ORIGINATION OF *DS* ELEMENTS FROM *AC* ELEMENTS IN MAIZE: CHARACTERIZATION OF *AC* DERIVATIVES FROM *BZ-M39(AC)*

by

HONGBO PAN

A Dissertation submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Master of Science

Graduate Program in Plant Biology

written under the direction of

Professor Hugo K. Dooner

and approved by

_____

_____

_____

_____

New Brunswick, New Jersey

[October, 2008]

ABSTRACT OF THE THESIS

Origination of *Ds* element from *Ac* elements in Maize

-----Characterization of *Ac* derivatives from *bz-m39(Ac)*

By HONGBO PAN

Thesis Director:
Professor Hugo K. Dooner, Ph.D.

The maize genome contains a large number of both DNA transposons and retrotransposons. It is already known that retrotransposons are located in the highly methylated intergenic regions where recombination does not take place. However, DNA transposons are located in the eukaryotic regions of the maize genome, where they are able to interact with the host genes. It has been recently known that two transposable elements in direct orientation in the same chromosome may recombine occasionally with each other, which leads to the deletion of the intervening fragment. Based on this fact, it would be interesting to find out whether two elements in different homologues are able to recombine at meiosis. A series of *Ac* derivatives that originated de novo from *bz-m39 (Ac)* have been generated previously by Dr Dooner's lab. Some of these *Ds* derivatives appear to be large (>4kb) based on Southern blots. It may be possible to recombine two *Ds* elements with non-overlapping deletions to reconstruct a 4565-bp *Ac* element. Hence these derivatives' exact sequence composition needs to be defined. In this research project, I characterized 23 Ac^{-} derivatives, and the results indicate that most of these

derivatives are composed of fragmented transposon structures. Some of them are *Ds* elements with or without filler DNA; some of them are single-ended fractured *Ac* (*fAc*) elements with either 5' or 3' missing ends, yet not transposed, which have not been reported before, indicating that *Ac* may undergo not only internal deletions but also terminal deletions that are sometimes accompanied by deletion of its flanking sequence. Microhomology, which is the main mechanism to explain the non-homologous end-joining repair of deletions in *Ac*, was discovered at the deletion junctions of most of these derivatives. The next step is to select suitable derivatives to carry out the recombination experiment by crossing and measuring the frequency of meiotic recombination within the transposable element (TE) in *bz,* which will be compared then to recombination in *bz*. Here, I present the molecular analysis of these 23 derivatives.

Acknowledgement

I wish to sincerely thank my adviser Professor Hugo K. Dooner for providing me the opportunity to conduct lab rotation and accepting me as his graduate student to carry out research project in his laboratory. I would not have finished my graduate study at Rutgers without his kind-hearted help. I have really learned a lot of theoretical knowledge and experimental expertise as well from my lab experience especially in the genetics and molecular biology field. It was my great pleasure and honor to work with everyone in Dr. Dooner's lab.

I particularly appreciate Mr. Jun Huang for his great help and advice on my experiment and Dr. Qinghua Wang for her guidance on my experiment during my lab rotation period.

I also appreciate my committee members, Dr. Joachim Messing and Dr. Pal Maliga for their generous help and guidance on my thesis defense.

I would like to appreciate the Plant Biology Graduate Program at Rutgers University for providing me the Teaching Assistantship and Graduate Assistantship during my two years' graduate study at Rutgers University. I would not have the chance to accomplish my graduate study without this funding.

Table of Contents

Lists of tables

List of illustrations

I. INTRODUCTION

1.1 *Ac/Ds* system

Transposons or transposable elements (TEs) are pieces of DNA that can move around in the genome, and this process is called transposition. Transposons can mutate genes' function when they jump into genes. Basically, transposons can be divided into two types based on the mechanisms of transpositions. Type I transposable elements, or retrotransposons, can move around in the genome by being first transcribed to RNA and then reverse transcribed back to DNA by reverse transcriptase; Type II TEs are DNA transposons which can directly move around in the genome by transposase-mediated "cut and paste" or "copy and paste" mechanisms. The transposon *Activator (Ac)* was the first autonomous element described in maize (*Zea mays*) by McClintock (1949). McClintock discovered that insertions, deletions and transpositions of these TEs could change the color of maize kernels. Autonomous elements like *Ac, Spm*, and *MuDR* are able to transpose themselves, however, their counterparts like *Ds*, *dSpm* and *Mu1*, respectively, are not able to transpose alone and require the presence of their progenitors for transposition (McClintock 1956a). McClintock (1956b, 1962, 1963) also reported that an *Ac* element mutated to *Ds*. Molecular characterization of such *Ds* elements revealed that *Ac* mutated to *Ds* by experiencing internal deletions (Fedoroff 1989). Unlike the highly conserved structure of *Ac* elements (Fedoroff et al.1983; Behrens et al.1984; Muller-Neumann et al. 1984; Pohlman et al. 1984), *Ds* structures are pretty diverse, which have also been proved by our study. State I or "*double Ds*" elements have nested *Ds* insertions or multiple *Ds* end sequences in close proximity (Doring et al. 1984, 1989, 1990; Weck et al. 1984; Ralston et al. 1989; Weil and Wessler 1993). State II elements usually suffer

deletions of *Ac* with various sizes and internal structures (Kunze and Weil 2002). For a long time, the mechanism of how *Ac* mutates to *Ds* was not clear. A series of *Ds* elements have been analyzed and Dooner et al (1986) found a 3-bp direct repeat adjacent to the deletion junctions of a *Ac* derivative *bz-m2 (D1),* which provided some clues to explain how the internal deletion arises in *Ds*. Subsequent molecular characterization of several *Ds* elements originated *de novo* from *Ac* further provided evidence by sequencing that *Ds* elements are generated by internal sequence deletion of *Ac* elements when DNA synthesis mechanism repairs the double strand break (DSB) due to *Ac* excision event, and they also found that direct repeats, or microhomologies, prevalently exist adjacent to the deletion junction regions (Yan et al, 1999). Basically these analyses established that *Ds* most likely arises *de novo* as the result of *Ac* excision.

Ac excision leads to DSB. DSB is generally repaired through two types of mechanisms. First is homologous recombination, which is sub-divided into two types of mechanisms (**Figure 1**). One is called "synthesis dependent strand annealing (SDSA)". One strand invades and initiates DNA synthesis. DNA unwinds from the template as the synthesis process goes on, and the other tail anneal to this unwound DNA strand. The other one is called "single-strand annealing (SSA)", a process initiated when DSB forms between two repeats with the same orientation. Single strand is generated adjacent to the break so as that complementary strands can anneal with each other. Flap is removed and the break is ligated to form a strand.

The second model for repair of DSB is not dependent on the homologous template and is called non-homologous end joining (NHEJ). NHEJ usually is based on microhomology to guide repair of DSB. Microhomologies are present on the single stranded tails of the DNA ends, which, if compatible, will join together. However NHEJ is also mutation-prone because NHEJ leads to the deletion of one short repeat and the internal sequence between the two direct short repeats, as well. A model of DSB repair by NHEJ is proposed and discussed in detail based on our research results (see discussion).

