1999

Genome reorganization and non-linear transposition in maize

Jianbo Zhang
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd
Part of the Genetics Commons, and the Molecular Biology Commons

Recommended Citation
Zhang, Jianbo, "Genome reorganization and non-linear transposition in maize" (1999). Retrospective Theses and Dissertations. 12189.
https://lib.dr.iastate.edu/rtd/12189

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

**The quality of this reproduction is dependent upon the quality of the copy submitted.** Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6” x 9” black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI®

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600
Genome reorganization and non-linear transposition in maize

by

Jianbo Zhang

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Genetics

Major Professor: Thomas Peterson

Iowa State University
Ames, Iowa
1999
This is to certify that the Doctoral dissertation of

Jianbo Zhang

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Major Professor

Signature was redacted for privacy.

For the Major Program

Signature was redacted for privacy.

For the Graduate College
DEDICATION

To my mother,

and my younger sister.
TABLE OF CONTENTS

CHAPTER 1. GENERAL INTRODUCTION 1
  Pericarp pigmentation and the maize PI gene 1
  Ac/Ds family 9
  Ac/Ds transposition and genome reorganization 17
  Dissertation organization 23
  References 24

CHAPTER 2. GENOME REARRANGEMENTS BY NON-LINEAR TRANSPOSONS IN MAIZE 32
  Abstract 32
  Introduction 33
  Methods and materials 35
  Results and Discussion 36
  Acknowledgment 45
  References 45
  Figure legends 50

CHAPTER 3. NESTED DELETIONS GENERATED BY NON-LINEAR TRANSPOSITION IN MAIZE 60
  Abstract 60
  Introduction 61
  Methods and materials 63
  Results 64
  Discussion 68
  Acknowledgment 71
  References 71
  Figure legends 74

CHAPTER 4. GENERAL SUMMARY 85
  References 91

ACKNOWLEDGMENTS 93
CHAPTER 1. GENERAL INTRODUCTION

Pericarp pigmentation and the maize PI gene

The structure of the maize kernel

The mature maize kernel consists of pericarp, endosperm, and embryo (Figure 1). Like other plants, maize kernel development starts from double fertilization. Two genetically identical sperm cells are introduced into the embryo sac through pollen tube. One of them fuses with the egg cell to form a zygote; the other fuses with the central cell that contains two polar nuclei that are also genetically identical, resulting in a triploid cell. The fertilized egg later develops into the embryo, the fertilized central cell develops into the nutritive endosperm, and the outer layer of endosperm later differentiate into aleurone. The protective tissue enclosing embryo and endosperm in the maize kernel is the pericarp, that is derived from the ovary wall (Kieselbach 1999).

The chromosomal content of each kernel tissue can be deduced from the processes of fertilization and kernel development. Endosperm originates from the fusion of a pollen cell (1n) and a central cell (2n), so it is composed of 3n tissues with 2 chromosome sets from the female parent and one chromosome set from the male parent (pollen). The embryo originates from the fusion of a pollen cell (1n) and an egg (1n), so it is a 2n tissue with one chromosome set from each parent. In contrast, the pericarp is derived from the ovary wall; it is a 2n tissue genetically identical to maternal plant (Kieselbach 1999). Thus the pericarp may have a different genotype from that of the embryo and the endosperm in the same kernel.
Figure 1. The mature maize kernel
Many genes that are expressed in the kernel are not essential for the viability of the plant. The genes involved in phlobaphene pigmentation are a good example, as the viability of the null mutants of these genes is normal. Phenotypic changes that are visible in the kernel are generally heritable, because embryo, endosperm, and pericarp are related by cell lineage. Thus, the maize kernel is a very good tissue for fundamental genetic research.

The biochemical pathway of phlobaphene pigmentation

Phlobaphene and anthocyanin are two major classes of pigments in maize. Anthocyanin pigment can be found in almost all maize tissues, while phlobaphene pigmentation is normally observed only in some maize floral tissues, including pericarp, cob glume, silk, and husk. The biosynthetic pathways of both pigments share three common steps (Figure 2): 1) synthesis of chalcone from 4-coumaroyl-CoA and 3-malonyl-CoA, 2) conversion of chalcone to flavanone, and 3) reduction of the C ring of flavanone (or dihydroflavonol). These three steps are catalyzed by chalcone synthetase, chalcone-flavanon isomerase, and NADPH-dependent reductase, respectively (Styles and Ceska 1989). The reduction of flavanone generates flavan-4-ol that can be converted to phlobaphene by polymerization (Styles and Ceska 1975).

The PI gene is required for phlobaphene synthesis. It encodes a transcription factor that shares high (~70%) homology to animal myb proteins (Grotewold et al. 1991). This transcription activator regulates the expression of the A1, C2, and Chi genes (Grotewold et al. 1994). A1, C2, and Chi encode NADPH-dependent reductase, chalcone synthetase, and chalcone-flavanone isomerase, respectively (Wienand et al. 1986, Grotewold and Peterson 1994, Schwarz-Sommer et al. 1987). Thus, loss of PI gene function would eliminate the
Figure 2. Biosynthetic pathway for flavonoids in maize

4-coumaroyl-CoA + 3-malonyl-CoA

\[ \text{Chalcone Synthase (C2)} \]

\[ \text{Chalcone} \]

\[ \text{Chalcone isomerase (Chi)} \]

\[ \text{Dihydroflavonol} \leftrightarrow \text{Flavanone} \]

\[ \text{Flavan-3,4-Diol} \]

\[ \text{Flavan-4-ol} \]

\[ \text{Flavanone} \]

\[ \text{Reductase (AI)} \]

\[ \text{Polymerization} \]

\[ \text{Flavan-4-ol} \]

\[ \text{Glucosides} \]

\[ \text{Anthocyanin} \]

\[ \text{3-hydroxyanthocyanidin} \]

\[ \text{Phlobaphene} \]
expression of Ai, C2, Chi, and thereby prevents the phlobaphene pigmentation in maize floral tissues (ATHMA and PETERSON, 1991).

The maize PI gene

The PI gene controls the phlobaphene pigmentation of maize floral tissues, including pericarp and cob glume. Alleles of the PI gene are conventionally assigned a two-letter suffix to indicate their expression pattern in pericarp and cob. The common PI alleles are listed in table 1 (ANDERSON 1924, STYLES and CESKA 1989).

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Phenotype</th>
<th>Genotype</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pericarp</td>
<td>Cob glume</td>
<td></td>
</tr>
<tr>
<td>PI-rr</td>
<td>red</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PI-vv</td>
<td>variegated</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PI-ovov</td>
<td>orange variegated</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PI-fw</td>
<td>white (colorless)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PI-wr</td>
<td>white (colorless)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The PI-rr allele and its derivatives are well characterized at both the genetic and molecular levels (ANDERSON 1924, EMERSON 1917, GREENBLATT and BRINK 1962, LECHLT et al. 1989, ATHMA et al. 1992). The PI-rr gene was cloned by Ac transposon tagging. The PI-rr allele has a complex structure, as its ~7 kb coding sequence is flanked by two 5.2 kb direct repeats (LECHLT et al. 1989) (Figure 3). The upstream 5.2 kb direct repeat contains cis-elements that regulate PI-rr expression (SIDORENKO et al. 1999). Differential splicing
Figure 3. The structure of the Pl-rr allele. The black boxes are the exons of the Pl-rr allele. The hatched boxes indicate the 5.2 kb repeats flanking the Pl-rr coding region. The gray box that overlaps with the 3' hatched box is fragment 15. The other gray boxes within or outside of the hatched boxes are fragment 15 homologous sequences. S = SalI, S* = methylated SalI, H = HindIII
generates two transcripts: an 1802 nt transcript encodes the PI gene regulatory function, while the function of the 945 nt transcript is unknown. Comparison of the 1802 nt cDNA sequence and genomic sequence of PI-rr revealed that PI-rr contains three exons and two introns; the 4.6 kb intron 2 is the biggest intron yet identified in plants (GROTEWOLD et al. 1991).

The standard PI-vv allele isolated by Emerson contains an Ac element inserted in the intron 2 of PI-rr (EMERSON 1917, LECHLT et al. 1989). Ac insertion prevents PI-rr function, resulting in a colorless pericarp and colorless cob phenotype; whereas Ac excision from PI-vv restores PI-rr function, giving rise to the red pericarp and red cob phenotype. For PI-vv ear, the pericarp and cob tissues are a mosaic of red and colorless cells: the colorless cells retain Ac insertion in the PI locus, while in the red cells, the Ac element excised from the PI locus.

Other PI alleles could be derived from PI-vv via Ac transposition, e.g., the revertant allele PI-rr-4B2 is derived from PI-vv by Ac excision. In another case, the allele PI-ovov-1114 is derived from PI-vv by intragenic Ac transposition; i.e., Ac excised from its site of insertion in PI-vv, and reinserted 165 bp upstream in opposite orientation (PETERSON 1990).

A number of different PI alleles with Ac elements inserted in PI-rr have been isolated by ATIMA et al. (1992). For these PI alleles, if the Ac element inserted in one of the introns and has the same transcription orientation as that of the PI-rr allele, it generally specifies variegated pericarp (PI-vv phenotype); if the Ac element inserted in one of the introns and has the opposite transcription orientation to that of the PI-rr allele, it generally specifies orange variegated pericarp (PI-ovov phenotype). For the PI-vv alleles, presumably the transcription terminator in the Ac element can efficiently terminate PI transcription, resulting in production
of a chimeric \( PI/\text{Ac} \) message which lacks \( PI \) gene exon 3 sequences (Lechelt et al. 1989).

For the \( PI-ovov \) alleles, presumably the \( PI \) gene can elongate through the entire \( Ac \) sequence, which can be subsequently the spliced out as intron sequence. In this case, an amount of functional \( PI \) message sufficient to confer orange pericarp is produced. \( Ac \) insertion in any of the three exons severely inhibits \( PI \) expression, and thus generates very light variegated pericarp phenotype regardless of the \( Ac \) orientation.

Alleles with \( Ac \) insertion in \( PI-rr \) specify a variegated phenotype such as \( PI-vv \) and \( PI-ovov \). Although many of these alleles condition a similar phenotype, they may have different gene structures (different \( Ac \) insertion sites and \( Ac \) orientation). The same holds true for alleles derived by excision of \( Ac \) because \( Ac \) excision generally leaves a small footprint (minor sequence change at the \( Ac \) insertion/excision site). Alleles derived by excision of \( Ac \) may condition red pericarp (\( PI-rr \)), pale pericarp, or white pericarp (\( PI-ww \)), even though their structures may be very different. To avoid confusion, alleles are designated by an origin number that is placed after the two-letter suffix of \( PI \) allele, i.e. in \( PI-ovov-1114 \), 1114 is the origin number. Origin numbers are used to identify the potential structure difference between the \( PI \) alleles that have similar phenotype. In this way, alleles can be identified based on their independent origins, and alleles conferring similar phenotype can be distinguished from one another.

Transposition of the \( Ac \) element in \( PI \) allele (\( PI-vv \) or \( PI-ovov \)) can induce homologous recombination between the 5.2 kb direct repeats flanking the \( PI \) gene (Athma and Peterson 1991; Xiao and Peterson, in preparation) (Figure 3). The recombination event removes the whole \( PI \) coding region, resulting in loss of \( PI \) gene function. This transposon-induced
homologous recombination can explain the multiple-kernel colorless pericarp sectors observed in the \textit{PI-vv} and \textit{PI-ovov} ears. Products of this recombination reaction are stable \textit{PI-ww} alleles specifying colorless pericarp and colorless cob. The structure of the \textit{PI-wr} allele is much more complex than that of \textit{PI-rr}, it contains 6 tandem repeats, each repeat has essentially the same sequence as in \textit{PI-rr} except the few hundreds basepair 3' coding region (Chopra \textit{et al.} 1996; Chopra \textit{et al.} 1998). It is unknown whether or not all the repeats are active in contributing to the cob phlobaphene pigmentation. For most of the \textit{PI} alleles, the colored phenotype is dominant to the colorless phenotype.

