it would become thymine instead of uracil, which is recognized more readily by repair enzymes. Hence the lack of methylation ability could account for some of the discrepancy observed between studies.

The other interesting observation with the M4 (G → T) sublines was the differing responses to growth conditions. Plants grown on soil showed lower spontaneous mutation than plants grown on plates (0.91 versus 1.54 events per plant, respectively), and the difference was statistically significant ($p = 0.0114$). A comparison of soil-grown plants grown in different day length conditions revealed another significant difference; continuous light led to lower mutation than photoperiod light (0.13 versus 0.46 events per plant, respectively; $p = 0.0025$). These mutation differences may be due to differences in photosynthesis levels, metabolism levels, and overall growth between the different photoperiodicities. Plants grown on soil depend much more on photosynthesis than plants grown on plates, as the MS medium contains sucrose. The plate-grown plants almost exclusively use sucrose from the plates as a carbon and energy source and thus have minimal need for photosynthesis.

This difference in levels of photosynthesis may influence endogenous levels of oxidation, both through different levels of ROS production and radical-scavenger expression.

Overall, spontaneous mutation appeared higher at G:C base pairs than A:T base pairs. Even when the highly reverting M4 (G → T) sublines were excluded from the comparison, the average G:C frequency was five-fold higher than the average A:T frequency. This supports previous findings that guanine is a better target for oxidation than adenine. This difference could be influenced by both the relative propensity for damage and the relative efficiency of repair for each base.

Differences in UV-C and heavy-metal-ion treatment

The two mutagens utilized in this study represented two forms of exposure. The UV-C treatment was given as a single acute dose when only four to six leaves were expanded on the plants. The smaller, developing eight to nine leaves and leaf progenitors were gathered above the shoot apical meristem and were thus shielded from the irradiation. Evidence of the shielding effect included the restriction of the majority of reversions to the oldest leaves. In one instance, events were detectable early in the staining process, and reversion events were observed in the necrotic tissue that was induced by UV-C treatment. In addition, plants were allowed to grow for two weeks after irradiation. The majority of reversion events were thus larger spots or sectors, due to the increased time for cell division and expansion. These large reversion events were easier to identify and their size suggested that the reversions were indeed induced by the UV-C treatment. Smaller spots were also observed, but

as the spontaneous mutation frequencies were so low (an average of 0.002 in T → C lines and 0.04 in C → T lines), the smaller UV-specific-mutation events were more likely due to UV-induced mutation in terminally-differentiated tissue. Heavy-metal-ion treatment, on the other hand, provided continuous exposure, most likely to all tissues. This is supported by the data, which revealed a variety of sizes and location of events after exposure to heavy metal ions (data not shown).

UV-C response

The UV-specific mutation induction I observed differed from the Kovalchuk/Hohn data [40]. The range of UV-specific induction among different sublines for their codon 112 lines was 0.01 to 0.31 events per plant (median of 0.105). Overall, their induced TT → TG mutation frequencies ranged from 0.01 to 0.78 events per plant (median of 0.20). In contrast, my M1 (T<u>T</u> → T<u>G</u>) reporters showed a relatively consistent but low response to UV-C of 0.03 to 0.09 UV-specific events per plant (median of 0.05). I cannot rule out the possibility that by chance I selected only low-responding lines for testing. If I use the published codon-112 data set for probability calculation, P = 0.03125 is the probability of isolating all low-reversion lines. It seems more likely that I simply observed a different response than the previous workers.

Also unexpected was the observed M5 (A → T) mutation induction. This line was included as a negative control, since dimer formation is not possible, yet UV-induced mutation was observed at levels equivalent to the M1 (T<u>T</u> → T<u>G</u>) lines (0.09 events per plant in both generations). The Cupples and Miller *E. coli* study [29]

showed UV-induced mutation in their A → T line, although this may have been due to SOS response rather than direct UV damage. There are two possible explanations for our results. First, in our plant lines, UV-C exposure led to growth inhibition and necrosis of exposed tissues. It may be this stressful condition and any endogenous mutagens it induces that the M5 (A → T) reporter is able to respond to instead of the direct UV damage. Second, the M5 (A → T) reporter line showed ubiquitous background staining. Although apparent reversions could be discerned above the background, it is possible that the calculated mutation frequencies were not accurate. Further analysis with this line should help to confirm the reversion data. For comparison, the Kovalchuk/Hohn 118 (A → T) reporter lines showed a range of UV-specific induction of 0.004 to 1.16 events per plant, but the median induction was 0.04. This base substitution occurred within a 5'CT3' sequence context that might allow dimer formation, yet T → A is not a highly observed mutation. Therefore my observed M5 (A → T) data agrees with previous observations.

A difference in response between our M6 (T<u>T</u> → T<u>C</u>) lines and the Kovalchuk/Hohn reporter lines was also observed. Two Kovalchuk/Hohn T → C sublines showed UV-specific-mutation induction of 0.05 and 0.56 events per plant, which was similar to the UV-specific induction of 0.01 to 1.04 events per plant (median of 0.11) that I observed. However, the spontaneous mutation frequencies observed in the two reporter lines were quite different, with an observed range of 0 to 0.01 (median of 0) in our M6 (T<u>T</u> → T<u>C</u>) lines versus 0.02 and 0.26 spontaneous events per plant in the Kovalchuk/Hohn sublines. This difference may be due to the

influence of the codon that was targeted for reversion and the local sequence context (5'TTTT3' in our T → C lines, 5'TTCT3' in the Kovalchuk/Hohn codon 166 sublines).

The other striking result was the relatively large number of T → C reversions relative to C → T. Our data showed a range of 0.14 to 0.42 UV-specific events per plant with a median of 0.15. Most previous studies showed C → T events to predominate in UV-mutation spectra with T → C occurring at a much lower rate [22]. The one striking exception is the work by Cupples and Miller, which showed slightly more T → C than C → T at higher UV doses [29]. This effect may have been influenced by other types of damage, however, as the survival of the tested bacteria was very low after UV treatment and likely reflected a stressed condition. The highest mutation frequencies, corresponding to 0.75% survival, showed a four-fold predominance of T → C relative to C → T. Yet the lowest UV exposure, equivalent to 14% survival, resulted in a two-fold excess of C → T relative to T → C. Therefore less toxic UV exposures showed a more typical UV response. The preferential UV induction of C → T mutations is suggested to be due to more efficient repair of T[CPD]T [19], as well as CPD-cytosine deamination [20] and SOS bypass.