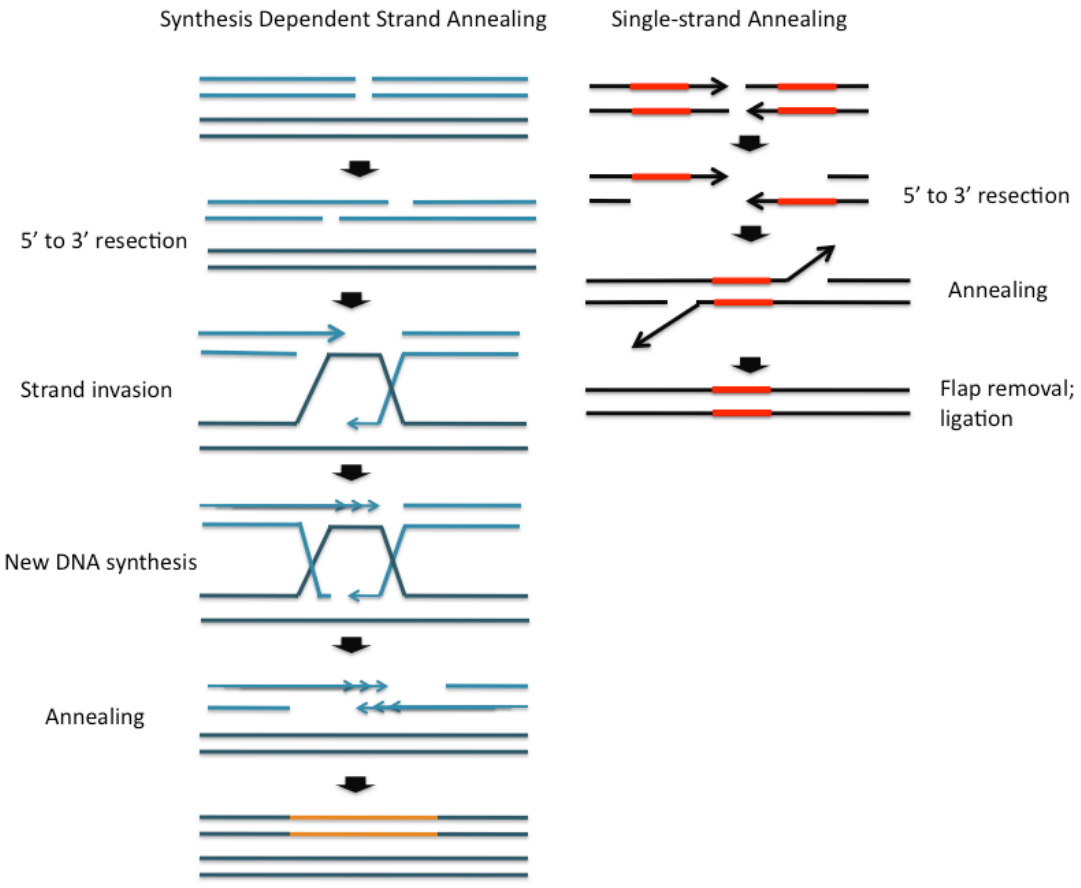


Figure 1. Two types of mechanisms of homologous recombination for repair of DSB. The left figure shows the repair is dependent on the DNA synthesis of the homologous template; the right one shows how

the repair precedes when DSB takes place between two short repeat orientated with the same direction. This figure was modified based on http://asajj.roswellpark.org/huberman/DNA_Repair/SDSA_SSA.gif.

1.2 Meiotic recombination

Currently we do not know whether *Ac/Ds* can recombine at meiosis. It has been shown that two elements may recombine occasionally with each other deleting intervening fragment, if these two elements are located on the same chromosome in direct orientation (Huang and Dooner, 2008). However, we do not know whether two elements are also able to recombine with each other in meiosis if they are in different homologues, and what the frequency is if they are able to, and this is the long-term research objective for our *Ds* element project. One way to test this is to reconstitute a 4.6-kb *Ac* element at the *bz* locus by recombination between two *Ds* elements with non-overlapping deletions. We can analyze the different colors of kernels with different genetic backgrounds to verify whether two elements on different homologous can recombine or not, since the *bz-m39 (Ac)*, if can be reconstituted, will show a purple spots on a bronze background, which will be very different from the bronze color of bz-s derivatives. Actually Yan *et al.* have generated several *Ds* derivatives from *bz-m2 (Ac)* previously; however, we cannot get the intact 4.6-kb *Ac* element since the deletion regions of these *Ds* derivatives all overlap.

Fortunately the new *Ds* derivatives from *bz-m39 (Ac)* can be used for our purpose to test whether two different *Ds* homologues can recombine in meiosis, since some of these *Ds* derivatives look like the *Ac* progenitor in size (over 4 kb), when tested by Southern blots

(Dooner and He, unpublished). And my master research project objective is to identify at least two transposable elements that can be used to reconstitute the 4.6-kb *Ac* element by recombination. *bz-m39(Ac)* is located in the *bz* gene 5' UTR (untranslated region), at a position 39 to 32 bp upstream of the start codon (Huang and Dooner, unpublished); therefore, *Ac* excisions restore *bz* gene function and appear as purple revertants (*Bz'*). *bz-m39 (Ac)* can also mutate to stable bronze phenotype. Among the 23 *bz-s* derivatives, the first 12 (*bz-s*39.1 to *bz-s*39.13) have all been tested by genetic study but not the later 11 selections (W8040-1 to W8042-4). Under this condition, the sequences of these *Ac* derivatives need to be defined so as to determine which derivatives are suitable to recombine to reconstruct the 4565-bp *Ac* element. All derivatives were characterized molecularly as described in Materials and Methods.

Generally the *Ac* derivatives, except for pollen contaminants and hypermethylated *Ac*, consist of two classes; *Ds* elements and single-ended *fAc* elements. These derivatives possess various structures, such as, some *Ds* have filler in the deletion junction, some do not; some single-ended *fAc* elements are 5' end defective, some are 3' end defective.

II. MATERIALS AND METHODS

2.1 Genetics stocks: all the stocks in the study have the same background of the inbred line W22.

Bz-McC (purple): the normal progenitor allele of the *bz-m39 (Ac)* mutation.

bz-m39 (Ac) (purple spots on a bronze background): an allele generated by an insertion of the 4565-bp *Activator (Ac)* element in the *bz* 5' UTR, at position 39 to 32 bp upstream of the start codon.

*bz-R (*bronze*):* the *bz* reference allele, associated with a 340-bp deletion that extends from within the single intron to the second exon of *bz* and includes the *Ac* insertion site in *bz-m2* (Rhoades 1952; Ralston *et al.* 1987,1989).

2.2 Selection and analysis of new *Ds* derivatives:

The mutations *sh* (shrunken endosperm) and *wx* (waxy endosperm) were used as markers flanking *bz* gene. Heterozygous *Sh bz-m39 (Ac) Wx*/ *sh bz-R wx* was crossed to *sh bz-R wx* pollen parents and then stable derivatives were selected as *Sh bz Wx* single-kernel events. The derivatives were then crossed to a *Ds* reporter stock to score for the presence of *Ac* and to an *Ac* stock to determine whether a new *Ds* element had originated at the *bz* locus. It should be noted that I first got *bz-s* derivatives in Oct 2007, and after that I got selections W8040, W8041 and W8042 in April 2008. Those *bz-s* derivatives have crossed with *Ac* and *Ds* testers so we know they are inheritably stable. However,

selections W8040, W8041 and W8042 have not been tested by crossing with a *Ds* reporter stock or an *Ac* stock yet. So we call them as selections and some of these selections are actually pollen contaminants, not our desired new derivatives, which will be described later on in results. For those selections that have been demonstrated as the new derivatives by sequencing, we have given them new *bz-s* derivatives names respectively.

