Saturation mutagenesis using maize transposons
Virginia Walbot

Transposon mutagenesis facilitates gene discovery by tagging genes for cloning. New genomics projects are now cataloging transposon insertion sites to define all maize genes. Once identified, transposon insertions are ‘hot spots’ for generating new alleles that are useful in functional studies.

Addresses
Department of Biological Sciences, 385 Serra Mall, Stanford University, Stanford, California 94305-5020, USA; e-mail: walbot@stanford.edu

Current Opinion in Plant Biology 2000, 3:103–107
1369-5266/00/$ – see front matter © 2000 Elsevier Science Ltd. All rights reserved.

Abbreviations
bp base pairs
MTM Maize Targeted Mutagenesis
NSF National Science Foundation
TIR terminal inverted repeat
TUSC Trait Utility System for Corn

Introduction
The analysis of mutant phenotypes yields an understanding of gene function in a whole-organism context. Consequently, the generation, evaluation, maintenance and distribution of seed containing verified mutations or seed with populations of mutations suitable for high-throughput screening are essential in the modern genetic analysis of maize and other plants.

Saturation mutagenesis of maize genes involves two ambitious goals: to define all of the genes and to find the phenotype of every individual gene. This review will outline how transposon mutagenesis can be used to achieve both goals. Transposons are already the primary tool for tagging and cloning maize genes. Studies so far demonstrate that specific transposons are well suited for either global mutagenesis and gene discovery or multiple rounds of mutagenesis at a defined target gene [1]. New genomics approaches, employing strategies for screening by PCR and for plasmid rescue, are now providing indexed collections of mutations and the sequences flanking transposon insertion sites. Many researchers will soon identify transposon-generated mutants in specific genes after querying a database rather than searching a corn field.

Insertional mutagenesis across the genome: primer on key transposon properties
The Ac/Ds and MuDR/Mu maize transposons have been used in most recent mutagenesis experiments [1–3]. In stocks with the transcriptionally active regulatory elements Ac or MuDR, family members with essential transposase-binding sites in the terminal inverted repeats (TIRs) are mobilized to create new insertion mutations. During a gene-tagging experiment, transposon activity is conveniently monitored with a reporter allele that has a phenotype that becomes visible after the excision of a transposon (Figure 1a). Saturation mutagenesis of all genes requires a robust mutagenized population. An estimate of the distinct units (genes, for example) and the confidence level of the coverage are used to calculate the population size required to provide one insertion per unit, assuming random mutagenesis (Table 1).

Ac/Ds are ‘cut and paste’ transposons (Figure 1b). New insertions can occur throughout the life cycle: somatic insertions that occur in the apical meristem can be apparent in the shoot and can be transmitted to the next generation [3]. Most newly identified mutations, however, are recovered in single progeny, indicating that, in practice, insertions in gametes are the main source of new mutations. Germinal revertants are readily recovered (at a frequency of 10–4 to 10–5), indicative of late somatic or gametophytic excision events. These germinal revertants often differ in sequence [4], and are a rich source of allelic diversity for functional analysis [5,6], as they have been over evolutionary time [7]. Because only a few mobile Ac/Ds exist in a mutagenic line and these elements insert preferentially into sites that are closely linked to the original element locations [3], Ac/Ds is an inefficient global mutagen (with a forward mutation frequency of 10–6 [1]) but can be highly efficient regionally. There are also maize transposons with a high insertion preference for particular genes [8]: these are a restricted tool but one that is highly useful for those studying the target(s).

Ac/Ds elements spaced at 10–20 cM intervals across the recombination map would facilitate intensive mutagenesis of the entire maize genome [9]. Because mobile Ac/Ds copy number is low, confirmation that a particular transposon insert is the cause of a mutant phenotype is easily obtained [3]. There are several strategies for exploiting Ac/Ds in maize that depend on enriching for transposed copies, which may be linked or unlinked to the original location [3] (Figure 2).
Mutator stocks typically contain several copies of mutation frequency (~10⁻⁴, range 10⁻³–10⁻⁵ in any gene and many 20–200 thousand base pairs (kbp) between maize genes found in complex arrays; these create the spacers containing genome, however, consists of inactive retrotransposons [1,16]). When engineered dicots expressing maize transposons [10–15]; with genetical-These strategies have also been implemented in transgenic events.

**Mu** insertion do occur preferentially into low-copy DNA; there is no apparent local bias [16]. Recent analyses of hundreds of somatic (M Raizada, V Walbot, unpublished data) and germinal Mu element insertions (K Edwards, personal communication) confirm insertion into or near exon-like sequences. Although high copy numbers make MuDR/Mu efficient mutagens, the resulting multiple mutant genes complicate the assignment of phenotypes to specific genes [2].

Transposon tagging in maize is much more efficient than would be predicted from the large genome size, ~2.3 × 10⁹ base pairs (bp) [18]. More than half of the maize genome, however, consists of inactive retrotransposons found in complex arrays; these create the spacers containing 20–200 thousand base pairs (kbp) between maize genes [19]; consequently, insertions would be expected to affect only one gene. With short introns, maize genes are as compact as those of Arabidopsis [20]. Thus, transcription, recombination [21] and DNA transposon insertion are all highly biased for the genes, and for these important processes absolute genome size is of relatively little importance.