There are several factors which may influence our observed results. First, the observed ratio parallels the dimer-induction-spectrum (TT and TC equally likely to form dimers) but differs from the reported mutation spectrum. Therefore the factors that contribute to the difference between frequencies of dimer formation and observed mutation may be different in *Arabidopsis* than in other systems. C → T transition

frequencies are suggested to be more prominent due to frequent cytosine deamination. Yet it is unknown how often cytosine deamination occurs in *Arabidopsis*, and it is possible that levels of cytosine deamination are lower than expected. Alternatively, since most mutations likely occur during the 24-hour dark period immediately after UV-C treatment, the cytosine deamination frequency may be too low or slow to make a difference in the observed mutation spectrum. The low mutagenicity of TT dimers relative to TC may be another possible factor in the reported mutation spectrum. T[CPD]T dimers show high propensity for correct adenine incorporation during DNA translesion synthesis in biochemical experiments with Pol η. However, T[CPD]T bypass is not completely error-free. Additionally, Pol η may not be expressed highly in vegetative tissue relative to other, more error-prone translesion polymerases that are preferentially expressed in somatic tissue [1]. If error-prone translesion polymerases are used for bypass of T[CPD]T dimers instead of Pol η, the occurrence of TT → TC mutations could increase.

Another factor which may contribute to our observed results is overall DNA repair efficiencies. The observed ratio of T → C versus C → T reversions tends to agree with expected repair efficiencies of photodimers. [6-4]s are repaired more efficiently than CPDs, which would lead to a two-fold excess of TT over TC dimers. TT dimers are primarily T[CPD]T and TC dimers form equivalent levels of T[CPD]C and T[6-4]C. When UV-treated plants are grown in photoreactivating (white or blue/UV-A) light, the half-life of [6-4]s is roughly 20 minutes, and that of CPDs is roughly 40 minutes. Conversely, dark repair of photodimers is more skewed. After

eight hours, up to two-thirds of [6-4]s are removed from DNA, whereas CPDs show as little as 13% repair after 24 hours, versus 23% repair of [6-4]s in the same experiments [25, 81]. Therefore, even if TT dimers are less mutagenic, an excess in overall numbers could result in the observed mutation spectrum. However, this is not observed in other organisms.

The transcription level of the transgene may be a factor in differential repair. Our reporter constructs are driven by a strong promoter, as opposed to previous forward mutation reporters which were integrated into the host genome but not transcribed. There are repair pathways specific to the template strand in transcriptionally active DNA. There might be some bias towards repair of one type of dimer over another that would result in the observed ratios.

Finally, the observed mutation spectrum could be influenced by the sequence context of the photoproduct site. Surrounding sequence, especially additional adjacent pyrimidines, has been shown to be a factor in photodimer formation. The TT site mutated in our reporter is at the 3' end of a 5'TTTT3' sequence context, and the TC site is in the middle of a 5'TTCC3' sequence context. Due to the high number of adjacent pyrimidines, these constructs should both have increased photodimer formation, but the relative sequence context contribution to increased dimer formation is unknown, and irradiation of 5'TTTT3' context may induce more photoproduct than irradiation of 5'TTCC3'. In one study of sequence-context dependent dimer formation, a 5'CTTC3' context showed two-fold higher dimer formation than 5'ATTG3', and 5'CTTTA3' formed four-fold more dimers than 5'ATTG3' [82].

Thus adjacent thymine bases may promote dimer formation more than adjacent cytosines. Further studies in our laboratory, including biochemical assays with various sequence contexts, need to be conducted to determine the relative mutagenic potential of these contexts. Additionally, crossing the M2 (T<u>C</u>C → T<u>T</u>C) and M6 (T<u>T</u> → T<u>C</u>) lines with a line deficient in [6-4] or CPD photolyase should help to reveal the relative contribution of each dimer type to the different mutation pathways. Another interesting experiment would be crossing the M2 and M6 lines with AtPolη-deficient plants to see if there is an increase in reversion.

Mutation response to heavy metal ions

One surprising result was the lack of a toxic response in the reporter lines after exposure to heavy metal ions. I saw no difference in the growth of plants when grown on MS plates versus MS + Cd^{2+} or MS + Zn^{2+} plates (heavy metal ion salts added before autoclaving). In contrast, a previous study showed exposure to equivalent concentrations of cadmium in liquid media to result in smaller, stressed plants with stunted root growth [83]. Even more intriguing is my preliminary observation that plants grown on cadmium plates to which the cadmium salt was added after autoclaving showed a healthier, more robust phenotype than plants grown on MS plates (Figure 28). The plates are from different batches of medium, so we cannot rule out potential differences between the two media, yet the difference would have to be drastic to induce the plants to grow better on medium supplemented with a toxic