2.3 DNA extraction:

Leaves were ground into powder. 25 ml DNA Extraction Buffer were added and mixed by shaking. Materials were kept on ice for 30 min. Chloroform was added into the ground materials till up to 50 ml, mixed thoroughly, and then were centrifuged at 4000 rpm for 15 min. The upper phase was taken into a fresh tube and the same volume of phenol/chloroform (1:1) was added, mixed by shaking and then centrifuge again at 4000 rpm for 15 min. The upper phase was taken into a fresh tube again and the same volume of ethanol was added, and centrifuge for the 3rd time at 4000 rpm for 15 min. The pellet were left in the air till it turned a little dry, and then the pellet were re-suspended in 5 ml 1X TE (pH8.0) with RNase (25 μg/ml), 37°C 30 min. 5.7 ml 2% CTAB and 0.7 ml 5 M NaCl were added and centrifuged at 4000 rpm for 15 min. 10 ml 70% ethanol containing 0.15 M NaCl were added and kept the pellet was shaked slowly in 4°C cold room over-night. The pellet was centrifuged at 4000 rpm for 5 min. 10 ml 70% ethanol was added in for washing purpose. 1X TE was added to dissolve the pellet.

2% CTAB: 2 g CTAB (N-Hexadecyl-N, N, N-trimethylammonium bromide), 50 mM Tris-Cl (pH7.5), 10 mM EDTA (pH8.0)

2.4 Southern blotting:

After electrophoresis, 0.8% agarose gels were rinsed with 500 ml 0.25 N HCl till the blue dye turned yellow. The gels were washed with water twice and then rinsed with 2000 ml buffer containing 0.5 N NaOH and 1.5 N NaCl for 30 min on shaker. The DNA was transferred from agarose gel onto membrane. 35 ml 65℃ Church buffer were mixed with the dried membrane for pre-hybridization and incubated in the oven for at least 1 hour. 25-100 ng DNA was mixed with ^{32}P, heated for 5 min at 100°C, and put back to ice immediately for 1 min. The probe was added into the bottle with the membrane, mixed well, and hybridized overnight. The membrane was washed by 0.5X SSC with 0.1X SDS or 0.2X SSC with 0.1% SDS till the signal was down to an acceptable strength for later detection.

2.5 PCR and sequencing:

For those derivatives whose *Ds* or single-ended *fAc* elements were not transposed, long range PCR was carried out with primer combination of *bz*-y5 and del2, which are located at the sequences flanking the *Ac* insertion site (**Figure 2**). For those derivatives whose *Ds* appeared to have transposed somewhere else in *bz*, several other *bz* primer combinations were tried to identify where it was transposed to with the common Qiagen PCR system first, and then long range PCR was performed to amplify those regions containing *Ds* elements. Primers sequences are listed in Table 1.

PCR amplification was carried out following the instructions of Roche Long Range PCR kit. The long range PCR program is conducted as follows: 1st round amplification for10 cycles (92°C 10 s; 56°C 15 s; 68°C 5.5 min); 2nd round amplification for 25 cycles (92°C 10 s; 56°C 15 s; 68°C 5.5 min with increased 20 s/cycle). 10 μl PCR amplification system consists of 2 μl 5X PCR buffer (+Mg^{2+}), 1 μl DMSO, 0.3 μl dNTP, 0.2 μl forward primer, 0.2 μl reverse primer, 0.1μl polymerase, 1 μl 150ug/ml DNA, 5.2 μl ddH_2O. For Qiagen PCR kit, 10 μl amplification system consists of 1 μl 1X PCR buffer (w/ Mg^{2+}), 2 μl 5X buffer Q, 0.2 μl forward primer, 0.2 μl reverse primer, 1 μl DNA, 0.3 μl dNTP, 5.6 μl ddH_2O. PCR products were fractionated on a 0.8% agarose gel containing ethidium bromide and visualized on a UV light machine.

The amplified PCR fragment was cut from the agarose gel and purified with Gel Extraction Kit (QIAGEN). 10 μl ligation system consists of 5 μl 2X ligation buffer, 3.5 μl PCR product, 1 μl T4 ligase and 0.5 μl p-GEMT easy vector (Promega, Madison, Wisconsin). 5 μl ligation product was mixed with 50 μl competent cells for 30 min on ice and then was put into 42°C water bath for 90s heat shock. After that, competent cells were put into ice for 5 min and then were screened on the plates with Ampicilin antibiotics, IPTG and X-gal overnight at 37°C. The putative positive clones were picked from the LB culture plates and cultured with 3 ml LB liquid medium containing 100 μg/ml Ampicilin for 37°C overnight at 250 rpm. Plasmid DNA was extracted with QIAprep Spin Miniprep Kit. Plasmid was then cut by *Eco*RI restriction enzyme to check insertion size. Cycle sequence reaction was conduced with ABI 96-well plates. For each well there is a 10 μl sequencing reaction containing 1 μl plasmid, 0.5 μl BigDye Terminator v3.1, 1.75 μl 5X ABI sequencing buffer, 1 μl primer (5pmol final) 5.25 μl

H_20 and 0.5 µl DMSO. Cycle sequence reaction was conducted as 96°C 10 s, 50°C 5 s, 60°C 4 min for 60 cycles. 10 µl H_20 and 5 µl 1% SDS was added to the bottom of each well. Plates were then incubated at 60°C for 6 min and 10 min at room temperature to cool down. 5 µl 125mM EDTA was added and then spun down. 60 µl 100% ethanol was added and the plate was sealed with Qiagen tape and inverted 3-4 times and then incubated for 30 min in the dark at room temperature. Plates were then centrifuged for 35 min at 3500 rpm at 4°C. Liquid was decanted carefully. 200 µl 70% ethanol was added and centrifuged for 15 min 3500 rpm at 4°C. Liquid was decanted carefully. Plate was dried in the dark hood for 2-3 hours or alternatively dried in the drawer overnight. 10 µl formamide was added to each well and briefly spun down. Finally the samples were sequenced by 3730 ABI sequencing system (Applied Biosystem, HITACHI).

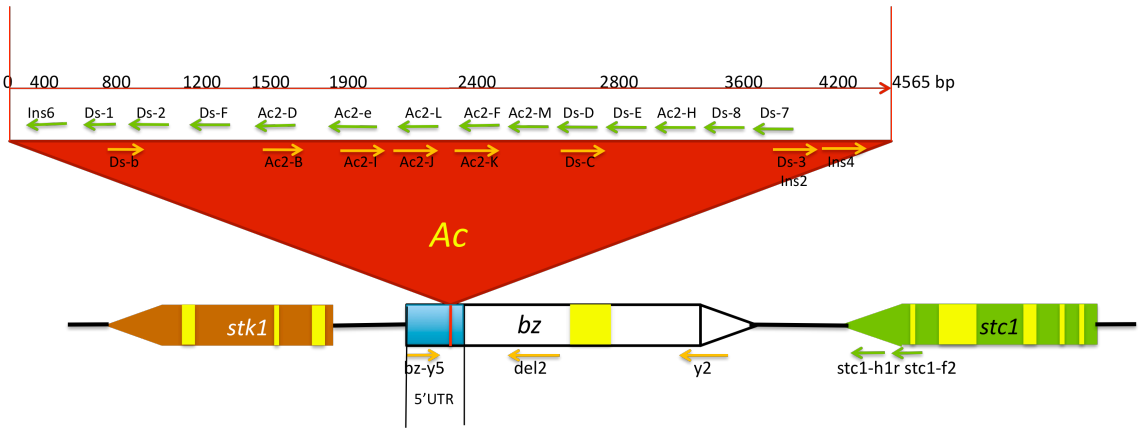


Figure 2. Structure of the parental chromosome, showing the location of genes, transposons and PCR primers. *Ac* is inserted at a position 39 to 32 bp upstream of the start codon of *bz*. *Ac* is shown in red triangle. Genes are represented as pentagons pointing in the direction of transcription and are not drawn to scale. Introns are drawn in yellow rectangles; exons and UTRs in *stk1, bz* and *stc1* are respectively drawn in brown, white and green. Intergenic sequence is represented by a black line. Primers are shown as arrows.