**DNA sequencing strategies for genomic DNA next to transposon insertions: methods used to analyze transposons in a single plant**

Initial applications of transposon tagging in maize relied on correlating the inheritance of a plant phenotype with a band on a DNA hybridization blot. Particularly for the multi-copy Mu elements, demonstrating tight linkage required some luck in restriction enzyme choice and in hybridization probes to resolve specific Mu elements (Mu1, Mu2, Mu3, Mu8 and MuDR); the Mu TIR probe detects too many bands to be useful. Proof that the correct gene was cloned after recovery of the transposon-tagged genomic fragment requires the analysis of independent mutants, or a complementation test. Given the difficulty of maize transformation, complementation has so far been useful only for traits that can be assayed in transient expression assays with protoplasts or tissues amenable to particle gun bombardment [22].

Several techniques for amplifying and sequencing genomic DNA next to transposon ends are amenable to genomics approaches. PCR primers anchored on transposons are ‘read out’ into genomic DNA [23,24,25*] with a wide choice of strategies for priming from the flanking genomic DNA of unknown or specified sequence (Figure 2) [26*]. Initial methods relied on the separation of the PCR products by size, with manual recovery and analysis of bands that segregated with a phenotype. Nowadays, simplifications include
Identifying transposon insertions. (a) Enriching for transposed Ac/Ds. In this diagram, kernels with excision sectors are white with spots of gray. Early somatic excisions that restore gene function result in a sector of gray kernels. By selecting seeds in a sector of kernels with a functionally revertant allele at the reporter gene (gray circles) investigators identify a somatic excision event from the original site; these kernels should contain the same transposed Ac at a new location in the genome. Because Ac shows a negative dosage effect on the frequency of somatic excision, kernels with fewer spots probably contain both the original donor site and a new transposed Ac. (b) The amplification of sequences next to transposons typically uses a ‘read out’ primer (arrows) complementary to the TIRs and a strategy to allow priming from the contiguous genomic DNA (star) such as a restriction site cloning combined with high-throughput sequencing. Selective PCR primer strategies can examine a subset of transposon family members or a subset of insertion sites (with primers spanning the joint with genomic DNA). Because transposon insertions result in characteristic host sequence duplications (9 bp for Mu [16], 8 bp for Ac/Ds [3]), compiling the continuous DNA sequence at any given insertion site uses the duplication for matching up the left and right genomic sequences.

‘Panhandle’ PCR, in which transposon TIRs form intramolecular duplexes in single-stranded DNA [26••] (Figure 2), is the foundation for a new National Science Foundation (NSF)-funded genomics project that aims to recover and sequence maize genomic DNA adjacent to transposed Ac [9•]. Plasmid rescue is an alternative strategy that is much simpler to use in practice but requires more effort to set up. Transgenic plants with transposons engineered to contain a bacterial plasmid allow the selective cloning of transposon insertion sites after transformation of Escherichia coli with total genomic DNA; only plant DNA with a bacterial origin of replication and an antibiotic resistance marker is maintained in the transformed bacteria. This approach was pioneered by using Ac/Ds-based elements introduced into dicots via Agrobacterium-mediated transformation [27]; now RescueMu plasmids based on MuI are being used to clone maize genes [28•].

### Indexed mutant collections and parallel searches for transposon insertions

Traditional genetics starts with the mutant individual and seeks the gene, whereas genomics methods start with the DNA sequence and search for the phenotypes. The first genomics approach, TUSC (Trait Utility System for Corn), was developed by Pioneer Hi-Bred International; it is available to academic researchers who sign material transfer agreements to receive the seeds containing putative ‘hits’ to user-supplied sequence motifs. TUSC contains DNA samples and progeny seed from ~44,000 Mutator plants that probably contain >10⁹ independent Mu insertions. DNA samples, from single plants or post-preparation pools, are screened for Mu insertions into a target gene or motif by using a PCR primer that reads out of Mu TIRs and a target-specific primer. All mobile Mu elements share a very highly conserved sequence in the terminal 25 bp of the TIRs [16]; consequently, a single ‘read out’ primer suffices. TUSC has been used successfully by both Pioneer Hi-Bred International [29] and by academic researchers [30,31]. One drawback of a PCR strategy with a high cycle number is that both germinal and somatic insertions are found; ~10–20% genuine germinal insertions are verified among candidate insertions from TUSC screening [28•].

A convenient field organization into rows and columns for collecting pooled plant (and hence DNA) samples was implemented at about the same time as TUSC [32]; this method overcomes the identification of ‘false positive’ somatic insertions. DNA samples representing pooled leaf punches from a row of plants and separate sets of leaf punches from each column of plants are screened by PCR; two successful reactions define the row and the column and hence a specific plant that has the germinal insertion mutation (Figure 3). The Maize Targeted Mutagenesis (MTM) project, funded in the NSF plant genomics program, is now offering screening of a Mutator collection of ~50,000 plants and providing seed for the phenotypic evaluation of ‘hits’ [33•]. Project personnel are evaluating plant phenotypes at the kernel, seedling and adult plant stages to build a database of maize mutant phenotypes. To use MTM services,
investigators supply a gene sequence with intron–exon annotation; after receiving seed, investigators are obligated to contribute phenotypic analysis to the database.