\textit{Ac/Ds} family

The discovery of transposable elements

Transposable elements were described in maize in the late 1940s by McClintock. The transposable element activity was observed by studying the variegation in endosperm phenotypes. In one case, the variegated phenotype was associated with marker loss in the short arm of chromosome 9. Cytogenetic analysis showed that the genetic marker loss was correlated with chromosomal fragment loss resulting from chromosome breakage. The chromosome break always initiated at a particular position in chromosome 9, and this site was designated the \textit{Dissociation (Ds)} locus (McClintock 1947). The \textit{Ds} locus alone could not cause chromosome breakage, but it required the presence of another locus, \textit{Activator (Ac)}, in the genome. Subsequently, it was found that both \textit{Ds} and \textit{Ac} can change their position in the genome, and the transposition of either \textit{Ac} or \textit{Ds} could cause mutations. Transposition of \textit{Ds}
is dependent on the presence of Ac, while Ac transposition does not require any other special trans factors (McCLINTOCK 1948). In the following decades, McClintock and others identified a number of other element that can change their positions in the genomes of maize and other organisms. Because their unique mobility, they are termed transposable elements.

**The structures of Ac and Ds elements**

Several Ac elements were cloned from different loci in maize. It was found that they have identical or very similar structures. The Ac element isolated from wx9 is 4565 bp in length and has 11 bp imperfect terminal inverted repeat (TIR, base 1 and base 4565 are exceptions) (POHLMAN et al. 1984a, 1984b, MULLER-NEUMANN et al. 1984).

Compared to Ac, the structures of Ds elements are much more heterogeneous. Some Ds elements are derived from Ac by internal deletion, while others contain internal rearranged Ac-homologous sequence and sequences unrelated to Ac (DORING and STARLINGER 1986, FEDOROFF 1989, VARAGONA and WESSLER 1990). Yet another type of Ds (Dsl) shares only 40 bp sequence homology to the Ac termini. Dsl is 400 bp in length, its internal sequence is unrelated to Ac. Unlike the other Ds elements, Dsl can be mobilized by Ubiquitous element (SUTTON et al. 1984, GERLACH et al. 1987, PISABARRO et al. 1991, CALDWELL and Peterson 1992). The first transposable element described by McClintock that causes chromosome breakage actually is a double Ds, one Ds element inserted into another copy of Ds in the opposite orientation.
The cis-element required for Ac/Ds transposition

Ac has imperfect 11 bp TIR, while most Ds elements have perfect 11 bp TIR, thus the sequence change of the last base of Ac/Ds from A to G does not affect Ac/Ds transposition (Kunze 1996). However, mutation of the last base from A to C dramatically decreases Ac transposition frequency (Athma and Peterson, unpublished results), and deletion of the first 5 or the last 4 bases immobilizes the Ac element (Hehl and Baker 1989, Healy et al. 1993). So the TIR sequence is important for Ac/Ds transposition.

Although the TIR is essential for Ac/Ds transposition, it was shown that the sub-terminal sequences are also required for Ac/Ds transposition. By deletion of various portion of the Ds sub-terminal sequence, it was found that the 5' 238 bp sequence and the 3' 209 bp sequence are required for efficient Ds transposition. Further deletions dramatically decrease Ds transposition frequency (Coupland et al. 1989).

DNA replication and Ac/Ds transposition

As mentioned previously, P1-ww conditions variegated pericarp due to insertion of Ac in the P1-rr gene. However, the P1-ww ear is not homogeneously variegated, but often contains twinned sectors composed of light variegated pericarp (smaller and less frequent red stripe in pericarp) adjacent to red pericarp. Genetic analysis revealed that the light component of the twinned sector always contains two copies of Ac elements in the genome: one in the P1 locus, and the other (termed trAc) located elsewhere. In contrast, the red component contains either one Ac (in type I twin) or no Ac at all (in type II twin). For type I twin, genetic and molecular analyses have shown that the Ac in the red sector and the trAc in the light sector are in the
same position in the genome (Greenblatt and Brink 1962; Chen et al. 1992).

The formation of the twinned sector could be explained as following (for details, see Figure 4): Ac transposition follows DNA replication, but only one of the replicated Ac elements is competent for transposition. The excised Ac could reinsert in either a replicated or unreplicated region. Insertion of the excised Ac into a replicated region of the sister chromatid with Pl-νν will lead to the formation of a type II twin (Figure 4B). Insertion of the excised Ac into a unreplicated region will lead to the formation of a type I twin (Figure 4C). It has been proposed that all the Ac transposition events lead to the formation of incipient twinned sectors. Although untwinned red sectors and untwinned light variegated sectors can be observed, it has been proposed that apparently the untwinned sectors are in fact twinned with a sector which lies in a non-visible portion of the cob (Greenblatt et al. 1985). If that is true, then the frequency of untwinned red sectors should equal to the frequency of the untwinned light sectors because either sector has an equal chance to be in the invisible portion.

Supporting this model, the numbers of the observed untwinned red sector and untwinned light sector are roughly equal (Greenblatt 1974). However, I believe the above conclusion to be in error for the following reason: The frequency of Ac reinsertion into genetically unlinked sites after excising from Pl-νν is 39% (Fedoroff 1989). Insertion at an unlinked site should leads to random segregation of the trAc and Pl-νν, resulting in an excess of untwinned red sectors relative to untwinned light sectors. It is hard to imagine that trAc always segregates with the Pl-νν chromatid no matter whether it is linked to the Pl locus or not.

Three conclusions were made from the above genetic study: First, Ac transposes by a cut-and-paste mechanism. The element excises from the original position (donor site) and
Figure 4. The relationship of DNA replication, Ac transposition, and twinned sectors. A. After being replicated, only one Ac copy is competent to transpose. Ac transposes shortly after DNA replication into B. target site 1, Ac transposition in this case creates two sister chromatids that are different from their progenitor Pl-vv (specifying medium variegated pericarp): one chromatid contains Pl-vv plus trAc (specifying light variegated pericarp because of the Ac negative dosage effect); while the other contains Pl-rr (specifying red pericarp) (type II twin). C. into target site 2. The results of Ac transposition in this case are similar to B except that the chromatid with Pl-rr contains a trAc (type I twin). D. into target site 3. Ac transposition in this case creates two sister chromatids: one is the same as the progenitor, while the other carries Pl-rr, so only an untwinned red sector can be seen.

Three possible Ac reinsertion sites (1, 2, and 3) are indicated.
Competent Ac
reintegrates in a new position, the donor site does not retain the Ac/Ds copy after transposition. Second, Ac transposes shortly after the Ac-containing region is replicated, and only one replicated Ac is competent for transposition; the excised Ac could reinserted in either a replicated region or an unreplicated region. Third, Ac tends to transpose locally, i.e., the excised Ac element has more chance to reinsert into a linked site. Molecular studies have confirmed that DNA replication is required for Ac/Ds transposition by transient assay. In maize protoplast transfected with constructs containing Ds, the Ds element could transpose only when the construct contained a functional viral DNA replication origin (Wirtz et al. 1997).

**Footprint and target site duplication**

Like other transposons, Ac/Ds elements generate target site duplications upon insertion that subsequently flank the transposon. The length of the Ac/Ds target site duplication is 8 bp. It was proposed that Ac transposase makes 8 bp staggered nicks at the target site, and the two termini of the excised transposon link to the 8 bp overhangs. Subsequently, DNA repair fills the gaps and thereby generates 8 bp target site duplication (Peacock et al. 1984, Saedler and Nevers 1985).

Ac/Ds elements transpose through a cut and paste mechanism. Ac/Ds excision generally is not precise; the excision site nearly always contains some minor sequence changes termed a footprint. The sequence change in the footprint is not random. It was shown that each specific insertion site generates different patterns of footprint, and the 39 bp 5' and 18 bp 3' Ac/Ds flanking sequences affect the formation of the footprint (Scott et al. 1996, Rinehart et al. 1997).
1997). Recently, Weil and Kunze showed that Ac/Ds system can function in yeast. Many of the footprints in yeast contain short deletion or inverted repeats (personal communication), which are very similar to those produced by tam3 and the rag1, rag 2 system responsible for immunoglobulin gene rearrangement (Coen et al. 1986, Hiom et al. 1998, Agrawal et al. 1998, Hagmann 1997, van Gent et al. 1996).

**Ac encodes Ac transposase**

Ac produces a 3.5 kb transcript that contains a long GC-rich leader sequence and 2421 bp ORF. The 807 amino acid polypeptide encoded by the ORF is thought to be the transposase which is essential for transposition of both Ac/Ds elements (Kunze et al. 1987, Kunze and Starlinger 1989). In heterologous plant systems, the Ac cDNA is necessary and sufficient for transposition of Ds, thus confirms that the Ac cDNA encodes the transposase (Swinburne et al. 1992, Scofield et al. 1992, Rommens 1992).

**Negative dosage effect**

The light variegated sector on the P1-vv ear contains two copies of Ac. The frequency and size of red stripe in the sector is much less than that in standard P1-vv that contains only one Ac (Brink and Nilan 1952, Greenblatt and Brink 1962, Chen et al. 1987, 1992). The homozygous P1-vv ear also has less frequent and smaller red stripes compared to heterozygous P1-vv. It seems that increasing Ac copy number tends to delay the timing of Ac/Ds transposition. This phenomenon is termed the Ac negative dosage effect. The negative dosage effect is also observed in other maize loci containing Ac insertions (McClintock...
The mechanisms underlying the negative dosage effect is largely unknown. It was observed that increasing \( Ac \) copy number increases the \( Ac \) transcription level and protein expression level (Kunze et al. 1987; Fusswinkle et al. 1991), so the regulation of the negative dosage effect must be a post-translational process. It was also observed that increasing \( Ac \) copy number increases the frequency of \( Ac \) transposase aggregates (Helein et al. 1994). If soluble transposase is active for transposition while transposase aggregates are inactive, then increasing \( Ac \) copy may decrease the concentration of soluble (active) transposase, and thus leads to the negative dosage effect.

**\( Ac/Ds \) transposition and genome reorganization**

**\( Ds \) elements and chromosome breakage**

According to their response to \( Ac \), \( Ds \) elements can be divided into two classes: state I \( Ds \) elements cause a high frequency of chromosome breakage and a low frequency of reversion, whereas state II \( Ds \) elements give little or no chromosome breakage and a high frequency of reversion (McClintock 1949). Molecular analysis shows that a state I \( Ds \) is actually composed of two state II \( Ds \) elements, with one copy of state II \( Ds \) inserted into another identical \( Ds \) copy in the opposite orientation (Courage-Tebb et al. 1983, Doring and Starlinger 1984, Doring et al. 1989, Weck et al. 1984, Doring et al. 1984). This so-called double \( Ds \) element is not the only configuration that could induce chromosome breakage. It was found that pair of tightly linked state II \( Ds \) elements, or a terminally deleted \( Ac \) and
adjacent intact Ac, also cause a high frequency of chromosome breakage (Weil and Wessler, 1993, Ralston et al. 1989, Dooner and Belachew, 1991). English et al. dissected the double Ds element, and found that a 5' Ac end and a 3' Ac end in direct orientation are sufficient to cause chromosome breakage in transgenic tobacco (English et al. 1993).

Two very similar models were proposed by English et al. and Weil and Wessler to explain the phenomenon of chromosome breakage (English et al. 1993, Weil and Wessler 1993). In these models, the Ds ends located on different sister chromatids are recognized as substrates by Ac transposase. Attempted transposition of these Ds ends on sister chromatids leads to the formation of a chromatid bridge which breaks in the subsequent cell division (Figure 5). Both English et al. (1993) and Weil and Wessler (1993) identified chromatid bridge predicted by the breakage model.