Table 1. Oligonucleotide primers used for PCR and sequencing reaction.

bz-y	GGGTGTGTCCAGAATGTACCTA
y15	GCGAGGACTCGTGGCCGCA
y1	ATCCGCTGGTCGCCGAAGAA
y2	ACCGTGCCGAAGCTGACGTA
y6	ATTGCGCGCGGGTTTGATGA
bz-3R	AAACCTCTGAACAGCAAGACGACC
bz-N	CACGCTCTCGTTCCTCTCCA
Del2	GGACGCGGTGGAGAGGAACGA
Stc1-H1r	AACGAGGAGTGGCTGAACAT
Stc1-F2	GGAGGTAAGATGGCGCGAACAAGG
Ds-1	GACACAACAGCTTGGTGCAA
Ds-2	GTGACTTGTGAGCACCTGGA
Ds-F	GAGGCTTCAGCTTTTTCTTTTCAAC
Ac2-D	GCTTCTTCTTCGGGTTCAGGTTGT
Ac2-E	GTGAGGGCGCAGAGACTT
Dc-E	GTGAGGGCGCAGAGACTT
Ac2-L	AAGCACACCTCACATGAAAGA
Ac2-F	CCCACTGCAAAGGAGAAGATT
Ac2-K	AATGAAGTGTGCTAGTGAATGTGA
Ac2-H	TTCAAAAGGGGTTCAGACAT
Ac2-J	TGTTCTTGCTGTAAAATCTTCTCC
Ac2-I	TCTTTCATGTGAGGTGTGCTTGTC
Ds-B	GATTGGCCAAGTTGATGTCTACC
Ac2-M	AAGGCATCCCTCAACATCAAATA
Ds-D	CCATCTTCCACTCCTCGGCTTTAG
Ds-3	ATCTCACTGCATGCGCCTTGTC
Ds-7	TAGCACCTCGAGATCACCAA

III. RESULTS

3.1 *Ds* elements

Ten *Ds* elements were found among the *bz-s* derivatives. They will be described first.

3.1.1 *bz-s39.3* (family D6232)

bz-s39.3 is a 3883-bp *Ds* element. It has a 682-bp deletion without any filler DNA at the junction. The red triangle represent the *Ac* progenitor and rectangle space in the triangle represent the deletion junction; GCCTTGTCT and TTGGAAA represent the 5' end and 3' end flanking sequences adjacent to the deletion junction in the *Ds* element, respectively; sequence ctagcggc marked in blue in the rectangle represents the footprint and target site duplication (TSD); the sequence in red in the rectangle represents the *bronze* sequence. The same representation applies for all of the following the schematic figures of derivatives structures.

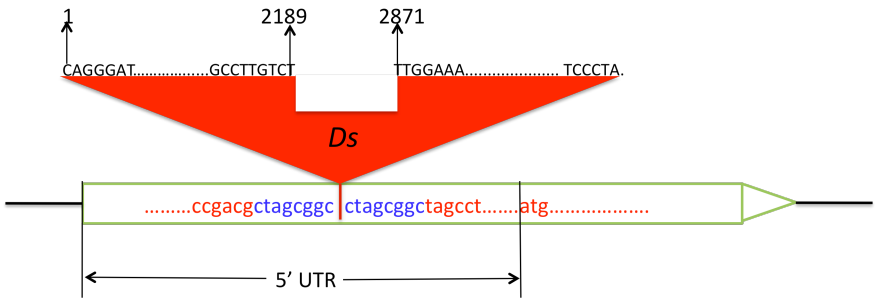


Figure 3. Schematic structure of *bz-s39.3*. The coordinates represent the positions of last nucleotide retained relative to the *Ac* sequence (GenBank X05424). The *bz* sequences are lowercased and in red color, except for the target site itself (ctagcggc), TSD and footprints are in blue color; the *Ac* and derivatives sequences are all capitalized and in black color. The same representation applies for all the following figures. *Ds* element is still retained in its original progenitor's insertion site, and this is consistent with all of the following derivatives except for three contaminants.

3.1.2 *bz-s39.5* (family D6234)

bz-s39.5 is a 2951-bp *Ds* element. In the internal deletion region a 200-bp sequence was discovered that was duplicated exactly starting from the first nucleotide at the 5' end of the original *Ac* (CAGGGATGAA…) and then inserted into the deletion junction. This filler has the same orientation as the 5' end. This element should not be a chromosome breaker since the two 5' ends' orientation is the same, so it would not be able to provide the substrate for alternative transpositions that would create dicentric chromosomes (Doring *et al,* 1984; English *et al,* 1993; Weil & Wessler *et al,* 1993).

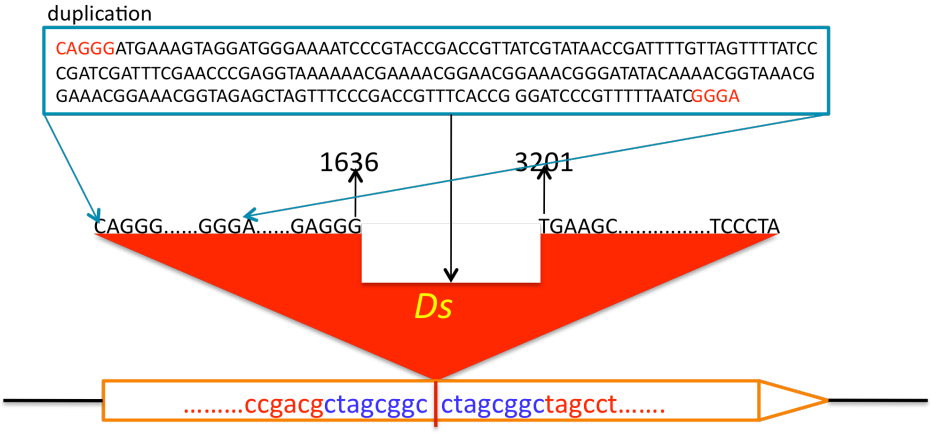


Figure 4. Schematic representation of *bz-s39.5*. BLAST search result indicates that the 200-bp duplication between the internal deletions endpoints comes from the 5' end of *Ac* progenitor in *bz-s39.5*. *bz-s39.5* is unusual in having three ends, two 5' ends and one 3' end.

3.1.3 *bz-s39.6* (family D6235)

bz-s39.6 is a 3361-bp *Ds* element without any extra sequence in the deletion junction.