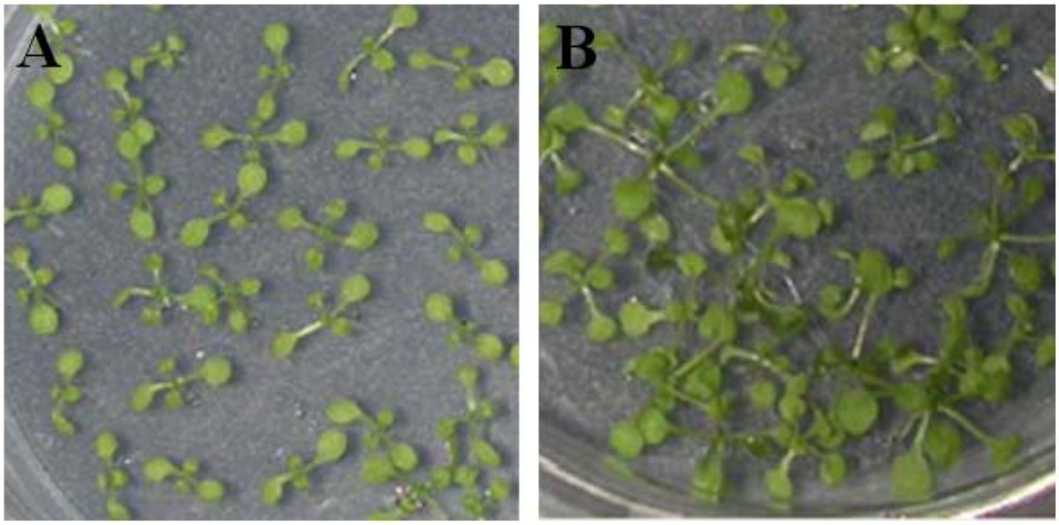


Figure 28. Growth of plants on cadmium-containing plates. (A) Plants grown on standard MS plates. (B) Plants grown on MS plates to which cadmium chloride was added to 1 mg/L after autoclaving.

carcinogen. Future experiments include testing plates with higher doses of cadmium and zinc salts in order to observe a toxic response, and then measuring the resulting reversion events.

As well as no toxic response, I observed no consistent induction of mutation by heavy metal ions. Previous frequencies of point-mutation reversion in response to heavy metal ions were as high as 0.65 events per plant with 1 mg/L $CdCl_2$ and 0.44 events per plant with 18 mg/L $ZnCl_2$ [9]. No significant heavy-metal response was observed in line M6 (T → C), even though the same reversion was assayed in the previous report. Line M3 (G → C) sublines showed a low heavy-metal response of 0.02 events per plant on 1 mg/L $CdCl_2$ medium. The high spontaneous mutation reporter line M4 (G → T) subline 5B showed a significant cadmium-specific induction of 0.62 events per plant on 1 mg/L $CdCl_2$ medium. However, a dose response was not observed for the heavy metal ions tested. Reversion frequencies observed were 1.02 events per plant on 0.4 mg/L $CdCl_2$ medium, 0.61 events per plant on 6 mg/L $ZnCl_2$ medium, and 0.39 events per plant on 18 mg/L $ZnCl_2$ medium. These results were all obtained using media in which heavy metal salts were added before autoclaving, so plants were tested on media in which heavy metal salts were added after autoclaving to determine if this would alter mutation induction. Surprisingly, line M4 (G → T) subline 5B plants grown on 1 mg/L $CdCl_2$ medium with cadmium added after autoclaving showed a reduced reversion frequency relative to previously measured spontaneous levels (1.03 and 1.54 events per plant, respectively). We mimicked the treatment conditions in the Kovalchuk/Hohn heavy-metal-ion study as accurately as

possible, and cannot explain why a heavy-metal response was not observed. Higher doses of heavy metal ions sufficient to induce a toxic response in the plant lines will be analyzed for induction of reversion in the reporter lines.

Photoperiodicity response

In the majority of plant lines tested, the only discernible differences between the two growth conditions were longer hypocotyls and slightly faster growth rates in continuous-light-grown plants. However, plants grown under either photoperiodic- or continuous-light conditions were equivalently stunted by UV-C treatment. There were a few lines, however, which showed reduced stunting of growth by UV-C radiation when grown in photoperiodic-light conditions. If this response is repeatable, it would be interesting to characterize transgene insertion sites in these lines to determine if there is a link to the enhanced post-UV recovery.

The more intriguing differences between photoperiod conditions were the differing mutation frequencies. In the M6 (T<u>T</u> → T<u>C</u>) sublines assayed at both day-length conditions, no significant difference was observed in either growth response or mutation induction. The M2 (T<u>C</u>C → T<u>T</u>C) line, however, showed significant photoperiodicity-dependent differences in reversion. The continuous-light-grown plants showed much higher UV-specific-mutation induction than photoperiodic-light-grown plants. This at first seems paradoxical as increased light exposure should correspond to increased photoreactivation. However, the key difference between the growth conditions may lie in growth leading up to UV treatment. Continuous-light-grown plants grow at a faster rate than the photoperiod-light-grown plants, and the

combination of UV-C treatment and a 24-hour dark period may be more stressful for the continuous-light-grown plants, which are unaccustomed to dark incubation. I suspect that the majority of mutations are fixed during the 24-hour dark period after treatment, hence a faster growth rate in the otherwise continuous light-grown plants may lead to a relatively higher rate of replication during the mutagenic dark period. In addition, as the plants grown in continuous light are constantly able to use photolyases for repair, dark repair systems may be expressed at lower levels relative to plants grown in photoperiodic light, which daily rely on dark repair systems for damage repair.

The other interesting observation was the differences in spontaneous reversion in the M2 (T<u>C</u>C → T<u>T</u>C) line due to photoperiod effects. Plants grown in photoperiodic light showed a roughly six-fold higher level of spontaneous mutation than plants grown in continuous light, although this seems inconsistent. Plants grown continuously in the light are growing at a faster rate and may operate under higher levels of stress. Different growth conditions might also lead to differences in DNA repair-system activity. Further studies need to be performed to confirm the day-length-dependent response in this line, as well as test other lines for the same response. Additionally, it would be interesting to ascertain if any repair systems are involved in the photoperiodicity response.