**Chromosome breakage and chromosome rearrangement**

In the chromosome-breakage model, Ac transposase recognizes the Ds ends on sister chromatids as substrates. Hence the DNA molecules that contains these Ds ends can be considered as an aberrant transposon. Excision of this aberrant transposon creates a chromatid bridge. Insertion of the excised aberrant transposon into the bridge will generate chromosome rearrangements such as deletion and inverted duplication (Figure 5); otherwise, chromatid bridge will break at the upcoming mitosis. Thus, chromosome breakage and chromosome rearrangements are different outcomes of the aberrant transposition. Indeed, English et al. (1993) isolated the predicted deletions and inverted duplications from several independent aberrant transposition events. Because the deletions and inverted duplications are derived
Figure 5. Aberrant transposition and chromosome breakage. The red arrows indicate the 3' and 5' Ac termini in direct orientation. The hatched ovals indicate centromeres, and the other colored lines indicate chromosome segments. A. DNA replication produces two sister chromatids. Ac transposase (small ovals) recognizes the 5' Ac end (filled red arrow) and 3' Ac (open red arrow) end on different sister chromatids as substrates. B. Transposase cuts at the Ac ends to generate two chromatid fragments with Ac ends (aberrant transposon). The two chromatid ends proximal to the excised Ac 3' and 5' ends are ligated together, forming a chromatid bridge. C. The aberrant transposon could reinsert anywhere in the genome. If the transposon does not integrate into the chromatid bridge, the bridge will break in the upcoming mitosis; otherwise, insertion of the transposon into the bridge generates reciprocal deletion and inverted duplication chromatids which will segregate to the daughter cells at the next mitosis.

Note: The map is modified from English et al. (1993) and Weil and Wessler (1993).
from independent events, they are unable to identify the target site duplication predicted as a consequence of transposon insertion (English et al. 1995).

**The sh-m5933 allele and aberrant transposition**

The *sh-m5933* allele isolated by McClintock contains complex chromosome rearrangements; this is one of the few chromosome rearrangement mutants that have received extensive molecular analysis (Courage-Tebbe et al. 1983, Burr and Burr 1982, Doring et al. 1989). The structure of *sh-m5933* is very complex: it contains 2 copies of "one and a half" *Ds*, and one copy of double *Ds*. It also contains inverted duplications of the 5' end of the *shrunken1* gene and adjacent sequences (Figure 6). The proximal copy of "one and a half" *Ds*, the double *Ds*, and the sequence in between can be excised as a composite *Ds* element. Excision of this composite *Ds* element restores *sh* function. Most revertants still retain the ability to break chromosome; this property is probably due to the remaining distal "one and a half" *Ds* element, because the chromosome breakage frequency dramatically decreases after excision of the *Ds* element in the distal "one and half" *Ds* element (Courage-Tebbe et al. 1983, Doring et al. 1989).

The *sh-m5933* allele is derived from a wild type *shrunken* allele with a double *Ds* distal to the *sh* locus. It has been proposed that *sh-m5933* was generated by an aberrant transposition event similar to the above model (Fedroff 1989, Doring et al. 1984). If the *sh-m5933* was generated by aberrant transposition, the target site for transposon insertion was probably located in the wild type *shrunken* allele; insertion of aberrant transposon destroyed the *shrunken* gene function and gave rise to the shrunken kernel phenotype. But the structure of *sh-m5933* is far more complex than the above models could explain, so other process must
Figure 6. The structure of the *sh-m5933* allele (lower portion) and its progenitor allele (upper portion). The red arrows indicate *Ds* elements, with the filled arrowheads and the open arrowheads indicating the 5' *Ds* and 3' *Ds* ends, respectively. The region between the double *Ds* and the proximal "one and a half" *Ds* is a 30 kb insertion sequence. The thick black lines in the 30 kb insertion are inverted duplications. Another copy of this sequence is located adjacent to the distal 1.5 *Ds* element. The hatched \( \leftrightarrow \) shapes indicate *sh* sequences. The vertical yellow line in the progenitor allele is an 8 bp sequence in intron 7 of the *sh* locus, the same 8 bp sequences are repeated at three sites in the *sh-m5933* allele. The small black circles indicate the centromere of maize chromosome 9. The *sh* coding region in *sh-m5933* is disrupted by the double *Ds*, the "one and half" *Ds*, and the 30 kb insertion. The 5' portion of the *sh* allele and its adjacent sequences are repeated in the distal location, but their orientation relative to the original *sh* locus is not certain.

Notes: 1. The map is modified from Doring et al (1989) and Fedoroff (1989). 2. The map is not to scale.
be involved in the formation of the *sh-m5933* allele. Based on the research in this dissertation, a molecular model to explain the generation of *sh-m5933* will be presented in Chapter 4 of this thesis.

**Ac/Ds transposition and chromosome rearrangements other than deletion and inverted duplication**

In addition to the deletion and inverted duplication described above, McClintock observed that *Ds* can generate reciprocal translocations, inversions, and direct duplication (McClintock 1953a, 1953b, 1978). These chromosome rearrangements are dependent on the presence of *Ac*, implying transposition is involved in their generation. However, normal *Ds* (state II) transposition only changes the position of *Ds* in the genome, and would not in itself lead to large-scale chromosome rearrangements. Hence the aberrant transposition and/or other events must participate in producing chromosome rearrangements. Since most of these chromosome rearrangements have not been analyzed at the molecular level, it is not clear how transposition was involved in their generation. The result presented in this thesis may help to elucidate the responsible mechanism.

**Dissertation organization**

The thesis contains 4 chapters: general introduction (Chapter 1), two research papers (Chapter 2 and Chapter 3), and general summary (Chapter 4). The general introduction provides some background for my Ph. D. thesis project. In the first paper, a non-linear
transposition (NLT) model was proposed to explain the complex chromosome rearrangements identified in a maize PI gene allele, and molecular evidence to prove the non-linear transposition model is presented. The second paper describes the molecular analysis of a series of NLT-generated deletions. These deletions provide new information on the genes nearby the PI locus and establishes the general utility of the NLT mechanism for deletion mapping in plants. Chapter 4 contains general summary of the research data in chapters 2 and 3. Chapter 4 also presents a modified NLT model to explain the mysterious chromosome rearrangements in the sh-m5933 allele.

References


KUNZE, R. and P. STARLINGER, 1989 The putative transposase of transposable element Ac from *Zea mays* L. interact with subterminal sequence of Ac. EMBO J. 8: 3137-3185


MCCLINTOCK, B. 1953a Induction of instability at selected loci in maize. Genetics 38: 579-599.


Swinburne, J., L. Balcells, S. R. Scofield, J. D. Jones and G. Coupland, 1992 Elevated levels of Activator transposase mRNA are associated with high frequencies of Dissociation excision in Arabidopsis. Plant Cell 4: 583-595


CHAPTER 2 GENOME REARRANGEMENTS BY NON-LINEAR TRANSPOSONS IN MAIZE

A paper accepted by Genetics

Jianbo Zhang and Thomas Peterson

ABSTRACT

Transposable elements have long been considered as potential agents of large-scale genome reorganization by virtue of their ability to induce chromosomal rearrangements such as deletions, duplications, inversions, and reciprocal translocations. Previous researchers have shown that particular configurations of transposon termini can induce chromosome rearrangements at high frequencies. Here, we have analyzed chromosomal rearrangements derived from an unstable allele of the maize Pi (pericarp color) gene. The progenitor allele contains both a full-length Ac (Activator) transposable element and a Ac terminal fragment termed fAc (fractured Ac) inserted in the second intron of the Pi-rr gene. Two rearranged alleles were derived from a classical maize ear twinned sector, and were found to contain a large inverted duplication and a corresponding deficiency. The sequences at the junctions of the rearrangement breakpoints indicate that the duplication and deletion structures were produced by a single transposition event involving Ac and fAc termini located on sister chromatids. Because the transposition process we describe involves transposon ends located on different DNA molecules, it is termed non-linear transposition (NLT). Non-linear transposition can rapidly break and rejoin chromosomes, and thus could have played an important role in generating structural heterogeneity during genome evolution.
INTRODUCTION

Transposition is essentially a biochemical reaction. The enzyme that catalyzes the reaction is transposase, and the substrates of transposase are the 5' and 3' termini of the transposon. Theoretically, non-contiguous 5' and 3' transposon termini could serve as transposase substrates, and transposition could involve transposon termini located on different chromosomes. Such transposition events involving dispersed transposon ends could lead to major chromosomal rearrangements, whereas ordinary transposition of a contiguous element changes only the location of the transposon in the genome. However, genomes containing multiple copies of related transposons are generally quite stable; this suggests that transposition involving non-contiguous transposon termini is rare. It has been estimated, based on genomic Southern blot hybridizations, that the maize genome contains approximately 30 to 50 copies of Ac/Ds-like transposons (Fedoroff et al., 1983). However, it is unclear how many of these copies are transposition competent, as a certain proportion may be immobile fragments (Kunze 1996) or inactivated by epigenetic modifications associated with DNA hypermethylation (Leu et al. 1992). In the case of maize Ac/Ds elements, the ability of dispersed transposons to participate in transposition reactions may be further restricted by differences in the timing of replication of individual transposon ends (Wirtz et al. 1997), and by the methylation state of the transposon ends (Wang and Kunze, 1998).

Transposition events resulting in chromosome breakage or other rearrangements can be detected by the use of appropriate genetic markers. In the early phases of transposon discovery, McClintock observed that transposition of Ds (Dissociation) was occasionally accompanied by chromosomal rearrangements, such as deletions, duplications, inversions, and
reciprocal translocations. Because these rearrangements occurred only in the presence of Ac
(Activator), it was believed that they were produced by Ac/Ds transposition events
(McClintock 1953a, 1953b, 1978). Subsequently, transposon-related rearrangements were
also observed in Antirrhinum and Drosophila (Martin and Lister, 1989; Lister et al., 1993;
Lim and Simmons, 1994); in some cases, however, the rearrangements were attributed to
recombination between dispersed copies of transposons (Lim and Simmons, 1994).

Most of the chromosomal rearrangements isolated by McClintock have not been
studied at the molecular level, with the exception of the sh-m5933 allele. This allele contains
a >30 kbp inverted duplication (Burk and Burr, 1982) flanked by “double Ds” elements (one
Ds inserted into a second Ds in opposite orientation; Courage-Tebbe et al., 1983; Doring et
al., 1989). Federoff (1989) proposed that the inverted duplication in the sh-m5933 allele was
generated by a transposition reaction involving Ds termini located on sister chromatids. Later,
English et al. (1993) and Weil and Wessler (1993) proposed similar models to account for
the phenomenon of Ds-induced chromosome breakage. These models predict that
transposition reactions involving Ds termini on sister chromatids should generate reciprocal
deletions and duplications. Such rearrangements were indeed identified in transgenic tobacco
(English et al., 1995). However, in no single case have all rearrangement junctions been
sequenced to identify the characteristic nucleotide sequence changes predicted to arise from
the transposition process. These include the so-called “footprint” at the site of transposon
excision, and the target site duplication of host sequences at the site of transposon insertion.

Here we report that non-contiguous transposon termini can serve as substrates for
unusual transposition events and thereby generate major genome rearrangements. We
analyzed rearranged chromosomes derived from a classical maize twinned sector (GREENBLATT and BRINK 1962). One chromosome has a large (4.6 cM) deficiency, while the other chromosome carries the deleted segment as an inverted duplication. The rearrangement breakpoints contain the footprints and target site duplications typically generated by Ac transposition; these sequences prove that the rearranged chromosomes are the reciprocal products of a single non-linear transposition (NLT) event.

MATERIALS AND METHODS

Genomic DNA extractions, Southern blot hybridization, and Genomic Cloning: Total genomic DNA was prepared using a modified CTAB extraction protocol (POREBSKI et al., 1997). Agarose gel electrophoresis and Southern hybridizations were performed according to SAMBROOK et al. (1989), except hybridization buffers contained 250mM NaHPO₄, pH7.2, 7% SDS, and wash buffers contained 20mM NaHPO₄, pH7.2, 1% SDS. Genomic libraries were prepared using λ Fix II vector and in vitro packaging reactions (Stratagene). Genomic fragments were subcloned in pBluescript (Stratagene).