In both TUSC and the row/column pooling method, DNA samples are finite. The NSF-funded Maize Gene Discovery Project uses RescueMu plasmid rescue to create immortalized collections of insertion sites in E. coli [28•]. Tagging fields with 2304 transgenic individuals are organized into 48 rows and 48 columns, generating 96 separate libraries of RescueMu plasmids for each field. Users receive a 96-well plate containing 48 row and 48 column libraries and design their own PCR strategy to identify insertions; germinal events are present in one row and one column (Figure 3). To obtain seed, users submit 1 kb of DNA sequence next to the Mu insertion site of the RescueMu plasmid of interest. Concomitantly, the Maize Gene Discovery Project will sequence row libraries to 95% coverage. Within a few years, users could first search the web Discovery Project will sequence row libraries to 95% coverage. Within a few years, users could first search the web, with small differences in sequence next to the Mu insertion site of the RescueMu plasmid of interest. Concomitantly, the Maize Gene Discovery Project will sequence row libraries to 95% coverage. Within a few years, users could first search the web Discovery Project will sequence row libraries to 95% coverage. Within a few years, users could first search the web, with small differences in sequence next to the Mu insertion site of the RescueMu plasmid of interest. Concomitantly, the Maize Gene Discovery Project will sequence row libraries to 95% coverage. Within a few years, users could first search the web, with small differences in sequence next to the Mu insertion site of the RescueMu plasmid of interest. Concomitantly, the Maize Gene Discovery Project will sequence row libraries to 95% coverage. Within a few years, users could first search the web.

### Saturation mutagenesis within a gene

One mutation, no matter how instructive, is rarely sufficient to lead to a comprehensive understanding of a gene’s function. Most transposon insertions into exons or introns are ‘knockouts’ of gene function; alternative splicing events that use sequences in the ends of the transposon can result in modest expression in a few cases [1]. Transposon insertions, unlike most mutations induced by physical agents or Agrobacterium insertions, are ‘hot spots’ for secondary mutations. Of greatest current use in maize are Ac/Ds and Spm/en mutants, from which germinal revertants are readily recovered; these typically contain one or a few base changes in addition to the host sequence duplication and, more rarely, larger deletions, ‘filler DNA additions’ or rearrangements. Of particular utility is the preference of Ac/Ds for local transposition; an Ac/Ds that transpose nearby can readily transpose back to the target gene, providing many new types of mutant allele [34,35]. For example, the promoter can be saturated with insertions at different sites followed by selection for minor alterations that affect regulation, or the requirements for splicing can be explored from insertion sites in or near the conserved intron motifs.

An important feature of all maize transposons is that they are somatically unstable in the presence of the transposase-encoding element. Consequently, somatic tissue is a mosaic of mutant and revertant cells of many different phenotypes. To produce somatic tissue of a single phenotype, revertants can be selected, but not all of these are ‘knockout’ alleles. Alternatively for Ac/Ds, individuals lacking Ac can be recovered to stop somatic excisions. This strategy is virtually impossible for multi-copy MuDR lines; however, epigenetic silencing of transposons can occur spontaneously and is frequent in Mutator lines [16]. In the case of Mu1 insertions in promoters, methylation after silencing can activate a ‘read out’ promoter in the TIRs, in effect restoring gene expression [36]. Stabilization of a fully mutant phenotype is possible by selecting for deletions, which occur with about 10−2 frequency from the ends of Mu1 elements [37].

### Conclusions

Transposon-induced phenotypes have long provided geneticists with beautiful materials and insights into gene expression, development and chromosome mechanics. Cloned transposons have facilitated gene discovery and cloning. In the genomics era, maize transposons have emerged as the premier method for gene discovery and sequencing, as well as the phenotypic analysis of gene expression in a whole-organism context. Transposons allow simultaneous effort in both phases of genomics, gene discovery and functional studies. Transposons are more efficient than a sequential approach to gene sequencing followed by the design of tools to study gene expression.

### Acknowledgements

I thank Michelle Facette for her comments on the manuscript. Research on Mutator and preparation of this article were supported by grants from the NSF (Maize Gene Discovery; DNA Sequencing and Phenotypic Analysis plant genomics grant 9872657) and the National Institutes of Health (GM49681).

### References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- **of outstanding interest**

A review of amplification and detection strategies for multi-copy insertion elements.


A review of the ways in which transposons are currently used to analyze genes.


Sundaresan V, Springer P, Volpe T, Haward S, Jones JDG, Dean C, Martienssen RA:

A review of amplification and detection strategies for multi-copy insertion elements.


A review of the ways in which transposons are currently used to analyze genes.