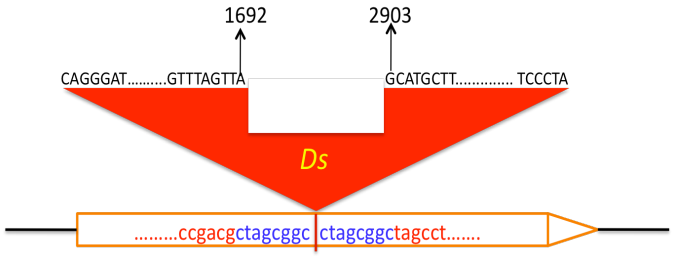
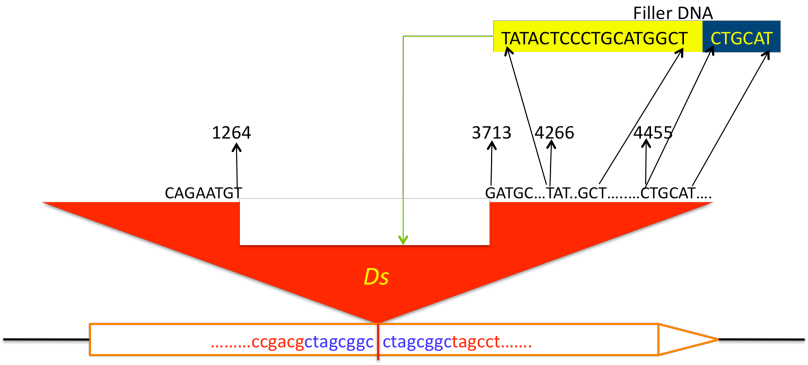


Figure 5. Schematic structure of *bz-s39.6*.

3.1.4 *bz-s39.7* (family D6236)

bz-s39.7 is a 2140-bp *Ds* element. A 24-bp filler was discovered in the deletion junction and BLAST showed that this filler DNA was duplicated from the *Ac* element. The first 18 nucleotides correspond to nucleotides 4266 to 4283, although the first 14 of these are also duplicated in *Ac* at position 4448 to 4461. The last six nucleotide CTGCAT in filler DNA (in red) are of unknown origin.




```
>gb|AF355378.1|AF355378  Zea mays bz-m2(Ac) mutable allele, genomic sequence
Length=5772

 Score = 36.2 bits (18),  Expect = 0.006
 Identities = 18/18 (100%), Gaps = 0/18 (0%)
 Strand=Plus/Plus

Query  1     TATACTCCCTGCATGGCT  18
             ||||||||||||||||||
Sbjct  4266  TATACTCCCTGCATGGCT  4283


 Score = 28.2 bits (14),  Expect = 1.4
 Identities = 14/14 (100%), Gaps = 0/14 (0%)
 Strand=Plus/Plus

Query  1     TATACTCCCTGCAT  14
             ||||||||||||||
Sbjct  4448  TATACTCCCTGCAT  4461
```

Figure 6. Schematic structure of *bz-s39.7* and BLAST search results of 24-bp filler DNA in the internal deletion region of *Ds bz-s39.7*.

3.1.5 *bz-s39.10* (family D5077)

bz-s39.10 is a 3961-bp *Ds* element. It has no inserted sequence in the deletion junction.

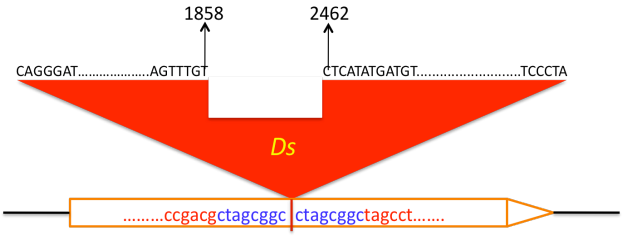


Figure 7. Schematic structure of *bz-s39.10*.

3.1.6 *bz-s39.11* (family D5078)

bz-s39.11 is a 3672-bp *Ds* element. No inserted sequence was found in the deletion junction.

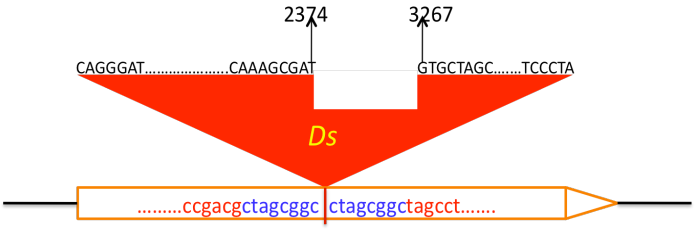


Figure 8. Schematic structure of *bz-s39.11*.

3.1.7 *bz-s39.14* (selection W8040-1)

bz-s39.14 is a 3105-bp *Ds* element. No filler exists in the deletion junction.

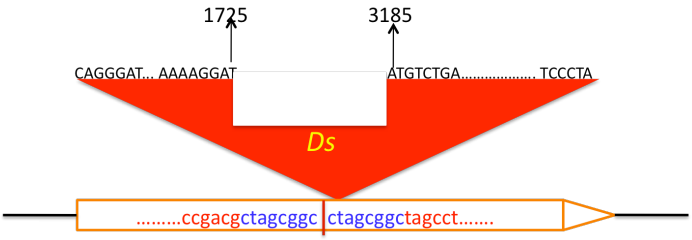
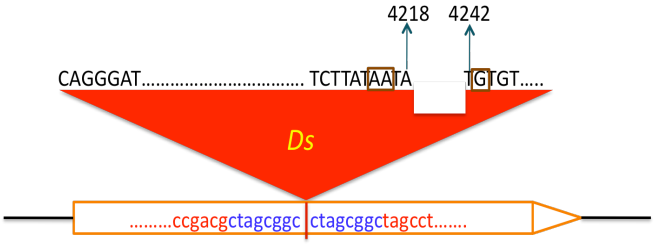


Figure 9. Schematic structure of *bz-s39.14*.

3.1.8 W8040-2

W8040-2 is a 4541-bp *Ds* element. It has the largest size of *Ds* among all of the derivatives in this research project. Three nucleotides adjacent to the deletion junction were found changed to other nucleotides. The 4th and 3rd nucleotides preceding the deletion junction were changed from the original GT to AA, and the 2nd nucleotide following the deletion junction was changed from the original A to G (**Figure 10**). The mechanism by which this happened is not clear yet. Unfortunately no seed were recovered from this selection, so it was not assigned a *bz-s* derivative number.



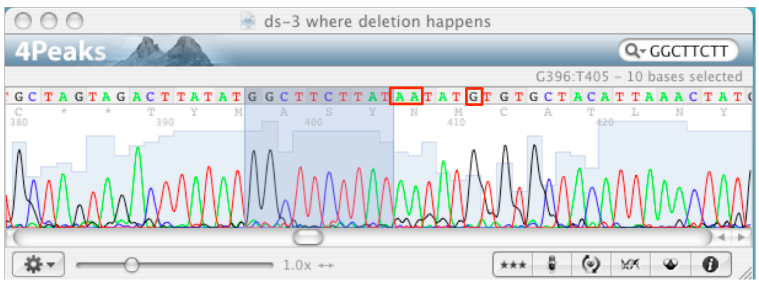


Figure 10. Schematic strucutre of W8040-2 (top) and sequencing result shows the statement is reliable (botom).

3.1.9 *bz-s39.18* (selection W8042-4)

bz-s39.18 is a 3886-bp *Ds* element. There is no filler DNA in the deletion junction.

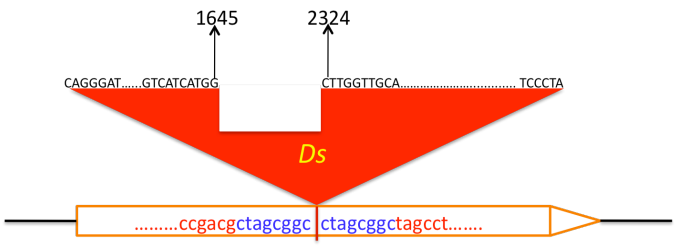


Figure 11. Schematic structure of *bz-s*39.18.