PCR amplifications: PCR amplifications were performed as described by SAIKI (1989) using the following oligonucleotide primers: Ac5, GGAATTCGTTTTTTACCTCGGGT; Ac6, GGAATTCTGCAACCCTTCCCCTC; A13, ATTGTGGATCCGCCCTCG. Reactions were heated at 94°C for 4 minutes; then cycled 35 times at 94°C for 20 seconds, 60°C for 30 seconds, and 72°C for 1 minute per 1 kb length of expected PCR product; then 72°C for 10 minutes. The rearrangement junction containing the 5' Ac end in the P-ww-defl allele was
isolated by Ligation-Mediated PCR (PRODHOM et al., 1998) as follows. Genomic DNA from plants of genotype P-ww-defl/P-wr was digested with SalI, ligated with SalI adaptor oligonucleotides (TCGCACTCCATTCAAGCTCTA and TCGATAGTAGCTTGAGAA), and used as template in PCR amplification using primer Ac5 and SalI adaptor primer. A single band was obtained that matched the expected size. (Genomic Southern analysis indicated that the rearrangement breakpoint in P-ww-defl is located on a 7.9 kbp SalI fragment, visible in Figure 4B, lane 1. Because the 3' flanking SalI site is located 3.0 kbp 3' of the 4.6 kbp Ac element, the 5' flanking SalI site should be located approximately 0.3 kbp 5' of Ac in the P-ww-defl allele.) The band amplified by LM-PCR was purified from an agarose gel and sequenced directly.

RESULTS AND DISCUSSION

Origin of a novel PI-ww allele with a large inverted duplication: The PI gene regulates the synthesis of a red phlobaphene pigment in maize floral organs, including the pericarp (outermost layer of the kernel derived from the ovary wall) and the cob (GROTEWOLD et al., 1994). The two-letter suffix of PI indicates its expression in pericarp and cob; i.e., PI-rr specifies red pericarp and red cob, PI-wr specifies white pericarp and red cob, and PI-ww specifies white pericarp and white cob (ANDERSON, 1924). The standard PI-w allele described by EMERSON (1917) conditions variegated pericarp and variegated cob. The PI-vv allele contains an Ac transposable element inserted in intron 2 of a PI-rr gene (LECHELT et al. 1989). The PI-vv allele gave rise to the PI-ovov-1114 allele (conditions orange-variegated pericarp and cob) by intragenic transposition of Ac (PETE 1990; ATHMA et al., 1992; numerals
placed after the two-letter suffix indicate the culture number of origin of each allele; alleles with the same phenotype but different culture numbers may have different gene structures.)

The *Pl-ovov-1114* allele in turn gave rise to *Pl-ww-9D9A*, which contains an *Ac* element, a 112 bp rearranged *Pl* gene fragment (rP), and a terminally-deleted *Ac* element termed *fAc* (fractured *Ac*) inserted in intron 2 of *Pl-rr* (Figure 1). The *fAc* element in *Pl-ww-9D9A* contains the 2039 bp 3' portion of *Ac*; a similar *fAc* element was described previously (RALSTON *et al.* 1989; DOONER and BELACHEW 1991). The intact *Ac* element can excise from *Pl-ww-9D9A* to give a revertant allele with the *Pl-rr* phenotype of red pericarp and red cob; this indicates that the rP and *fAc* insertions in *Pl-rr* intron 2 do not interfere with *Pl-rr* expression (J. ZHANG and T. PETERSON, unpublished). From a large multi-kernel white pericarp sector on a *Pl-ww-9D9A/Pl-ww* ear, we isolated a novel *Pl-ww* allele termed *Pl-ww-12:27-3*.

Genomic Southern analysis (not shown), genomic cloning and DNA sequencing indicates that *Pl-ww-12:27-3* contains a very large (>30 kb) insertion in *Pl-ww-9D9A* at the junction of the 3' end of *fAc* and the 5' portion of the *Pl* gene (hatched box in Figure 1). The insertion is an inverted duplication derived from the *Pl-ww-9D9A* sequence: from rP, it extends upstream through *fAc* and beyond the 5' end of the *Pl* gene. The sequences at the junctions of the *Ac/fAc* termini and the *Pl* gene in *Pl-ww-9D9A* and *Pl-ww-12:27-3* are identical except for a 2 bp change in *Pl-ww-12:27-3* at the junction of rP and the 5' portion of the *Pl* gene (Figure 1). The sequence changes, A to T and C to G, are similar to a typical footprint created by *Ac* transposition (RINHART *et al.* 1997), suggesting that the complex structure of *Pl-ww-12:27-3* probably resulted from an unusual transposition event.

**Non-linear transposition:** Based on its inverted duplication structure and putative
Ac transposition footprint, we propose that the \textit{Pl-ww-12:27-3} allele arose by non-linear transposition (NLT) (Figure 2; compare to models by \textsc{Weil} and \textsc{Wessler} 1993; \textsc{English et al.} 1993; \textsc{Gary et al.} 1996). The central feature of the model is that transposon termini located on different DNA molecules (sister chromatids are shown) can be utilized as transposase substrates. The resulting transposon (i.e., DNA internal to the Ac termini) is nonlinear and very large; in this case, comprising the terminal ~70 cM of the short arm of maize chromosome 1 (Figure 2B). Upon excision, the sequences originally flanking the Ac/fAc termini join to form a chromatid bridge, and some minor sequence changes occur at the junction to create a typical transposon excision footprint (Figure 2B). Insertion of the nonlinear transposon at a target site in the chromatid bridge will generate two unequal chromatids: one containing an inverted duplication (\textit{Pl-ww-id}), and the other with a corresponding deficiency (\textit{Pl-ww-def}) (Figure 2C). \textit{Pl-ww-id} contains a transposition footprint and a target site duplication (TSD), while \textit{Pl-ww-def} contains the other TSD. The inverted duplication structure of \textit{Pl-ww-12:27-3} and the position of the 2 bp sequence change (footprint) is exactly what would be predicted for a \textit{Pl-ww-id} allele produced by the non-linear transposition model (compare Figure 1 and Figure 2C).

We did not detect, however, the \textit{Pl-ww-def} allele predicted as the reciprocal product of the NLT reaction, possibly because the cell clone containing the \textit{Pl-ww-def} allele gave rise to the non-heritable internal portion of the cob (\textsc{Greenblatt} 1985). Therefore, we initiated a search for the reciprocal products of a single non-linear transposition event. Such reciprocal products could be detected and recovered in maize due to the cell lineage relationship between kernel pericarp and embryo (\textsc{Greenblatt} 1985). If non-linear transposition events
were to occur during ear development, the rearranged sister chromatids should segregate at mitosis into two daughter cells. Subsequent mitotic divisions of the daughter cells would generate a twinned sector, in which one twin carries the inverted duplication chromosome \((P-ww-id)\) and the other twin carries the corresponding deficiency \((P-ww-def)\). Because both deletion and insertion would destroy \(P1\) gene function, both twinned alleles should specify colorless pericarp instead of the variegated pericarp specified by the progenitor allele \(P1-vv-9D9A\). Following meiosis, each kernel in the twinned sector has a 50% chance to carry either \(P1-ww-id\) or \(P1-ww-def\).

To screen for the reciprocal products of a non-linear transposition event, we crossed \(P1-vv-9D9A/P1-wr\) with \(P1-wr, r-m3::Ds\) pollen. The \(r-m3::Ds\) is an Ac tester allele: Ac-induced excision of \(Ds\) from the \(r\) locus gives rise to purple anthocyanin pigment in aleurone cell clones (Kermicle 1980). Among approximately 1500 ears screened, one ear had a large colorless pericarp sector in which the kernels within the sector were phenotypically twinned for anthocyanin pigmentation: one twin had kernels with large purple aleurone sectors, while the other twin had kernels with small purple aleurone sectors (Figure 3). Two alleles, \(P1-ww-def1\) and \(P1-ww-id1\), were recovered from the twinned sector. Like \(P1-ww-12:27-3\), \(P1-ww-id1\) exhibits a dominant delay in the developmental timing of Ac-induced \(Ds\) excisions from \(r-m3::Ds\) as evidenced by small purple aleurone sectors. The relationship between the delayed timing of Ac-induced \(Ds\) excisions and the structures of \(P1-ww-12:27-3\) and \(P1-ww-id1\) is under investigation. In contrast, \(P1-ww-def1\) exhibits normal Ac-induced \(Ds\) excisions as evidenced by large purple aleurone sectors. \(P1-ww-def1\) is transmitted at normal frequencies through both pollen and ovum, but no homozygous \(P1-ww-def1\) plants could be obtained.
Molecular analysis of twinned alleles: As predicted by the NLT model, Southern blot analysis indicated that both a PI locus probe (fragment 15) and a PI-linked probe (p1.5B22, 3.5cM from PI locus) are deleted in P1-ww-defl and duplicated in P1-ww-idl (Figure 4). To test whether the P1-ww-idl allele has an inverted duplication, we screened a genomic P1-ww-idl library with both a PI gene probe (probe 10') and an Ac probe (1.6 kb Ac internal HindIII fragment). Eight clones hybridizing with both probes were obtained, and these were grouped into two types: type I (5 clones) contain fAc and the 5' portion of the PI gene, whereas type II (3 clones) contain Ac, fAc and the 3' portion of the PI gene. No clones contained both 5' and 3' portions of the PI gene. This result is predicted by the non-linear transposition model because the large insertion in P1-ww-idl separates the 5' and 3' portions of the PI gene.

Southern blot and sequence analysis of the P1-ww-idl clones indicate that P1-ww-idl, like P1-ww-12:27-3, contains an inverted duplication that begins at rP and extends beyond the 5' end of the PI gene. However, the P1-ww-idl duplication extends beyond the distal endpoint of the P1-ww-12:27-3 duplication.

We tested several additional predictions of the NLT model. The P1-ww-idl allele should contain an Ac-type footprint at the junction of rP and the 5' portion of the PI gene (Figure 2C); such a footprint was identified by PCR as follows: In P1-vv-9D9A, oligonucleotide primers A13 and Ac6 cannot produce a PCR product because the Ac6-homologous sequence in fAc has the same orientation as that of A13, and Ac6 in Ac is 5.2 kb from A13 in P1-vv-9D9A (Figure 1). However, following non-linear transposition, Ac6 in the P1-ww-id insertion lies opposite to A13, and the distance between them is 1.4 kb (Figure 1). PCR amplification produces the predicted product from P1-ww-idl, but not from P1-vv-
or Pl-ww-defl/Pl-wr (data not shown). Compared to the sequences of the Pl-ww-9D9A progenitor allele (Figure 5A and Figure 5B), the Pl-ww-idl allele contains a typical Ac-type footprint precisely at the junction of rP and the 5' portion of the Pl gene (Figure 5C); 2 basepairs are changed (A to T and C to G).

Upon transposition, Ac elements generate an 8 bp target site duplication (TSD) of the host sequence at the insertion site. The NLT model predicts that an 8 bp TSD will likewise be formed at the transposon insertion site; however, because the NLT transposon consists of two sister chromatids, the TSD should be found at the rearrangement breakpoints of the chromosomes carrying the twinned alleles. In the Pl-ww-idl allele, the TSD should be adjacent to the fAc; in the Pl-ww-defl allele, the complementary eight basepair sequence should be found adjacent to the Ac 5' end (Figure 2C). We subcloned and sequenced the DNA fragment containing the 3' end of fAc from Pl-ww-idl type II clones. The sequence of the suspected TSD adjacent to the 3' end of fAc is AGCGAGGC (Figure 5D). We cloned the DNA fragment containing the suspected TSD in Pl-ww-defl by modified ligation-mediated PCR (Prod'hom et al. 1998). The sequence of the PCR product contains the expected TSD (GCCTCGCT, the complementary sequence of AGCGAGGC) at the junction of the Ac 5' end and the rearrangement breakpoint in Pl-ww-defl (Figure 5E). The presence of the identical TSD at the Ac 5' end in Pl-ww-defl and the fAc 3' end in the Pl-ww-idl strongly supports the hypothesis that the rearrangements were generated by a single Ac transposition event.