Future experiments

The point-mutation-reversion lines can be utilized for a variety of future studies. Crosses with lines deficient in various DNA repair or tolerance pathways

could create progeny with enhanced sensitivity of reporting reversion. New insights into the effect of a given repair pathway on the repair of specific lesions could also be ascertained. Potential crosses of interest include lines deficient in (or overexpressing) ROS repair, MMR, and TLS. One of the more striking results was the high frequency of G → T reversion, which may reflect 8-oxoG-induced mutation. Overexpression of the *Arabidopsis* homologs for OGG1 or MutM, which are responsible for the removal of 8-oxoG, should lower the cellular levels of 8-oxoG, perhaps resulting in lower spontaneous mutation frequencies in the M4 (G → T) lines. Mismatch repair is responsible for correcting misinsertions during DNA replication, and the loss of MMR activity increases mainly transitions and frameshifts. To follow-up on our previous work [69], it would be interesting to see how the reversion frequencies of the point-mutation lines would respond to an AtMSH2-deficient background. Crosses of the reporter lines with lines deficient in various translesion polymerases would also yield interesting progeny, whose induction of specific point mutations could be used to identify TLS target lesions *in vivo*. Deficiency in TLS could potentially increase the sensitivity of the reporter lines, such as an increase in UV-induced T → C and C → T reversion with pol η deficiency.

The reversion-reporter lines are also useful for assays of known and potential mutagens to study their mutation spectra *in vivo*. Previous mutation-reporter assays, such as the Kovalchuk/Hohn reporter assays, could be corroborated and all point-mutation pathways could be observed for a given compound. For example, this is the first known plant reporter for G → T reversion, which is expected to be highly

responsive to reactive oxygen species. We have already been able to show a very high spontaneous level of reversion, likely due to endogenous DNA oxidation. We have also documented relatively high levels of potential germinal G → T mutations. This reporter line would therefore be ideal for identifying mutagens that induce oxidative damage of DNA.

Use of the reversion-reporter lines for biomonitoring environmental damage is also possible. The reporter lines could be treated with a variety of mutagenic compounds and analyzed to determine which reversions were induced. The reporter set could then be used to assay complex soil mixtures containing multiple damaging agents. Analysis of the observed mutation spectra may provide insights as to what the contributing mutagens in the soil are. Since plants are able to take up and concentrate even trace amounts of chemicals, they may be ideal for detecting rare mutagens. The reporter lines could be used for monitoring sites of environmental damage clean-up by tracking the decline of mutation induction. Assays using the reversion reporter lines are fast, simple, and fairly inexpensive compared to previous plant biomonitors.

Among the different lines tested, we observed variation between the different sublines and between different generations of the same subline. Different frequencies were also observed in comparison to previous reporters, such as the Kovalchuk/Hohn lines. It is unclear why such variation exists, and more extensive studies with reporters such as these need to be performed to gain a clearer picture of the factors involved in line-to-line and generation-to-generation mutation variation. There was no obvious correlation between total protein expression and mutation frequency among

different lines. Protein expression remained fairly constant between generations (Figure 29), even when there were significant changes in observed mutation frequency. An intriguing study would be an extensive analysis of mutation frequency across several generations with and without transgene silencing capabilities. Reversion frequencies would be measured with or without mutagen treatment and compared across several sublines from each of the reporters, enabling a better picture of the relative stability of the mutation reporters. The various generations of reporter lines could also be analyzed for factors that contribute to alterations in reversion frequency, such as transgene silencing factors or alterations in chromatin state. Overall, the reporter system defined by this study has great potential for point mutation analysis.

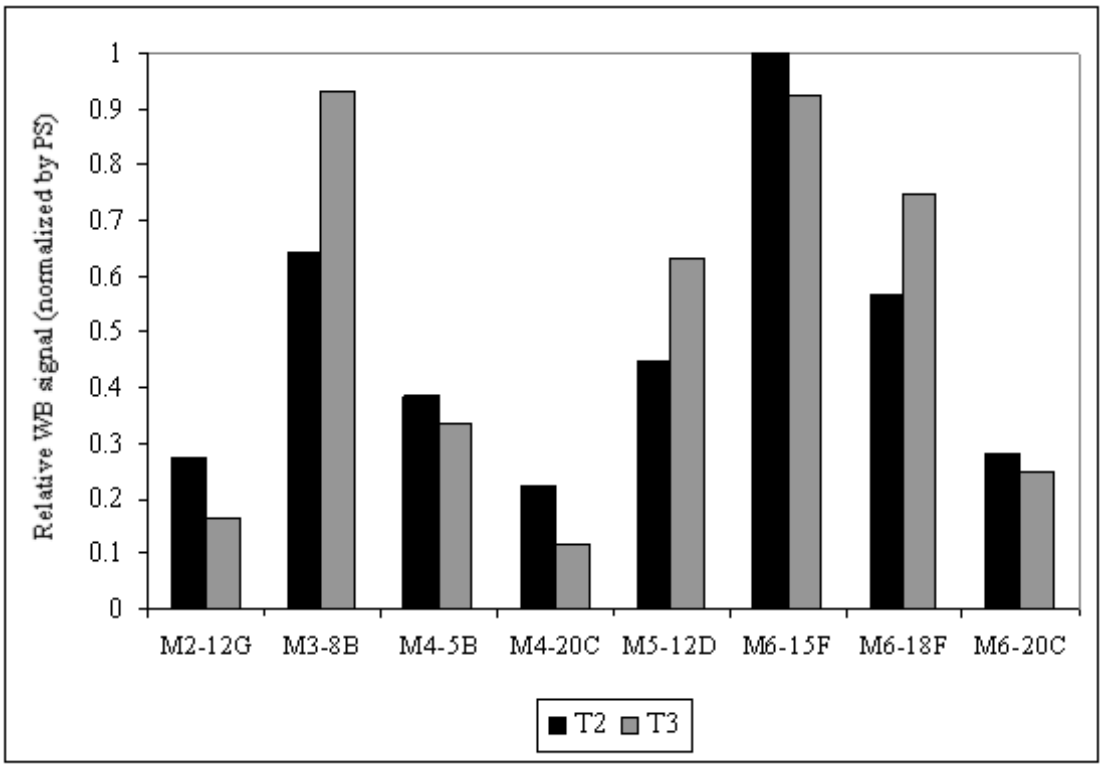


Figure 29. GUS protein expression in successive generations. Levels of GUS-protein expression in indicated T_2 and T_3 lines were determined as described in Figure 12 legend.