3.2 Single-ended *fAc*

Five single-ended elements were found. These elements have been referred to as *fractured Ac (fAc)* (Ralston *et al,* 1989; Zhang & Peterson, 1999; Zhang & Peterson, 2004). Two were deletions of the 5' end, similar to the previously described *fAc*s, and three of the 3' end, which had not been described before.

3.2.1 *bz-s39.1* (family D6242)

bz-s39.1 has a 1595-bp long *fAc*, which is defective in its 5' end. A 99-bp *bz* sequence (shown in light blue box) exists between *fAc* and the 8-bp footprint (CTAGCGGG). This 99-bp sequence was duplicated from the *bz* sequence immediately following the 1595-bp fractured element (shown in red). There is a typical 8+8 excision footprint at the prior site of *Ac* insertion: one of the bases has been changed to its complement (CTAGCGGG/CTAGCGGC). The repair of *Ac/Ds* excision sites usually creates some sequence alternations, or footprints at the excision sites (Britt & Walbot; Doseff et al; Rinehart, et al; Scott et al), and this was also found in *bz-s*39.1. The last base C in TSD---target site duplication (CTAGCGGC) was changed to G.

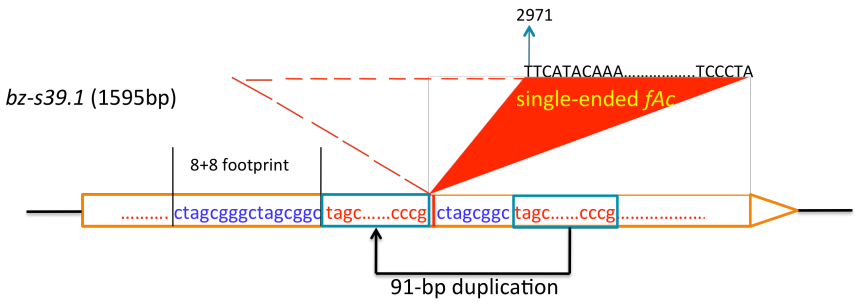


Figure 12. Schematic structure of *bz-s39.1*. The full sequence of filler DNA is shown in the light blue rectangle.

3.2.2 *bz-s*39.4 (family D6233)

Sequencing results indicate that *bz-s39.4* lacks the *Ac* 5' end, which makes it similar to *bz-s39.1*. *bz-s39.4* has a 1730-bp long *fAc* sequence and a 27-bp duplicated *bz* sequence (TAGCCTAGCCGAACAGCCTGAGCGCGC) exists between the *fAc* and a target site duplication (CTAGCGGC). These 27 bases were duplicated starting from the 9th base downstream of this fractured element.

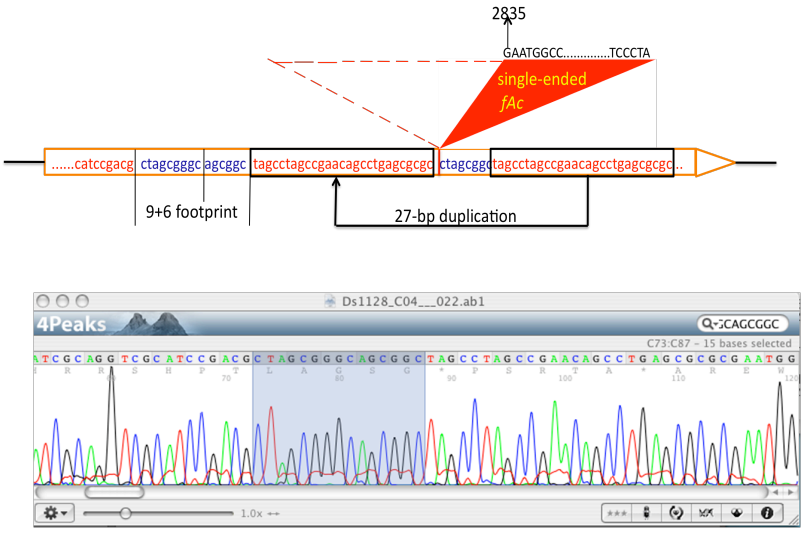


Figure 13. Schematic structure of *bz-s*39.4.

3.2.3 *bz-s*39.9 (family D5076)

A previously unreported phenomenon was discovered in *bz-s*39.9 and, later on, in *bz-s*39.13. First, I tried to carry out long range PCR with primers *bz*-y5 and del2, however the PCR did not work. Then, I realized that some deletion events might have happened leading primers not to work, so I sub-cloned the 5' and 3' ends separately. Several *Ac* internal primers were used, like Ac2-E, Ac2-L, Ac2-M, etc, combining with *bz*-y5 (**Figure 2**). From the agarose gel, we can clearly see that sequence starting from the 5' end to Ac2-M is intact, and then I used Ac2-I, a primer located upstream of Ac2-M, combined with Stc1-h1r, a primer located in the *stc1* gene, to clone its 3' end. Finally, the sequencing results indicate that *bz-s39.9* has some unusual characteristics that had not been reported yet. Interestingly, *bz-s39.9* is missing 1912 bp from the *Ac* 3' end and a 2447-bp sequence that includes the entire *bz* gene and most of the intergenic sequence

between *bz* and the distal *stc1* gene; meanwhile, this fractured element was not transposed and a 16-bp filler DNA (AAAGTCTCAAATCTCA), indicated in the cartoon below, was found in the junction region. BLAST is not feasible to find where this filler comes from; the filler is too short thus microhomology cannot be analyzed since no sequence context is available. This is the first report of both a 5'-end *fAc* and of a single-ended *fAc* element not associated with an excision footprint; meanwhile, the deletion of the 3' end of *Ac* caused a large adjacent deletion of host DNA.