According to the non-linear transposition model, the endpoints of the rearrangements in Pl-ww-defl and Pl-ww-idl should be adjacent to each other in the progenitor allele Pl-ww-9D9A. We designed primers nearby each endpoint, and used this primer pair to PCR amplify
the genomic sequence from the $P1$-$vv$-$9D9A$ template. The size of the PCR product (240 bp) matched the size inferred from the primer positions (data not shown). The PCR product contains a single copy of the GCCTCGCT target site, and the sequences flanking GCCTCGCT are the same as those from the endpoints of $P1$-$ww$-$idl$ and $P1$-$ww$-$defl$ (Figure 5F). These results show that the rearrangement breakpoints identified in the $P1$-$ww$-$defl$ and $P1$-$ww$-$idl$ alleles are derived from insertion of $Ac$ transposon ends into the GCCTCGCT target site in the progenitor chromosome.

The $P1$-$ww$-$12$:27-3 and $P1$-$ww$-$idl$ alleles both contain inverted duplications that begin at the rP in the $P1$-$vv$-$9D9A$ allele, and extend upstream beyond the 5' end of the $P1$ gene. According to the NLT model, the duplications should end at the transposon insertion site located in the chromatid bridge (Figure 2B). Indeed, RFLP mapping shows that the endpoints of the inverted duplications in $P1$-$ww$-$12$:27-3 (p1.5B22) and $P1$-$ww$-$idl$ (pJZPX) map 3.5 cM and 4.6 cM, respectively, proximal to the $P1$ locus, in the order: $P1$ 3' end, $P1$ 5' end, p1.5B22, pJZPX, centromere (M. McMullen and T. Muskett, personal communication).

Relative frequencies of normal and non-linear transposition: McClintock (1949) described two alternative states of $Ds$ elements: State I, which produces frequent chromosome breakage and rearrangement events, and State II, in which breakage events are rare, and reversions are frequent. McClintock’s original chromosome-breaking State I $Ds$ has since been associated with the compound double $Ds$, a structure in which one $Ds$ element is inserted in reverse orientation into a second identical $Ds$ copy (Döring et al., 1990; Martínez-Férez and Dooner, 1997). The inverse relationship between the frequencies of chromosome breakage vs. normal excision has also been observed in tobacco plants containing engineered
Ds constructs: double Ds elements promote chromosome breakage at high frequencies, but their rates of excision are much reduced compared to that of simple Ds elements. Thus, the presence of the double Ds configuration appears to inhibit simple excision of the individual Ds subunits (English et al., 1993). The reason for this is unclear, but it has been proposed that directly repeated 3' and 5' Ds ends are preferred to ends in normal orientation as substrates of Ac transposase when they are present together in double Ds elements (English et al., 1993). In the case of the P1-ww-9D9A allele, simple excision of the intact Ac element is easily recognized by the occurrence of red revertant sectors and germinal P1-rr revertants. The pattern of variegation (red stripes) given by the P1-ww-9D9A allele is very similar to that of the standard P1-ww allele which contains a single Ac insertion in the P1-rr gene. Thus, simple excision of Ac is not noticeably inhibited by a nearby fAc element in the P1-ww-9D9A allele, even though this allele contains directly repeated 3' and 5' termini in the same configuration as the natural (Döring et al., 1990) and engineered (English et al., 1993) double Ds elements.

In contrast, non-linear transpositions of the Ac/fAc elements in P1-ww-9D9A will fragment the P1 gene and generate colorless pericarp sectors. Among 1500 ears carrying the P1-ww-9D9A allele, we obtained 15 ears with large multi-kernel colorless pericarp sectors which gave rise to rearranged alleles characteristic of the non-linear transposition reaction (J. Zhang, P. Zhang, and T. Peterson, unpublished). This is an underestimate of the actual frequency of NLT events, for several reasons: First, we selected for study only large, easily recognized multi-kernel sectors. Colorless sectors smaller than one kernel in size are difficult to distinguish from the background of variegated pericarp which itself is a mosaic of red stripes on a colorless pericarp background. Moreover, some of the NLT events would be
predicted to be inviable and hence would not have been analyzed. Nevertheless, for the PI-vv-9D9A allele, the non-linear transposition events appear to be much rarer than the frequency of simple Ac excision. Further research will be required to determine the parameters that influence the propensity of individual Ac/Ds termini to participate in normal or aberrant transposition reactions.

**Significance of NLT-induced rearrangements:** There are a number of physical and genetic agents that can induce deletions in plants at random sites. In contrast, sister chromatid NLT has the unique property of producing deletions which extend from a single site in the genome (in this case the PI locus) to multiple flanking sites. The deletion endpoints will represent the insertion sites of the non-linear transposon; in the case of Ac-mediated transposition, these sites will likely be relatively close due to the tendency of Ac to transpose to nearby locations (Greenblatt 1984). The resulting nested deletions can be used to rapidly map molecular markers in a relatively small genetic interval (J. Zhang, P. Zhang and T. Peterson, unpublished results) a process that is difficult by standard meiotic mapping due to the limited resolution of most mapping populations. Generation of deletions by sister chromatid transposition could be extended to sites throughout the genome by transformation with transgene constructs containing Ac termini and a reporter gene whose loss is easily detected. Additionally, NLT events can generate duplications of varying sizes that may be useful for studies of the effects of gene dosage on expression levels (Guo et al. 1996).

The NLT model predicts that insertion of a non-linear transposon into a target site on another chromosome would generate other chromosomal rearrangements including translocations, acentric fragments, and dicentric chromosomes. Thus, non-linear transposition
could have contributed to the major genome rearrangements observed between related species (Bennetzen and Freeling 1993; Bennetzen et al. 1998). Moreover, because chromosome rearrangements can often lead to semisterility among progeny heterozygous with the progenitor genotype, NLT could lead to reproductive isolation and thus be an important first step in speciation (Lewis 1966). As we have shown here, NLT events generate characteristic chromosomal structures and sequences at the rearrangement junction. These molecular signs of NLT events may yet be recognizable in the genome sequences of closely related species.

ACKNOWLEDGMENTS

We thank Suzy Cocciolone and Dan Voytas for helpful comments on the manuscript, and T. Muskett and M. McMullen for mapping of pJZPX and p1.5B22. Journal Paper No. J— of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa, Project No. 3297, and supported by Hatch Act and State of Iowa funds.

REFERENCES


**FIGURE LEGENDS**

Figure 1. Schematic representation of structure of the progenitor *P1-vv-9D9A* allele (upper) and the *P1-ww-12:27-3* allele. Green arrows indicate the 5' portion of the *P1* gene and its associated upstream sequence. Red lines with arrow(s) indicate Ac or fAc, and the open and filled arrows indicate the 3' and 5' ends, respectively, of Ac/fAc. The short black line between Ac and fAc in *P1-vv-9D9A* indicates a 112 bp rearranged *P1* sequence (rP) that is duplicated in *P1-ww-12:27-3* (not to scale). The hatched box containing rP, fAc and the 5' portion of the *P1* gene and its upstream sequence indicates the large insertion in *P1-ww-12:27-3*. The arrows below the DNA structure indicate the positions and orientation of PCR primers. The X at the junction of rP and the green arrow indicates the position of the 2 bp sequence change in *P1-ww-12:27-3*. 
Figure 2. Non-linear transposition model. The two lines indicate sister chromatids joined at the centromere (oval). All the symbols have the same meaning as in Figure 1.

A. Ac transposase (small ovals) binds to the 5' terminus of Ac in one sister chromatid and the 3' terminus of fAc in the other sister chromatid.

B. Cuts are made at the Ac and fAc termini. The two nontransposon ends join together to generate a chromatid bridge, and minor sequence changes occur at the junction to form the transposon footprint.

C. The excised transposon termini insert at the target site (the junction between the black line and the green arrow) to generate one sister chromatid (Pl-ww-id) with an inverted duplication (green arrows), and a second sister chromatid with a corresponding deficiency (Pl-ww-def). The Pl-ww-id and Pl-ww-def sister chromatids will segregate to adjacent daughter cells at the subsequent mitotic division, forming a potential twinned sector. The small boxes labeled a to f indicate the rearrangement junction sequences shown in Figure 5, including footprint, target site, and target site duplication (TSD). Notes: in A and B, interactions between transposase molecules are not shown for clarity. In B, a fully-excised transposon is depicted, but the transposition reaction could proceed through sequential cut and ligation steps in which no free intermediate is formed. In B and C, the outcome of insertion of the excised transposon at a target site on the upper chromatid is depicted. Alternatively, insertion at a target site on the lower chromatid could occur; this would give the opposite orientation of fAc and rP in Pl-ww-id (not shown).
Figure 3. Ear with twinned sector produced from the cross: *P1-vv-9D9A/P1-wr* X *P1-wr, r-*
*m3::Ds*. Both sides of a single ear are shown. The left side has a sector of kernels (lower) with variegated pericarp and purple spotted aleurone containing the progenitor *P1-vv-9D9A* allele. The right side has a sector of kernels (lower) with colorless pericarp and large purple aleurone spots, containing the *P1-ww-defl* allele. Encompassing the entire upper portion of the ear is a large sector of kernels with colorless pericarp and tiny, barely visible purple aleurone spots; this sector contains the *P1-ww-idl* allele. Genotypes were confirmed by molecular analysis as described in the text. The unequal sizes of the twinned *P1-ww-idl* and *P1-ww-defl* sectors is most likely an indirect result of the very early formation of the original twinned daughter cells. In general, larger twinned sectors tend to be more irregular and unequal in size than smaller twins, probably because twinned sectors that are formed early in development are more likely to be affected by differences in the subsequent growth and development of the daughter cell clones. As stated by Greenblatt (1985), irregularities can also result from the fact that pericarp twinned sectors are formed in the three-dimensional structure of the ear, whereas they are visible only at the surface of the ear. Kernels with colorless aleurone are of *P1-wr* genotype and lack *Ac* due to meiotic segregation, hence the *R* gene required for aleurone pigmentation remains nonfunctional due to *Ds* insertion.

Figure 4. Structural analysis of the twinned sector alleles.

A. Restriction map of *P1-vv-9D9A*. The backbone is *P1-rr*, the solid boxes indicate exons, the numbered boxes indicate hybridization probes, and hatched boxes indicate sequences
homologous to probe 15. \( S=S_{II}, S^*=\)methylated \( S_{II} \) site. Insertion of fractured \( Ac \) (\( fAc \), smallest triangle), \( rP \) (small gray rectangle), and \( Ac \) (medium triangle) within intron 2 of \( P1-rr \) gives rise to \( P1-vv-9D9A \); the 5' and 3' ends of \( Ac/fAc \) are indicated.

B. Genomic DNA was digested with \( S_{II} \) and hybridized with probe 15. Lane 1, \( P1-ww-def1/P1-wr \); lane 2, \( P1-vv-9D9A/P1-wr \); lane 3, \( P1-ww-id1/P1-wr \). The 12.6 kb and 1 kb bands are derived from the \( P1-wr \) allele (CHOPRA et al. 1996; CHOPRA et al. 1998), and the 1 kb band is a single copy sequence used as internal control for DNA loading. The similar intensity of the 1 kb band in lanes 1, 2, and 3 indicate that these lanes contain equivalent amounts of DNA. The location of the 3.0 and 1.2 kb bands is shown in panel A. In the \( P1-ww-def1/P1-wr \) genotype (lane 1), the 3.0 kb band is missing, and the 1.2 kb band is less intense than in the progenitor \( P1-vv-9D9A/P1-wr \) genotype (lane 2). In contrast, the same 1.2 kb and 3.0 kb bands are more intense in the \( P1-ww-id1/P1-wr \) genotype (lane 3). Lanes 1, 2 and 3 contain bands at 7.9 kb, 10 kb, and ~12 kb, respectively. These bands arise from cutting at the \( S_{II} \) site at the 3' boundary of probe fragment 15, and at \( S_{II} \) sites upstream of the \( Ac \) and/or \( fAc \) insertion in the \( P1-ww-def1, P1-vv-9D9A, \) and \( P1-ww-id1 \) alleles. The \( S_{II} \) site in fragment 10' is shown for the \( P1-vv-9D9A \) allele in panel A. This \( S_{II} \) site is removed by the rearrangements in the \( P1-ww-def1 \) and \( P1-ww-id1 \) alleles and replaced by other \( S_{II} \) sites which alter the sizes of the corresponding fragments.