Bibliography

1. Britt, A.B., *DNA Damage and Repair in Plants.* Annu Rev Plant Physiol Plant Mol Biol, 1996. **47**: p. 75-100.
2. Kovalchuk, I., O. Kovalchuk, and B. Hohn, *Biomonitoring the genotoxicity of environmental factors with transgenic plants.* Trends Plant Sci, 2001. **6**(7): p. 306-10.
3. Yoshihara, R., C. Nakane, and K. Takimoto, *A new system for detecting mutations in arabidopsis thaliana and the mutational spectra resulting from ethylmethanesulfonate treatment.* J Radiat Res (Tokyo), 2006. **47**(3-4): p. 223-8.
4. Hartwell, L., et al., *Genetics: from genes to genomes*. 1 ed. 2000: McGraw-Hill Higher Education.
5. Bray, C.M. and C.E. West, *DNA repair mechanisms in plants: crucial sensors and effectors for the maintenance of genome integrity.* New Phytol, 2005. **168**(3): p. 511-28.
6. Kimura, S. and K. Sakaguchi, *DNA repair in plants.* Chem Rev, 2006. **106**(2): p. 753-66.
7. Barone, F., et al., *Replication of 2-hydroxyadenine-containing DNA and recognition by human MutSalpha.* DNA Repair (Amst), 2007. **6**(3): p. 355-66.
8. Skinner, A.M. and M.S. Turker, *Oxidative mutagenesis, mismatch repair, and aging.* Sci Aging Knowledge Environ, 2005. **2005**(9): p. re3.
9. Kovalchuk, O., et al., *A sensitive transgenic plant system to detect toxic inorganic compounds in the environment.* Nat Biotechnol, 2001. **19**(6): p. 568-72.
10. Clemens, S., *Toxic metal accumulation, responses to exposure and mechanisms of tolerance in plants.* Biochimie, 2006. **88**(11): p. 1707-19.
11. Sanita di Toppi, L. and R. Gabbrielli, *Response to cadmium in higher plants.* Environmental and Experimental Botany, 1999. **41**: p. 105-130.
12. Valko, M., et al., *Free radicals, metals and antioxidants in oxidative stress-induced cancer.* Chem Biol Interact, 2006. **160**(1): p. 1-40.
13. Hartwig, A., *Current aspects in metal genotoxicity.* Biometals, 1995. **8**(1): p. 3-11.
14. Filipic, M., T. Fatur, and M. Vudrag, *Molecular mechanisms of cadmium induced mutagenicity.* Hum Exp Toxicol, 2006. **25**(2): p. 67-77.
15. Sirover, M.A. and L.A. Loeb, *Infidelity of DNA synthesis in vitro: screening for potential metal mutagens or carcinogens.* Science, 1976. **194**(4272): p. 1434-6.
16. Fatur, T., T.T. Lah, and M. Filipic, *Cadmium inhibits repair of UV-, methyl methanesulfonate- and N-methyl-N-nitrosourea-induced DNA damage in Chinese hamster ovary cells.* Mutat Res, 2003. **529**(1-2): p. 109-16.
17. Jin, Y.H., et al., *Cadmium is a mutagen that acts by inhibiting mismatch repair.* Nat Genet, 2003. **34**(3): p. 326-9.

18. Britt, A.B., *Molecular genetics of DNA repair in higher plants.* Trends Plant Sci, 1999. **4**(1): p. 20-25.
19. Chiu, R.K., et al., *Lysine 63-polyubiquitination guards against translesion synthesis-induced mutations.* PLoS Genet, 2006. **2**(7): p. e116.
20. Douki, T. and J. Cadet, *Individual determination of the yield of the main UV-induced dimeric pyrimidine photoproducts in DNA suggests a high mutagenicity of CC photolesions.* Biochemistry, 2001. **40**(8): p. 2495-501.
21. Mouret, S., et al., *Differential repair of UVB-induced cyclobutane pyrimidine dimers in cultured human skin cells and whole human skin.* DNA Repair (Amst), 2008.
22. Nijhof, J.G., et al., *Epidermal stem and progenitor cells in murine epidermis accumulate UV damage despite NER proficiency.* Carcinogenesis, 2007. **28**(4): p. 792-800.
23. Tornaletti, S. and G.P. Pfeifer, *UV damage and repair mechanisms in mammalian cells.* Bioessays, 1996. **18**(3): p. 221-8.
24. Kovalchuk, I., et al., *Molecular aspects of plant adaptation to life in the Chernobyl zone.* Plant Physiol, 2004. **135**(1): p. 357-63.
25. Dany, A.L., et al., *Repair of the main UV-induced thymine dimeric lesions within Arabidopsis thaliana DNA: evidence for the major involvement of photoreactivation pathways.* J Photochem Photobiol B, 2001. **65**(2-3): p. 127-35.
26. Sarasin, A., *The molecular pathways of ultraviolet-induced carcinogenesis.* Mutat Res, 1999. **428**(1-2): p. 5-10.
27. Van der Auwera, G., et al., *Development and application of novel constructs to score C:G-to-T:A transitions and homologous recombination in Arabidopsis.* Plant Physiol, 2008. **146**(1): p. 22-31.
28. Levin, D.E. and B.N. Ames, *Classifying mutagens as to their specificity in causing the six possible transitions and transversions: a simple analysis using the Salmonella mutagenicity assay.* Environ Mutagen, 1986. **8**(1): p. 9-28.
29. Cupples, C.G. and J.H. Miller, *A set of lacZ mutations in Escherichia coli that allow rapid detection of each of the six base substitutions.* Proc Natl Acad Sci U S A, 1989. **86**(14): p. 5345-9.
30. Grant, W.F., *The present status of higher plant bioassays for the detection of environmental mutagens.* Mutat Res, 1994. **310**(2): p. 175-85.
31. Grant, W.F., *Higher plant assays for the detection of chromosomal aberrations and gene mutations-a brief historical background on their use for screening and monitoring environmental chemicals.* Mutat Res, 1999. **426**(2): p. 107-12.
32. Schairer, L.A., et al., *Exploratory monitoring of air pollutants for mutagenicity activity with the Tradescantia stamen hair system.* Environ Health Perspect, 1978. **27**: p. 51-60.
33. Knasmüller, S., Gottmann, E., Steinkellner, H., Fomin, A., Pickl, C., Paschke, A., Göd, R. and Kundi, M. , *Detection of genotoxic effects of heavy metal contaminated soils with plant bioassays.* Mutation Research, 1998. **420**: p. 37-48.