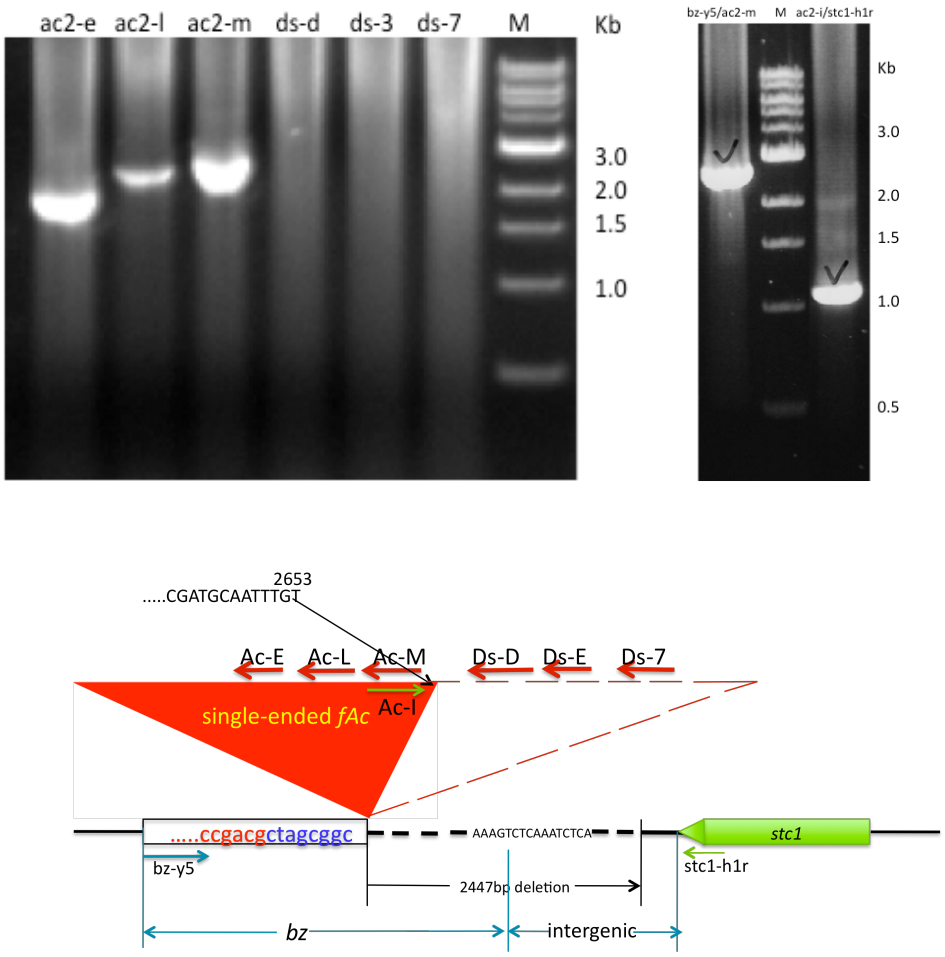


Figure 14. Long range PCR was carried out to sub-amplify *Ac* derivative in *bz-s39.9* (top); Schematic representation of *bz-s39.9* (bottom). The 2447-bp deletion region includes entire the *bz* gene fragment immediately following the *fAc* and partial intergenic region between *stc1* gene and *bz* gene.

AAAGTCTCAAATCTCA is filler DNA existing in the deletion junction. The origin of this extra sequence needs to be defined.

3.2.4 *bz-s39.13* (family D5080)

bz-s39.13 shares some characteristics with *bz-s39.9*, with defects in the *Ac* 3' end, no excision footprint, and with 3' end adjacent *bz* sequence deleted consequently. The same sub-clone strategy was used to amplify its 5' end and 3' end parts separately (**Figure 15**), and sequencing results indicate that a 1677-bp adjacent sequence was deleted, including coding and 3'UTR *bz* sequences downstream of this single-ended *fAc*. The *fAc* element in *bz-s39.13* has a smaller size (1728-bp) than the *fAc* in *bz-s39.9* (2653-bp) and caused a shorter adjacent sequence deletion (1677-bp) than *bz-s39.9* (2447-bp). The difference between *s39.13* and *s39.9* is that no filler DNA exits in the junction.

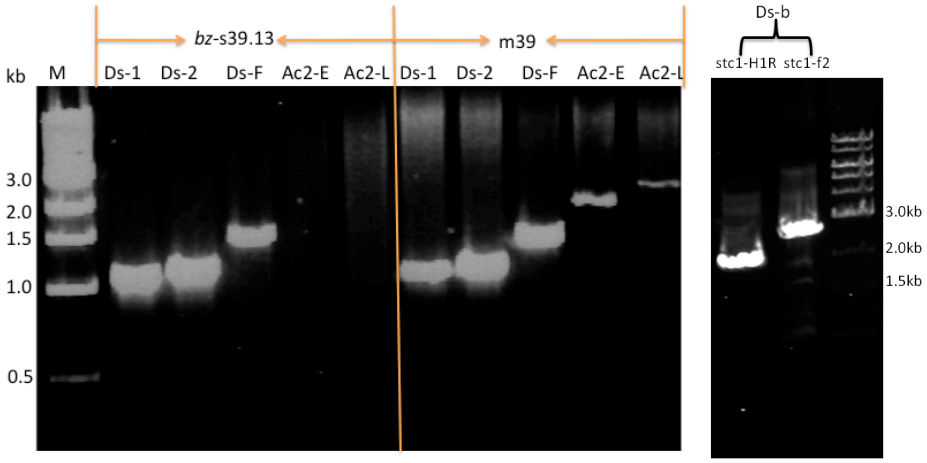


Figure 15. *Ac* derivative *bz-s39.13* was amplified with the same strategy as have been used in *bz-s*39.9 by long range PCR. Primers positions can be iendified with Figure 2.

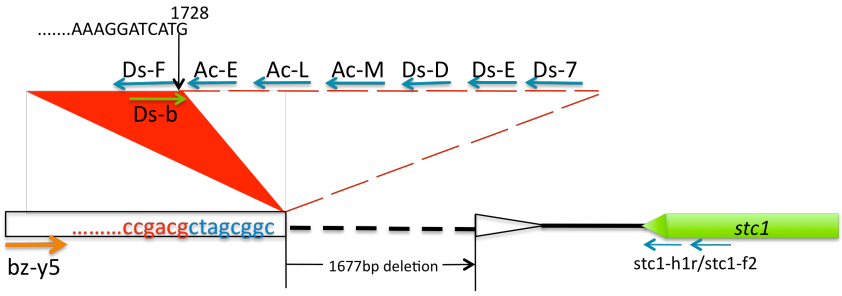


Figure 16. Schematic representation of *bz-s39.13*. It is a 1728 5' *fAc* element accompanied by 1677-bp deletion of *fAc* 3' end adjacent *bz* sequence.

3.2.5 W8040-3

W8040-3 is missing 2546 bp from the *Ac* 3' end plus 58 bp of adjacent *bz*, which is similar with what has been found in *bz-s*39.9 and *bz-s*39.13, though the deletion suffered in *bz* gene was much shorter than the previous two. In addition, the 3'end deletion is not associated with an excision footprint and no filler exists in the deletion junction. The 5' *fAc* element in W8040-3 is 2019 bp long. Like W8040-2, this selection was lost in the greenhouse, so it was not assigned a *bz-s* number.

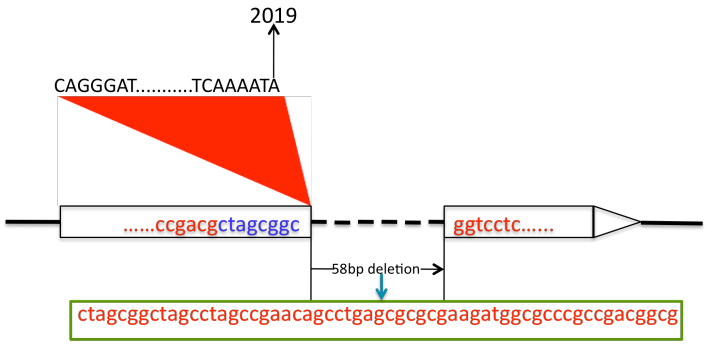


Figure 17. Schematic representation of W8040-3.

3.3 Hypermethylated *Ac*

Five possible cases of *Ac* inactivation by hypermethylation were found.