C. Genomic DNA was digested with the indicated enzymes, and hybridized with p1.5B22, a probe located 3.5 cM from the \( P1 \) locus. The loading is the same as in panel B (same preparation and amount of DNA). The bands in \( P1-ww-def1/P1-wr \) genotype (lanes 1) are the weakest, and the bands in \( P1-ww-id1/P1-wr \) genotype (lanes 3) are the strongest. The signal
intensities reflect the copy number of the probe sequence: the $P1$-wr haplotype contains one copy of the probe sequence, whereas $P1$-ww-defl, $P1$-vv-9D9A, and $P1$-ww-idl contain 0, 1, and 2 copies, respectively, of the probe sequence.

Figure 5. Footprint and target site duplications of the $P1$-ww-defl and $P1$-ww-idl alleles created by nonlinear transposition. The sequences from panels A to F correspond to the sequences from boxes a to f in Figure 2; the color and orientation of the sequence letters match the color and orientation of lines in Figure 2.

The footprint and target site duplication sequences are underlined. See text for further details.

A. Junction of rP and Ac 5' terminus in $P1$-vv-9D9A

B. Junction of 5' portion of $P1$ gene and the 3' terminus of $fAc$ in $P1$-vv-9D9A

C. Junction of 5' portion of $P1$ gene and rP in $P1$-ww-idl containing 2 bp transposon footprint.

D. Junction of inverted duplication and $fAc$ 3' end in $P1$-ww-idl.

E. Junction of deletion endpoint and Ac 5' end in $P1$-ww-defl.

F. Original target site sequence in $P1$-vv-9D9A.
Figure 1
Figure 2
Figure 4
Figure 5

A  
GGTTGTAGTCAGGGA  
CCCAACATCAGTCCCT

B  
CAACTACAACTAGGGA  
GTGATGTGATCCCT

C  
CAACTACAAAGCTCTACAACC  
GTGATGTTCAAGATGTTGGG

D  
TGAGCGAGCGAGCTAGGGA  
ACTCGCTCGCTCCGATCCCT

E  
GCCGCCGCTCGCTCAGGGA  
CGGCCGGAGGCCAGTCCCT

F  
GCCGCCGCTCGCTCGCTCA  
CGGCCGCGGGACCGAGGATG
CHAPTER 3. NESTED DELETIONS GENERATED BY NON-LINEAR TRANSPOSITION IN MAIZE

A paper to be submitted to Genetics

Jianbo Zhang, Peifen Zhang, and Thomas Peterson

Department of Zoology/Genetics, and Agronomy, Iowa State University

ABSTRACT

Ten interstitial deletion mutants derived from a \textit{P1} allele (P1-vv-9D9A) were analyzed. All ten deletions start at the same point in the second intron of \textit{P1} gene, but end at various sites up to >4.6 cM proximal to the \textit{P1} gene. The structures of the deletions suggest that they were generated by non-linear transposition (NLT) events involving \textit{Ac} termini located on sister chromatids. Genetic and molecular analysis of the deletion series provides several new insights into structure and function of genes in the immediately proximal vicinity of the \textit{P1} locus. First, the \textit{P2} gene (a duplicate of \textit{P1}) that is tightly linked to \textit{P1} contributes to the silk browning phenotype and levels of silk maysin, a C-glycosyl flavanone that deters feeding by corn earworm. Second, there appears to be an essential gene responsible for a zygotic lethal phenotype located upstream of the \textit{P2} locus. And, probe npi286 and the \textit{dekl} locus that were previously thought to be proximal to \textit{P1} are probably distal to the \textit{P1} locus. We conclude that NLT events can efficiently generate interstitial deletions, and that the resulting nested deletions are potentially useful for dissection of local intergenic regions, and for rapid correlation of genetic and physical maps.
INTRODUCTION

Deletions have been appreciated as very efficient tools for genetic mapping. In several classic studies, BENZER (1961, 1962) mapped the relative order and genetic positions of ~2400 phage T4 rII mutants by crossing these mutants with seven overlapping deletions that span the rII region. Based on their ability to complement the deletion series, all the mutants were easily and unambiguously localized to one of the seven major deletion regions. Using smaller deletions in each major deletion region, BENZER more precisely mapped the relative order of the mutants. BENZER completed the mapping using 25,000 crosses. Whereas more than 2 million crosses would have been required to accomplish the task using traditional two or three-factor method.

A similar approach was successfully used for the physical mapping of part of the drosophila X chromosome (SNYDER et al. 1985) and localization of the lettuce dm3 mutation (MEYERS et al. 1998). In addition to genetic mapping, deletion mutants are also very useful for mutation screening due to their pseudo-dominance phenotype. If using a deletion heterozygote as starting material to perform mutagenesis, any non-lethal recessive mutation located within the deletion region can be uncovered in the M0 generation; whereas otherwise recessive mutations can be detected only in the M1 generation.

The most widely-used treatment to induce deletions is gamma irradiation (CECCHINI et al. 1998, ANDERSON et al. 1996). However, high energy irradiation can also induce other undesirable chromosome rearrangements and point mutations, which can complicate the recovery and analysis of deletion mutants. In maize, the r-XI allele can induce terminal
deletions, but the viability of large terminal deletions is poor. If gene of interest is far from the telomere, it will not be possible to recover a viable deletion large enough to include the gene.

Recently, the cre/lox system was used to generate deletions in mouse, and deletions up to 3-4 cM were obtained (Ramírez-Solis et al. 1995, Wagner et al. 1997, Li et al. 1996, Zeh et al. 1998). The cre/lox system has also been applied to plant species such as tobacco and Arabidopsis to generate deletions, inversions and reciprocal translocations (Dale and Ow 1990, Medberry et al. 1995, Russell et al. 1992, Bayley et al. 1992, Odell et al. 1994, Osborne et al. 1995). Deletions were generated using cre/lox system as follows: Plants were transformed with a construct containing two lox sites: one lox site is located in the Ds element in the construct, the other is outside the Ds element. In the presence of Ac, the Ds element can transpose to a new site. If the lox site in the transposed Ds element and the lox site in the original construct are in the same chromosome arm and in direct orientation, the DNA sequence between the two lox sites can be deleted by cre/lox recombination.

Previously we have shown that chromosome rearrangements could be generated from the P1-vv-9D9A allele via non-linear transposition (Zhang and Peterson, 1999) (Figure 1). P1-vv-9D9A carries an intact Ac element and a terminally deleted Ac element. Ac transposase can excise the 3'fAc terminus and the 5'Ac terminus on two sister chromatids and thereby generate a non-linear transposon composed of the two half-arms of the chromosome 1. The two chromatid ends flanking the Ac/fAc ends joined together to form a chromatid bridge. The reintegration of the excised transposon ends into the chromatid bridge generates a reciprocal deletion and inverted duplication. In this paper, we present the molecular characterization of 10 interstitial deletions derived from P1-vv-9D9A by the NLT mechanism. As the NLT model
predicted, all the deletions start at the fAc insertion site in \textit{P1-vv-9D9A}, and end at various sites in the proximal region flanking the \textit{P1} locus. The largest deletion is > 4.6 cM. These results demonstrate the utility of the NLT mechanism for production of deletions in plants.

\textbf{METHODS AND MATERIALS}

\textbf{Mutation screening:} To recover chromosome rearrangements, we screened 1,500 ears derived from the progenitor allele \textit{P1-vv-9D9A}, for multiple-kernel sectors of colorless pericarp or whole colorless pericarp ears because the chromosome rearrangements predicted by the NLT model would be expected to destroy \textit{P1} gene function. 20 ears with multiple-kernel colorless pericarp sector, 20 colorless pericarp ears and one ear containing twinned colorless pericarp sector were obtained from the cross \textit{Pl-w-9D9A/Pl-wr} with \textit{Pl-wr, r-m3::Ds}. The seeds with purple spotted aleurone from these sector or ears are planted in the summer nursery field or greenhouse for further analysis.

\textbf{Genomic DNA extractions, Southern blot hybridization:} Total genomic DNA was prepared using a modified CTAB extraction protocol (\textit{Porebski et al.}, 1997). Agarose gel electrophoresis and Southern hybridizations were performed according to \textit{Sambrook et al.} (1989), except hybridization buffers contained 250mM NaHPO$_4$, pH7.2, 7% SDS, and wash buffers contained 20mM NaHPO$_4$, pH7.2, 1% SDS.

\textbf{PCR amplifications:} PCR amplifications were performed using the following oligonucleotide primers: Ac3, GGAATTCGTTTTTTACCTCGGGTTC; Ac6, GGAATTCTGCAACCCTTCCCCTCC; JZ1, GGAATTCCGCGTGCTGGTCTCTCACTC. Reactions were heated at 94°C for 4 minutes; then cycled 35 times at 94°C for 20 seconds,
60°C for 30 seconds, and 72°C for 1 minute per 1 kb length of expected PCR product; then 72°C for 10 minutes.

RESULTS

Identification of the NLT-generated deletion mutants: 20 ears with multiple-kernel colorless pericarp sector, 20 colorless pericarp ears and one ear containing twinned colorless pericarp sector were obtained from the cross Pl-vv-9D9A/Pl-wr with Pl-wr, r-m3::Ds. The colorless pericarp phenotype indicates that these Pl-9D9A derivatives could have different structures from that of progenitor Pl-vv-9D9A allele that conditions variegated pericarp. However, other structural changes in the Pl-vv-9D9A allele could also eliminate Pl gene expression. First, Ac is known to be able to induce intrachromosomal recombination between the two 5.2 kb Pl-flanking direct repeats, leading to the loss of the whole Pl coding region and eliminates Pl expression (ATHMA and PETERSON 1991). Second, transposition of Ac to an essential site in the Pl gene could knockout Pl function.

The NLT model predicts that sequences upstream of the Ac insertion site in Pl-vv-9D9A should be missing in the NLT-generated deletions. This region includes the 3.0 kb and 1.2 kb SalI fragments that are homologous to probe 15, and the 8.6 kb SalI fragment containing a part of probe 10'. Southern analysis was performed to distinguish the NLT-generated deletions from other types of mutations. Genomic DNA samples were digested with SalI and hybridized with probe 15. The 3 kb SalI band is missing in 10 colorless pericarp derivatives of Pl-vv-9D9A, and the intensity of the 1.2 kb band is also decreased for those 10 mutants compared to that in Pl-vv-9D9A. The loss of these two fragment is expected for
deletions generated by the NLT mechanism. A further prediction is that the 8.6 kb *SalI* fragment detected by fragment 10' should be missing in these alleles (Figure 2A). Southern analysis with probe 10' confirmed that the 8.6 kb *SalI* fragment is missing in these 10 lines (data not shown).

The NLT model also predicts that the deletion generated by the NLT mechanism should retain the 3' portion of the *P1* gene. PCR analysis was performed by using the primer-pair JZ1 and Ac3 (Figure 2A). All the 10 deletions give the predicted 300 bp PCR product, which suggests that the deletion retains the junction sequence between *Ac* 3' end and 3' of *P1* portion in *P1-vv-9D9A* (Data not shown here). Moreover, Southern analysis with *HindIII* digestion and probe 15 shows that the deletion alleles still contains a 7 kb band as their progenitor allele *P1-vv-9D9A* (Data not shown here).