34. Steinkellner, H., et al., *Genotoxic effects of heavy metals: comparative investigation with plant bioassays.* Environ Mol Mutagen, 1998. **31**(2): p. 183-91.
35. Nohmi, T., T. Suzuki, and K. Masumura, *Recent advances in the protocols of transgenic mouse mutation assays.* Mutat Res, 2000. **455**(1-2): p. 191-215.
36. Amanuma, K., et al., *Transgenic zebrafish for detecting mutations caused by compounds in aquatic environments.* Nat Biotechnol, 2000. **18**(1): p. 62-5.
37. Jefferson, R.A., T.A. Kavanagh, and M.W. Bevan, *GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants.* Embo J, 1987. **6**(13): p. 3901-7.
38. Novel, G. and M. Novel, *[Mutants of E. coli K 12 unable to grow on methyl-beta-D-glucuronide: map l ocation of uid A. locus of the structural gene of beta-D-glucuronidase].* Mol Gen Genet, 1973. **120**(4): p. 319-35.
39. Sudan, C., et al., *Ubiquitous presence of beta-glucuronidase (GUS) in plants and its regulation in some model plants.* Planta, 2006. **224**(4): p. 853-64.
40. Kovalchuk, I., O. Kovalchuk, and B. Hohn, *Genome-wide variation of the somatic mutation frequency in transgenic plants.* Embo J, 2000. **19**(17): p. 4431-8.
41. Kovalchuk, I., et al., *Transgenic plants are sensitive bioindicators of nuclear pollution caused by the Chernobyl accident.* Nat Biotechnol, 1998. **16**(11): p. 1054-9.
42. Filkowski, J., et al., *Genotoxicity of 2,4-D and dicamba revealed by transgenic Arabidopsis thaliana plants harboring recombination and point mutation markers.* Mutat Res, 2003. **542**(1-2): p. 23-32.
43. Swoboda, P., et al., *Intrachromosomal homologous recombination in whole plants.* Embo J, 1994. **13**(2): p. 484-9.
44. Boyko, A., J. Filkowski, and I. Kovalchuk, *Homologous recombination in plants is temperature and day-length dependent.* Mutat Res, 2005. **572**(1-2): p. 73-83.
45. Schultze, M., T. Hohn, and J. Jiricny, *The reverse transcriptase gene of cauliflower mosaic virus is translated separately from the capsid gene.* Embo J, 1990. **9**(4): p. 1177-85.
46. Kozak, M., *Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes.* Cell, 1986. **44**(2): p. 283-92.
47. Chen, P.Y., et al., *Complete sequence of the binary vector pBI121 and its application in cloning T-DNA insertion from transgenic plants.* Mol Breed, 2003. **11**: p. 287-293.
48. Schubert, D., et al., *Silencing in Arabidopsis T-DNA transformants: the predominant role of a gene-specific RNA sensing mechanism versus position effects.* Plant Cell, 2004. **16**(10): p. 2561-72.
49. De Buck, S., et al., *Single-copy T-DNAs integrated at different positions in the Arabidopsis genome display uniform and comparable beta-glucuronidase accumulation levels.* Cell Mol Life Sci, 2004. **61**(19-20): p. 2632-45.

50. Nagaya, S., et al., *Expression of randomly integrated single complete copy transgenes does not vary in Arabidopsis thaliana.* Plant Cell Physiol, 2005. **46**(3): p. 438-44.
51. Gelvin, S.B., *Agrobacterium-mediated plant transformation: the biology behind the "gene-jockeying" tool.* Microbiol Mol Biol Rev, 2003. **67**(1): p. 16-37, table of contents.
52. Dunoyer, P., C. Himber, and O. Voinnet, *Induction, suppression and requirement of RNA silencing pathways in virulent Agrobacterium tumefaciens infections.* Nat Genet, 2006. **38**(2): p. 258-63.
53. Meza, T.J., et al., *The frequency of silencing in Arabidopsis thaliana varies highly between progeny of siblings and can be influenced by environmental factors.* Transgenic Res, 2001. **10**(1): p. 53-67.
54. Qin, H., Y. Dong, and A.G. von Arnim, *Epigenetic interactions between Arabidopsis transgenes: characterization in light of transgene integration sites.* Plant Mol Biol, 2003. **52**(1): p. 217-31.
55. Lee, L.Y., et al., *Novel plant transformation vectors containing the superpromoter.* Plant Physiol, 2007. **145**(4): p. 1294-300.
56. Ni, M., et al., *Strength and tissue specificity of chimeric promoters derived from the octopine and mannopine synthase genes.* Plant J, 1995. **7**(4): p. 661-676.
57. Leisner, S.M. and S.B. Gelvin, *Structure of the octopine synthase upstream activator sequence.* Proc Natl Acad Sci U S A, 1988. **85**(8): p. 2553-7.
58. Becker, D., et al., *New plant binary vectors with selectable markers located proximal to the left T-DNA border.* Plant Mol Biol, 1992. **20**(6): p. 1195-7.
59. Carrington, J.C. and D.D. Freed, *Cap-independent enhancement of translation by a plant potyvirus 5' nontranslated region.* J Virol, 1990. **64**(4): p. 1590-7.
60. Day, C.D., et al., *Transgene integration into the same chromosome location can produce alleles that express at a predictable level, or alleles that are differentially silenced.* Genes Dev, 2000. **14**(22): p. 2869-80.
61. Medberry, S.L., B.E. Lockhart, and N.E. Olszewski, *The Commelina yellow mottle virus promoter is a strong promoter in vascular and reproductive tissues.* Plant Cell, 1992. **4**(2): p. 185-92.
62. Hensgens, L.A., et al., *Translation controls the expression level of a chimaeric reporter gene.* Plant Mol Biol, 1992. **20**(5): p. 921-38.
63. Guevara-Garcia, A., et al., *A 42 bp fragment of the pmas1' promoter containing an ocs-like element confers a developmental, wound- and chemically inducible expression pattern.* Plant Mol Biol, 1998. **38**(5): p. 743-53.
64. De Bolle, M.F.C., Butaye, K.M.J., Coucke, W.J.W., Goderis, I.J.W.M., Wouters, P.F.J., van Boxel, N., Broekaert, W.F. and Cammue, B.P.A. , *Analysis of the influence of promoter elements and a matrix attachment region on the inter-individual variation of transgene expression in populations of Arabidopsis thaliana.* Plant Science, 2003. **165**: p. 169-179.