**Endpoint mapping of the NLT-generated deletions:** The 2.8 kb *SalI* fragment in the *P2* locus (a duplication of *P1*) also can hybridize with the probe 10' (P. Zhang and T. Peterson, unpublished results). This fragment is retained in deletion *P1-ww-774*, but is absent in the other 9 deletions (data not shown). Thus, the deletion in *P1-ww-774* does not reach the close-linked *P2* locus. Southern blot and PCR analysis indicated that the endpoint of *P1-ww-def2* lies in 3' portion of the *P2* gene. To determine the relative sizes of the other deletions, two *P1*-linked probes (p1.5B22, pJZPX) were used for genomic Southern analysis. The probe p1.5B22 is the DNA fragment from the endpoint of *P1-ww-12:27-3*, and the probe pJZPX is the DNA fragment from the endpoint of *P1-ww-idl*. These probes mapped at 3.5 and 4.6 cM proximal to *P1*, respectively. Genomic DNA was digested with *HindIII*, and hybridized with these two probes. For probe 1.5B22, four mutants (*P1-ww-756, P1-ww-755, P1-ww-760, and*...
Pl-ww-defl) have weaker signals although the DNA loading in different lanes is similar. The plant DNAs were heterozygous with the Pl-wr allele, and we conclude that the lower signal intensity means that probe 1.5B22 is deleted in these 4 alleles (Figure 3). For probe pJZPX, Southern analysis shows that Pl-ww-755, Pl-ww-756, Pl-ww-defl have weaker signal, so pJZPX is deleted in these 3 alleles. In our previous study, we demonstrated that pJZPX is deleted in Pl-ww-defl because this probe is cloned from the endpoint of Pl-ww-idl (the reciprocal inverted duplication of Pl-ww-defl) (Zhang and Peterson 1999). The deletions in mutants Pl-ww-755 and Pl-ww-756 extend further beyond pJZPX.

Summarizing these data, the endpoint of Pl-ww-774 is placed between Pl and P2, and the endpoint of Pl-ww-def2 is placed within P2. The endpoints of Pl-ww-759, Pl-ww-775, Pl-ww-761, Pl-ww-765 are placed between P2 and p1.5B22, and the endpoint of Pl-ww-760 is placed between p1.5B22 and pJZPX. Pl-ww-defl has its endpoint adjacent to pJZPX, and Pl-ww-755 and Pl-ww-756 have their endpoints beyond pJZPX (Figure 4).

The P2 locus is tightly linked to the Pl locus: Although P2 is 95% homologous to Pl (P. Zhang and T. Peterson, unpublished results), their Southern patterns are different for specific probes, we can identify P2 specific bands only in non-Pl-wr allele in Southern analysis. Numerous crosses between Pl-wr and non-Pl-wr alleles, subsequent selfing, and Southern analysis shows that the P2 gene always segregates with the non-Pl-wr alleles (S. Chopra, P. Athma, and T. Peterson, unpublished results). Thus, P2 is tightly linked to Pl. Analysis of the NLT-generated deletions verified that. Among 10 NLT-generated deletions, only one retains the intact P2 sequence, and one retains partial P2 sequence, while the others lost the P2 sequence completely. The closest marker proximal to Pl is p1.5B22 that is 3.5 cM
from \textit{P1}. From our mapping results, \textit{P2} is placed between \textit{P1} and p1.5B22, so the distance between \textit{P1} and \textit{P2} is less than 3.5 cM.

**The relative orientation of the \textit{P2} locus to the centromere:** Previously, we have shown the transcriptional orientation of the \textit{P1} gene relative to the centromere is following: 3' \textit{P1} 5', centromere (\textit{Zhang} and \textit{Peterson} 1999). The \textit{P1-ww-def2} allele has a deletion of the 5' portion of the \textit{P1} gene, the 3' portion of the \textit{P2} gene, and the intervening sequences. Thus, the \textit{P2} gene has the same orientation as that of the \textit{P1} gene, and it is located between \textit{P1} and centromere in the following arrangement: 3' \textit{P1} 5', 3' \textit{P2} 5', centromere (Figure 5).

**Zygotic lethal phenotype of NLT-generated deletions:** Although the \textit{P1-ww-def1} allele is transmitted normally through both pollen and egg, we were unable to recover plants homozygous for this allele among 60 progenies from the self pollination of a \textit{P1-ww-def1/P1-wr} plant. There is a negligible probability (3.19x10^{-4}) that we did not recover a homozygous plant from a planting of this size. The self-pollinated \textit{P1-ww-def1/P1-wr} ears show some empty space and irregular row which is typical of 25% semisterility. This suggests that the \textit{P1-ww-def1} deletion has removed a gene essential for viability of the zygote. Among the mutants analyzed here, all four deletions (\textit{P1-ww-def1, 760, 755, 756}) which remove p1.5B22 are homozygous inviable, and one of the mutants whose deletion endpoint lies between \textit{P2} and p1.5B22 (\textit{P1-ww-765}) is also homozygous inviable. These results are consistent with the placement of a sporophyte-essential gene in the interval between \textit{P2} and p1.5B22. This also suggests that among the four deletions in the \textit{P2}-p1.5B22 interval, the deletion in \textit{P1-ww-765} should be larger than that in \textit{P1-ww-759, 761, and 775}. Interestingly, \textit{Emerson} (1939)
identified a zygotic lethal mutation (zll) which mapped to 1.2 cM proximal to Pl. The zll mutant stock has apparently been lost.

DISCUSSION

NLT generates more deletions than inverted duplication: The NLT model predicts that equal numbers of inverted duplications and deletions should be found. From the ~1500 ears screened (P1-vv-9D9A/P1-wr cross to P1-wr, r-m3::Ds), 10 NLT-generated deletions and 6 NLT-generated inverted duplications were recovered. The deletion frequency seems higher than the duplication frequency, which is surprising considering the potential that large deletions may not be heritable due to loss of essential gene(s). However, these results are consistent with a modification of the NLT model which takes into account the fact that transposition occurs during, or shortly after, DNA replication (Figure 6). In the NLT model showed in Figure 1, the chromatids are depicted as fully replicated. However, if transposition occurs before replication is complete, then the target site could be located in an unreplicated region. If the excised transposon ends insert into a unreplicated site in the chromatid bridge, the product containing the duplication will be located on an acentric chromatid fragment which will be lost in the subsequent cell division. In contrast, the chromatid containing the deletion will always contain a centromere and hence can be recovered if genetically viable (Figure 6). Thus, the relative frequencies at which deletions and duplications are recovered will depend to some extent on the timing of replication of the donor Ac sequence. The genetic analyses of Greenblatt and Brink (1962) support two important conclusions regarding transposition of Ac from the P1-vv allele. First, Ac tends to transpose more
frequently to sites near the *Pl* locus. Second, *Ac* inserts more frequently into unreplicated sites, than into sites that are already replicated. These two conclusions are consistent with the idea that the *Pl-vv* allele usually replicates earlier than the flanking chromosomal regions. Similarly, the non-linear transposition of *Ac/fAc* termini in the *Pl-vv-9D9A* allele may also result in preferential insertion into nearby unreplicated sites, thereby leading to a higher frequency of deletions than duplications, as discussed above.

Most of the NLT-generated deletions or inverted duplications are derived from the untwinned sectors: According to the NLT model, inverted duplications result from transposon reinsertion into a replicated site in the chromatid bridge; this should generate a corresponding deletion for each duplication allele. We identified 6 inverted duplications, but only one case did we detect the corresponding deletion. This reciprocal deletion and inverted duplication were derived from a twinned sector. Both components of the twinned sector had colorless pericarp, but this twins were recognizable based on difference in the *Ac*-induced *Ds* excision from the *R* locus which give purple aleurone sectors.

The deletion twin produced large aleurone sectors typical of a single dose of *Ac*, whereas the duplication twin produced much smaller aleurone spots which are a hallmark of the *Ac* negative dosage effect. Thus, these twinned sectors were recognizable only on the basis of difference in the *Ac*-activity. For those genetic twinned sectors that condition the same pericarp and aleurone phenotypes, it would be difficult to distinguish the twins. This probably is the case for most twinned sectors, because both deletion and inverted duplication give rise to the colorless pericarp, and contains a single active *Ac* in their genomes. Also,
recovery of duplication would be favored because very large deletions would remove essential
gene(s) and are gametophyte lethal.

**PI genetic neighborhood:** The probe npi286 is tightly linked to the *PI* locus, but its
proximal/distal relationship to the *PI* locus is uncertain (Neuffer et al. 1997). Southern
analysis shows that npi286 is not removed by even the largest deletion (>4.6 cM); thus, it is
most likely that npi286 is located in the distal side of *PI* (Figure 7). Likewise, the *dekl* locus
was mapped 1.0 cM proximal to *PI* (Neuffer et al. 1997), but none of our deletions affected
*dekl* function as assessed by crosses to *dekl*/Dekl heterozygote. Thus, it is probable that the
*dekl* locus is located distal to *PI*. This prediction is not unreasonable because original
mapping data for *dekl* and *PI* was only from the two-point test cross and were not conclusive
regarding the relative positions of *PI* and *dekl* (E. H. Coe, personal communication).

**Silk Browning and silk maysin:** Silk browning and maysin are quantitative traits.
Previous reports shows the *PI* region contribute ~57% of the total genetic variability for silk
maysin level (Byrne et al. 1996, Lee et al. 1998, McMullen et al. 1998). The *PI-rr* allele
clearly contributes to silk browning and silk maysin level, because a mutation of *PI* (e.g., the
*PI-ww-1112* allele in which the whole *PI* coding region is deleted), specifies much weaker
silk browning and much lower silk maysin levels compared to the *PI-rr* allele. The residual
silk browning and silk maysin level in the *PI-ww-1112* allele must result from other factor(s).
The *PI-ww-774* has a deletion which ends somewhere between *PI* and *P2*, and it shows silk
browning and maysin levels similar to that of *PI-ww-1112*. In contrast, the *PI-ww-def2* allele
has a deletion which extends into the 3' portion of the *P2* gene. Interestingly, the *PI-ww-def2*
allele has no silk browning and negligible silk maysin levels. The *P2* gene is expressed in silk,
and it encodes a protein that is 95% homologous to the P1-encoded protein. These results indicate that the P1 and P2 genes are essential co-regulators of maysin biosynthesis. (Byrne et al. 1996).

ACKNOWLEDGMENTS

Mapping data courtesy of T. Muskett and M. McMullen, U. Missouri. The dek1/Dek1 stock is a kind gift from Dr. Karen Cone. We thank W. Sheridan and J. Clark for helpful discussion.

REFERENCES


EMERSON, R. A., 1939 A zygotic lethal in chromosome 1 of maize and its linkage with neighboring genes. Genetics 24: 368-384


Zeh, K., M. Andahazy, S. O'Gorman and H. Baribault, 1998 Selection of primary cell cultures with cre recombinase induced somatic mutations from transgenic mice. Nucleic Acids Res. 26: 4301-4303

FIGURE LEGENDS

Figure 1. Non-linear transposition model. The two lines indicate sister chromatids joined at the centromere (oval). The cyan, purple, and green lines indicate the 5' portion of the P1 gene and its associated upstream sequence. Red lines with arrow(s) indicate Ac or fAc, and the open and filled arrows indicate the 3' and 5' ends, respectively, of Ac/fAc. The short black line between Ac and fAc in P1-vv-9D9A indicates a 112 bp rearranged P1 sequence (rP) (not to scale). A. DNA replication creates sister chromatids. Ac transposase (small ovals) binds to the 5' terminus of Ac in one sister chromatid and the 3' terminus of fAc in the other sister chromatid. B. Cuts are made at the Ac and fAc termini to release the non-linear transposon. The two non-transposon ends join together to form a chromatid bridge. C, D, and E. Reintegration of the non-linear transposon into target sites 1 or 2 or 3 will generate nested...
deletions. All the deletions start at the fAc or Ac insertion site in P1-vv-9D9A, and end at various sites on the centromere-proximal side of the P1 gene. The size of each deletion is determined by the target site where the transposon end inserts. Insertion of the non-linear transposon at target site 1 generates the largest deletion; Insertion of the non-linear transposon at the target site 2 generates medium deletion; while Insertion of the non-linear transposon at the target site 3 generates smallest deletion.

Figure 2. Southern Analysis of the P1-vv-9D9A derivatives with probe 15. A: Restriction map of P1-vv-9D9A. The backbone is P1-rr, the solid boxes indicate exons, the numbered open boxes indicate hybridization probes, and hatched boxes indicate sequences homologous to probe 15. The short black arrows indicate the orientation and the approximate position of primers used in PCR analysis. S = SalI, S* = methylated SalI site, H = HindIII. Insertion of fractured Ac (fAc, smallest triangle), rP (small gray rectangle), and Ac (medium triangle) within intron 2 of P1-rr gives rise to P1-vv-9D9A. The 5' and 3' ends of AcIfAc are indicated in the map. The arrows indicate the approximate position and the orientation of the primers used in PCR analysis. B. Genomic DNA was digested with SalI and hybridized with probe 15. The 12.6 kb and 1 kb bands are derived from the P1-wr allele (CHOPRA et al. 1996; CHOPRA et al. 1998), and the 1 kb band is a single copy sequence used as internal control for DNA loading. See Figure 2A for the location of the 3.0 and 1.2 kb SalI bands in the P1 locus.


Note: The two \textit{P1}-linked probes are single copy sequence. The DNA loading in different lanes is approximal equal. A weaker signal in a lane indicates deletion of the sequence corresponding to this probe. The residual signal is derived from the homologous chromosome in these heterozygous plants. A more intense signal in a lane indication duplication of the corresponding region.

Figure 4. Endpoint mapping of the NLT-generated deletions. The black boxes indicates the relative position of the probes. The endpoints of the deletions are located in the indicated region, and the dashed line inside the parenthesis means the endpoint is not certain.

Figure 5. The relative \textit{P2} orientation to \textit{P1} and the centromere. The solid boxes are the \textit{P1} and \textit{P2} gene. The small circle is the centromere. The structure of the \textit{P1-ww-9D9A} is shown in the upper portion, while the structure of the \textit{P1-ww-def2} allele is shown in the lower portion.
Figure 6. The modified NLT model. The small circles are centromere. The other symbols have
the same meaning as those in Figure 1. A: Only a part of the chromosome, including the \( P I-vv-9D9A \) allele, is replicated. B: Excision of the non-linear transposon and formation of the
chromatid bridge. The target site for the non-linear transposon integration is located in the
unreplicated region. C: The green end of the target site joins to the \( Ac \) 5' end, while the black
end of the target site joins to the \( fAc \) 3' end. D: Further DNA replication will generate an
intact deletion-carrying chromosome (with two telomere and a centromere) and a acentric
chromosome fragment. The acentric chromosomal fragment will eventually be lost due to lack
of the centromere.

Figure 7. Mapping the relative position of the probe npi286 to the \( P I \) locus. Genomic DNA
was digested with \( HindIII \) and hybridized with npi286. The signal in each lane is very similar,
indicateing that the locus corresponding to npi286 is not affected in these deletions.
Figure 1
Figure 2
Figure 3
Figure 4
Figure 5
Figure 6
Lane 1 P1-ww-773
Lane 2 P1-ww-739
Lane 3 P1-ww-660
Lane 4 P1-ww-772
Lane 5 P1-ww-765
Lane 6 P1-ww-771
Lane 7 P1-ww-758
Lane 8 P1-ww-775
Lane 9 P1-ww-755
Lane 10 P1-ww-756
Lane 11 P1-ww-761
Lane 12 P1-ww-796
Lane 13 P1-ww-666
Lane 14 P1-ww-774
Lane 15 P1-ww-1084

Figure 7
CHAPTER 4. GENERAL CONCLUSIONS

Transposable elements have long been considered a potential force in driving genome evolution due to their ability to induce large-scale chromosomal rearrangements such as deletions, duplications, inversions, and reciprocal translocations (McCLINTOCK 1953a, 1953b, 1978). The study of the molecular mechanism underlying these rearrangements did not begin until the early 1980s. Some of the first studies were focused on the maize sh-m5933 allele isolated by McCLINTOCK, but only partial structure was resolved because of the complexity and large-scale nature of the rearrangements (BURR and BURR 1982, COURAGE-TEBBE et al 1983). Two models had been proposed to explain the generation of the sh-m5933 allele, but neither of them can be tested easily (DORING and STARLINGER 1984, FEDOROFF 1989).

Study of chromosome breakage led to the discovery of transposable elements. Although Ds-induced chromosome breakage occurs during the whole life cycle of maize, broken chromosome are usually not passed to the next generation because chromosome breakage leads to large-scale deletion and loss of essential telomeres or centromeres. Therefore, Ds-induced chromosome breakage can be analyzed only in the somatic development stage. Using PCR and DNA sequencing (technologies available only in the middle of 1980s), ENGLISH et al. (1993) and WEIL and WESSLER (1993) found that Ds-induced chromosome breakage probably results from aberrant transposition events involving Ds ends on sister chromatids. They proposed that the aberrant transposition event creates a chromatid bridge, and the bridge breaks in the subsequent mitotic division (the observed chromosome breakage).

The aberrant transposition model predicts that Ds-induced chromosome breakage and
Ds-induced chromosome rearrangements (deletion and reciprocal inverted duplication) are two different outcomes of the aberrant transposition. If the deletion and the reciprocal inverted duplication were recovered from a single aberrant transposition event, both target site duplication and footprint resulting from that transposition event could be identified. English et al. (1995) identified deletions and duplications presumably generated by aberrant transposition events. They did not identify the characteristic target site duplication because these deletions and duplications resulted from independent transposition events.

By analyzing the Pl-ww-12:27-3 allele with complex chromosome rearrangement, we found this Pl allele could be generated by a transposition event involving Ac ends on sister chromatids. A model was proposed to explain the chromosome rearrangement which occurred in Pl-ww-12:27-3. This model is very similar to models for Ds-induced chromosome breakage proposed by English et al. (1993) and Weil and Wessler (1993). In our model, we refer to the transposition event involving Ac termini on sister chromatids as non-linear transposition rather than aberrant transposition, because this kind of transposition is normal from the biochemical point of view. We identified both the characteristic footprint and target site duplication resulting from a single non-linear transposition event. Because non-linear transposition can cause large-scale chromosomal rearrangements, it might have been contributing to the evolution of genome.

With some modification, the non-linear transposition model can explain the mysterious structural changes that occurred in generation of the sh-m5933 allele (Figure 1). The basic difference between the original NLT model and the modified NLT model is that the target site in the modified model is located in the unreplicated sh locus instead of the replicated region.
Figure 1. The modified NLT model. The small filled black and purple boxes indicate the telomeres of the short arm and long arm of the maize chromosome 9, respectively. The small black circles indicate the centromere. The red arrows are Ds elements, and the filled and open arrowheads indicate the Ds 5' and 3' ends, respectively. The hatched shapes indicate the sh sequences, and the vertical green line in the progenitor allele is the 8 bp sequence in the intron 7 of the sh locus. A. The DNA sequence at both sides of the sh locus are replicated, but the sh locus is still not replicated. The blue line is the DNA fragment between the double Ds and the sh locus. B. Ac transposase (small oval) recognizes a Ds 5' end on one sister chromatid and a 3' Ds end on the other sister chromatid as substrates. Cuts are made at these sites, so the non-linear transposon excises. C: The excised transposon inserts at the target site (the vertical green line) located in the sh locus. D. Further DNA replication generates a normal chromosome carrying a double Ds and deletion, a chromosome lack one telomere, and an acentric chromosomal fragment carrying inverted duplication. E. The free ends of two monotelomeric fragments fuse to form the sh-m5933 allele. The X indicates the position of the footprint created by transposition, and the vertical green lines are the target site duplication. The black line between the double Ds and the "one and a half" Ds in sh-m5933 is the 30 kb insertion sequence. The thick black lines in this region are inverted duplications, an additional copy of this sequence is located in the far distal region.

Note: The map is not to scale, lines which are shorter in the sh-m5933 allele than in the progenitor allele does not indicate a deletion; only the absence of a specific colored line indicates a deletion.
After reintegration, the non-linear transposon termini and some of their flanking sequences can replicate one more time in the same cell cycle by utilizing the replication origin of the *sh* locus. This additional round of DNA replication does not contradict the DNA replication license model because it does not require reinitiating an already replicated replication origin (Stillman 1996, Su *et al.* 1995). Completion of DNA replication will generate three molecules: a chromosome carrying deletion and a double *Ds* element, a chromosome with only one telomere, and an acentric chromosomal fragment carrying the inverted duplication. The free ends of the monotelomeric chromosome and the acentric chromosomal fragment may join together to restore the normal chromosome structure (one centromere and two telomeres). Fusion of the two molecules generate the complex rearrangements found in the *sh-m5933* allele. If the free ends do not fuse, both molecules will probably be lost because of large deletion or lack of centromere.

The *sh-m5933* structure predicted by the modified NLT model is the same as that of the partially solved *sh-m5933* structure, except that the modified NLT model predicts the existence of an additional half *Ds* located between the large inverted duplication (the sequence between the Double *Ds* and the *sh* locus in the *sh-m5933* progenitor allele). The half *Ds* element may be there, but it is far from the *sh* sequence, so it was not cloned by the earlier researchers. The target site duplications were found in the exact positions as the modified NLT model predicted. The footprint should be at the junction of the non-terminus end of the half *Ds* and one of the large inverted duplication. Because of the same reason that the half *Ds* element was not cloned, the footprint has not been resolved yet.
The research presented in Chapter 3 shows that non-linear transposition is an efficient way to generate large-scale nested deletions. We obtained 10 NLT-generated large-scale deletions from 1500 ears screened. All the 10 alleles are derived from large multiple-kernel colorless sectors (more than 8 kernels). If the smaller colorless sector were screened, we would probably obtain more NLT-generated deletions from the 1500 ear population.

Non-linear transposition probably also occurred at other loci in the maize genome. By crossing $C$, $ds$, $shI$ (homozygous for each allele) to $I$, $Ds$, $Sh$, $Bz$, $Ac$ ($I$, $Ds$, $Sh$, $Bz$, are on the short arm of chromosome 9, they are homozygous, but $Ac$ is hemizygous), McClintock isolated many $sh$ mutants in the F1 generation, and most of the $sh$ mutants could not revert to the wildtype phenotype (1952, 1953c). These new $sh$ mutants could have occurred by $Ac$ or $Ds$ insertion into the $Sh$ locus. However, some of the $sh$ mutants were also $bz$ mutant (1953c). It is very possible that these double mutants was caused by a NLT-induced deletion, because the double $Ds$ element distal to both $sh$ and $bz$ is in the right configuration to cause non-linear transposition.

The NLT events we analyzed were observed in the maize $PI$ locus; however, using plant transformation technology, a transgene construct containing the $Ac/fAc$ configuration as in the $P1$-$v 9D9A$ allele could be integrated anywhere in the genomes of a variety of plants. Therefore, deletions covering the whole genome could be generated. This would enable deletion mapping of the entire plant genome, following Benzer's classic approach (1961, 1962). The deletion can also be used to screen mutations, and the interstitial deletions would also be useful for screening mutations, and would facilitate gene cloning.
REFERENCES


Döring, H. P. and P. Starlinger, 1984 Barbara McClintock’s controlling element: now at DNA level. Cell 39: 253-259


McClintock, B. 1953a Induction of instability at selected loci in maize. Genetics 38: 579-599.

McCUNTOCK, B. 1953c Mutation in maize. Carnegie Institution of Washington Year Book. 52: 227-237


ACKNOWLEDGMENTS

First, I would like to acknowledge my mother, who let me go to school even though she could not afford so. Without her support and encouragement, I would be probably farming in the field right now.

I especially want to thank my major professor, Dr. Thomas Peterson for his excellent guidance and support, and I also want to thank him for correcting the English errors in my manuscript and thesis, which should not be part of his job. I wish to thank the current members of my POS committee, Dr. John Imsande, Dr. Jack Girton, Dr. Allen Miller, Dr. Daniel Voytas, and former POS committee member Dr. Alan Myers for their helpful input. I want to thank all the members in Tom's lab for their helpful discussion and cooperation.

Finally, I want to thank my wife who has done almost all the housekeeping work and taking care of our daughter, Angela, so I have time to do my lab work. I owe her too much to express in the form of language.