65. Butaye, K.M., et al., *Stable high-level transgene expression in Arabidopsis thaliana using gene silencing mutants and matrix attachment regions.* Plant J, 2004. **39**(3): p. 440-9.
66. Lawrence, S.D., N.G. Novak, and J.M. Slack, *Epitope tagging: a monoclonal antibody specific for recombinant fusion proteins in plants.* Biotechniques, 2003. **35**(3): p. 488-92.
67. Fan, H., et al., *Myc-epitope tagged proteins detected with the 9E10 antibody in immunofluorescence and immunoprecipitation assays but not in western blot analysis.* Biochem Cell Biol, 1998. **76**(1): p. 125-8.
68. Bannon, L.J., M.S. Stack, and K.J. Green, *Limitations of comparative detection of proteins via epitope tagging.* Anal Biochem, 2001. **293**(1): p. 139-42.
69. Leonard, J.M., S.R. Bollmann, and J.B. Hays, *Reduction of stability of Arabidopsis genomic and transgenic DNA-repeat sequences (microsatellites) by inactivation of AtMSH2 mismatch-repair function.* Plant Physiol, 2003. **133**(1): p. 328-38.
70. Hoffman, P.D., et al., *Rapid accumulation of mutations during seed-to-seed propagation of mismatch-repair-defective Arabidopsis.* Genes Dev, 2004. **18**(21): p. 2676-85.
71. Ausubel, F.M., et al., eds. *Current Protocols in Molecular Biology*. ed. V.B. Chanda. Vol. 1-4. 2005, John Wiley & Sons, Inc.
72. Puchta, H. and B. Hohn, *A transient assay in plant cells reveals a positive correlation between extrachromosomal recombination rates and length of homologous overlap.* Nucleic Acids Res, 1991. **19**(10): p. 2693-700.
73. Koncz, C. and J. Schell, *The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of Agrobacterium binary vector.* Mol Gen Genet, 1986. **204**(3): p. 383-396.
74. Clough, S.J. and A.F. Bent, *Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana.* Plant J, 1998. **16**(6): p. 735-43.
75. Ream, W. and K.G. Field, *Molecular Biology Techniques: an Intensive Laboratory Course*. 1999, San Diego: Academic Press.
76. Jefferson, R.A., *Assaying chimeric genes in plants: the GUS gene fusion system.* Plant Molecular Biology Reporter, 1987. **5**: p. 387-405.
77. Filkowski, J., O. Kovalchuk, and I. Kovalchuk, *Genome stability of vtc1, tt4, and tt5 Arabidopsis thaliana mutants impaired in protection against oxidative stress.* Plant J, 2004. **38**(1): p. 60-9.
78. Liu, Y.G. and R.F. Whittier, *Rapid preparation of megabase plant DNA from nuclei in agarose plugs and microbeads.* Nucleic Acids Res, 1994. **22**(11): p. 2168-9.
79. Brown, T.A., ed. *Molecular Biology Labfax*. ed. B.D. Hames and D. Rickwood. 1991, Academic Press, Inc.: San Diego.
80. Curtis, M.J. and J.B. Hays, *Tolerance of dividing cells to replication stress in UVB-irradiated Arabidopsis roots: requirements for DNA translesion polymerases eta and zeta.* DNA Repair (Amst), 2007. **6**(9): p. 1341-58.

81. Chen, J.J., D.L. Mitchell, and A.B. Britt, *A Light-Dependent Pathway for the Elimination of UV-Induced Pyrimidine (6-4) Pyrimidinone Photoproducts in Arabidopsis.* Plant Cell, 1994. **6**(9): p. 1311-1317.
82. Gordon, L.K. and W.A. Haseltine, *Quantitation of cyclobutane pyrimidine dimer formation in double- and single-stranded DNA fragments of defined sequence.* Radiat Res, 1982. **89**(1): p. 99-112.
83. Van Belleghem, F., et al., *Subcellular localization of cadmium in roots and leaves of Arabidopsis thaliana.* New Phytol, 2007. **173**(3): p. 495-508.

APPENDIX

Appendix. Raw data.

Subline	Gen.	Growth medium	Treatment	Light: dark	Total # plants	Total # zeroes	Total # events	p value[a]
M1(T→G)								
1A	T_2	Soil		24:0[d]	49	49	0	
		Soil	UV-B[b]	24:0	46	45	1	0.429
		Soil	UV-C[c]	24:0	42	40	2	0.348
1E	T_2	Soil		24:0	43	43	0	
		Soil	UV-B	24:0	47	47	0	0.500
		Soil	UV-C	24:0	45	42	4	0.298
3F	T_2	Soil		24:0	31	30	1	
		Soil	UV-B	24:0	28	28	0	0.421
		Soil	UV-C	24:0	28	26	2	0.401
6A	T_3	Soil		24:0	35	35	0	
		Soil	UV-B	24:0	31	30	1	0.413
		Soil	UV-C	24:0	31	30	1	0.413
11A	T_2	Soil		24:0	46	46	0	
		Soil	UV-B	24:0	45	45	0	0.500
		Soil	UV-C	24:0	47	43	4	0.242
11D	T_2	Soil		24:0	47	47	0	
		Soil	UV-B	24:0	47	42	5	0.189
		Soil	UV-C	24:0	46	42	4	0.236
12D	T_2	Soil		24:0	32	31	1	
		Soil	UV-B	24:0	24	20	4	0.198
		Soil	UV-C	24:0	36	33	3	0.359
M2(C→T)								
8D	T_2	Soil		16:8[e]	114	110	4	
		Soil	UV-C	16:8	115	96	21	0.044[i]
10C[m]	T_2	Soil		16:8	90	90	0	
		Soil		24:0	144	143	1	
		Soil	UV-C	16:8	87	59	37	<.0001[i]
		Soil	UV-C	16:8[f]	86	78	10	0.145
12G[m]	T_2	Soil		24:0	132	131	1	
		Soil	UV-C	16:8	91	78	14	0.043[i]
		Soil	UV-C	16:8[f]	92	88	4	0.323
	T_3	Soil		16:8	74	66	9	
		Soil	UV-C	16:8	81	71	15	0.425
		Soil		24:0	89	87	2	
		Soil	UV-C	24:0	73	55	30	0.007[i]
15G	T_3	Soil		16:8	100	98	2	
		Soil	UV-C	16:8	98	82	16	0.041[i]
20B	T_3	Soil		16:8	81	77	4	
		Soil	UV-C	16:8	96	76	24	0.033[i]
p1804-3G	T_2	Soil		16:8	139	123	23	
		Soil	UV-C	16:8	137	74	92	<.0001[i]

Appendix (Continued). Raw data.

Subline	Gen.	Growth medium	Treatment	Light: dark	Total # plants	Total # zeroes	Total # events	p value
M3(G→C)								
1G	T_2	MS plate		24:0	152	145	7	
		MS plate	1 mg/L Cd^{2+} [g]	24:0	149	139	10	0.749
8B	T_2	MS plate		24:0	148	143	5	
		MS plate	1 mg/L Cd^{2+} [g]	24:0	138	132	7	0.889
M4(G→T)								
5B	T_2	MS plate		24:0	146	55	226	
		MS plate	.4 mg/L Cd^{2+}	24:0	146	67	149	0.027 [i]
		MS plate	1 mg/L Cd^{2+} [g]	24:0	142	20	308	0.0001 [i]
		MS plate	1 mg/L Cd^{2+} [h]	24:0	99	47	102	0.048 [i]
		MS plate	6 mg/L Zn^{2+}	24:0	137	82	83	<.0001 [i]
		MS plate	18 mg/L Zn^{2+}	24:0	145	109	57	<.0001 [i]
		Soil		16:8	87	46	79	0.011 [i,j]
	T_3	Soil		16:8	89	55	41	0.067 [k]
		Soil		24:0	89	78	12	0.002 [i,l]
7D[m]	T_2	MS plate		24:0	142	121	25	
9F	T_2	MS plate		24:0	64	7	133	
14A[m]	T_2	MS plate		24:0	136	94	53	
20C	T_2	MS plate		24:0	131	40	168	
		MS plate	1 mg/L Cd^{2+} [g]	24:0	117	31	158	0.363
M5(A→T)								
12D[m]	T_2	Soil		16:8	58	55	3	
		Soil	UV-C	16:8	79	68	11	0.192
	T_3	Soil		16:8	93	87	6	
		Soil	UV-C	16:8	125	109	19	0.181
M6(T→C)								
5C	T_2	Soil		24:0	115	115	0	
		Soil	UV-C	24:0	116	112	4	0.326
6B	T_2	Soil		24:0	145	144	1	
		Soil	UV-C	24:0	139	136	3	0.417
15F[m]	T_2	Soil		24:0	123	122	1	
		Soil		24:0[f]	81	81	0	
		Soil	UV-C	16:8	74	67	7	0.156
		Soil	UV-C	24:0	87	79	10	0.152
	T_3	Soil		24:0	94	94	0	
		Soil	UV-C	24:0	74	63	12	0.049 [i]
18F	T_2	Soil		24:0	122	122	0	
		Soil	UV-C	24:0	130	49	135	<.0001 [i]
	T_3	Soil		16:8	103	103	0	
		Soil	UV-C	16:8	89	79	14	0.090
		Soil		24:0	102	102	0	
		Soil	UV-C	24:0	98	93	5	0.268

Appendix (Continued). Raw data.

Subline	Gen.	Growth medium	Treatment	Light: dark	Total # plants	Total # zeroes	Total # events	p value
M6(T→C)								
18F	T_3	MS plate	1 mg/L Cd^{2+} [h]	24:0	105	104	1	
18G	T_2	Soil		24:0	39	39	0	
		Soil	UV-C	24:0	41	35	10	0.131
20C	T_2	Soil		24:0	127	127	0	
		Soil	UV-C	24:0	108	65	65	<.0001 [i]
	T_3	Soil		24:0	88	87	1	
		Soil	UV-C	24:0	86	79	8	0.212
21C	T_2	Soil		24:0	132	132	0	
		Soil	UV-C	24:0	131	125	6	0.261
22B	T_2	Soil		24:0	47	47	0	
		Soil	UV-C	24:0	36	25	12	0.009[i]
24B	T_2	Soil		24:0	71	71	0	
		Soil	UV-C	24:0	55	50	6	0.192

[a] Mann-Whitney calculated p-value (UV comparisons use a one-tail test, others use a two-tail test).
[b] Three doses of 5000 J/m^2 UV-B.
[c] One dose of 1000 J/m^2 UV-C.
[d] Plants grown under continuous light.
[e] Plants grown under photoperiod light.
[f] Repeated experiment.
[g] Cadmium chloride added to final 1 mg/L before autoclaving.
[h] Cadmium chloride added to final 1 mg/L after autoclaving.
[i] Calculated p-value is statistically significant (<0.05).
[j] Comparison of MS plate to soil growth conditions.
[k] Comparison of T_2 to T_3 generation.
[l] Comparison of continuous to photoperiod day length.
[m] Sequence analysis confirmed the presence of the correct transgene.