3.3.1 *bz-s39.2* (D6231), *bz-s39.16* (W8042-1), W8042-2, *bz-s39.17*(W8042-3)

These four stable bronze derivatives were postulated to carry hyper-methylated *Ac*s since no deletion or insertion in the *Ac* element discovered by sequencing. Conrad *et al* (2007) also reported several derivatives containing an apparently full-length *Ac* element that is inserted at the original locus based on PCR product size, though no *Ac* activity was detected. They accounted for the loss of *Ac* activity as due to either too short deletions (<50 bp) suffered in *Ac*, which cannot be detected in PCR assay, or epigenetic silencing assisted by increased methylation of internal *Ac* sequences (Schwartz and Dennis 1986; Chomet *et al.* 1987; Kunze *et al.* 1987, 1988; Fusswinkel *et al.* 1991; Brutnell and Dellaporta 1994; Brutnell *et al.* 1997). Our sequencing results have eliminated the first possibility, and the second possibility needs to be proved in our case with some methylation-sensitive restriction enzyme, such as *Pvu*II or *Pst*I, and later on Southern blots. At least Conrad *et al.* (2007) did find evidence of heavy methylation in *Ac,* which leads to losses of *Ac* activity, by Southern blots.

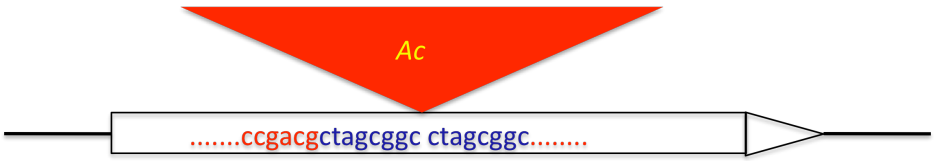


Figure 18. Schematic representation of *bz-s39.2* W8042-1 W8042-2 W8042-3 and later on *bz-s39.12.*

3.3.2 *bz-s39.12* (family D5079)

This derivative presented problems for analysis, as described below. D5079 is the pedigree number of *Ac* derivative *bz-s39.12* that was analyzed firstly. D5079 is the self

progeny of a putative *Sh bz-s39.12*/*sh bz-R* heterozygote (**Figure 19**). The genetic distance between *shrunken* gene and *bronze* genes is 2 cM, which means that the frequency of recombination between *Sh* and *bz-R* is 2%. Primers *bz*-y5 and del2 were used to amplify the insertion sites and sequencing revealed that the material that I analyzed, D5079, was actually *Sh bz-R/ sh bz-R* recombinant.

I tried another material later on, D4282, which is heterozygous *Sh bz-s*39.12/*sh bz*-R (**Figure 19**). Several primer combinations were used to amplify the 5' and 3' ends respectively. From **Figure 21**, it is hard to postulate any deletion happened in D4282 from the agarose gel. Then those two bands with the largest size (yellow rectangle in **Figure 21**) were cut from the gel and sequenced. No deletion was found in this derivative, so it may be hypermethylated.

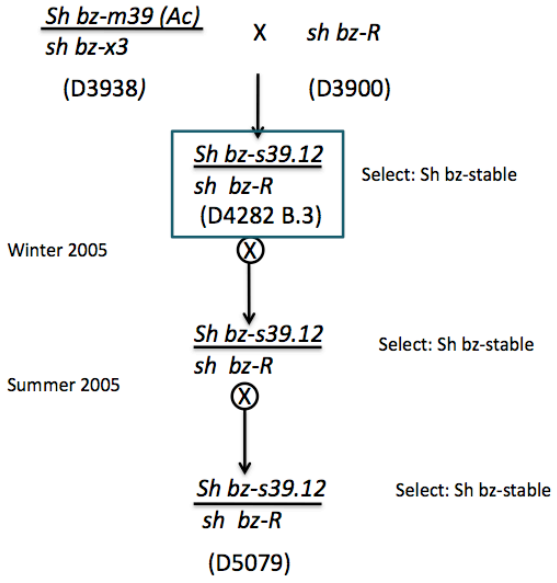


Figure 19. Genetic selection strategy of stably inherited *Sh bz-s39.12*.

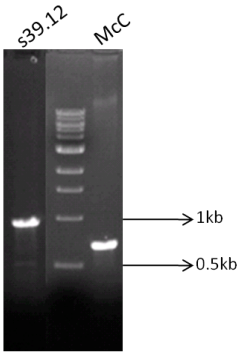


Figure 20. PCR was carried out with *bz*-y5 and del2 to amplify the insertion site with normal Qiagen enzyme. Considering the bands' size difference, the band of *bz-s39.12* was cut off and sequenced later on. It finally turned out to be from *bz-R* line, unfortunately.

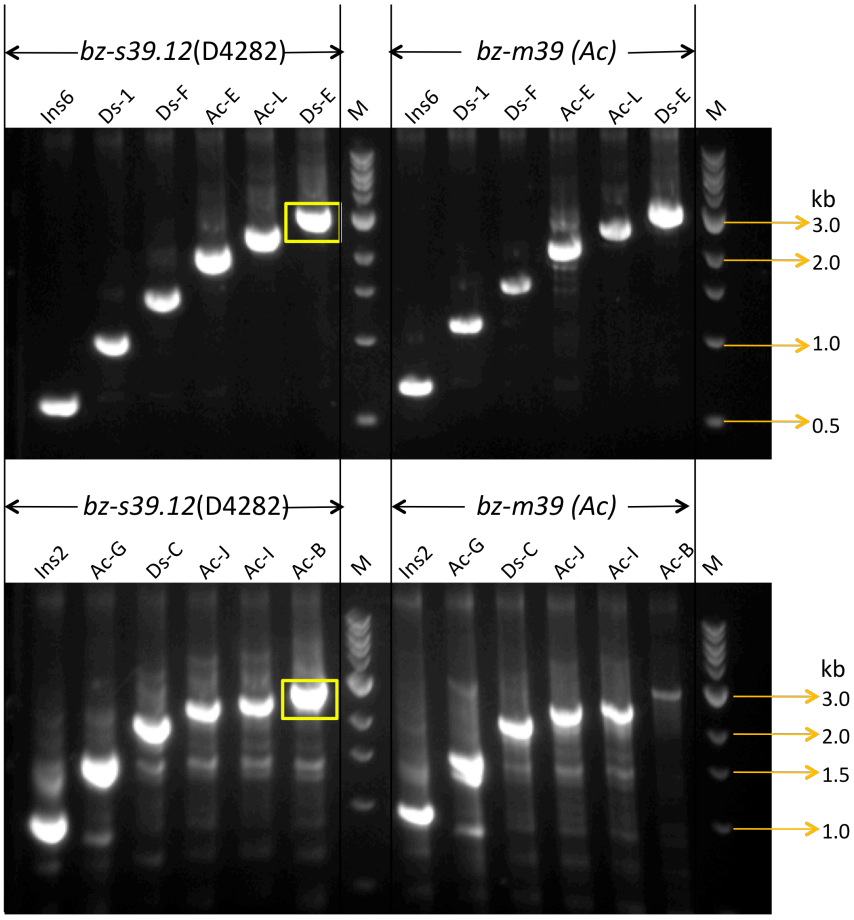


Figure 21. Long range PCR to amplify the 5' & 3' ends of *bz-s39.12* (D4282), respectively. The forward primer used in the top one is *bz*-y5; the reverse primer used in the bottom one is del2. The primers location in *Ac* and *bz* refer to **Figure 2**. *bz-m39(Ac)* was set as control.

3.3.2 Two ways to verify the hypermethylation status within *Ac* element

For those derivatives assumed to hypermethylated *Ac* elements, Southern blots will be carried out to verify that *Ac* was inactivated due to hypermethylation. Besides Southern blots, another method can be used to verify the hypermethylation status within the *Ac* element. I choose three selections W8042-1, W8042-2 and W8042-3 and digested the genomic DNA from them with three methylation sensitive restriction enzymes *Pvu*I, *Pvu*II and *Sac*II, respectively. Then PCR was performed with primers flanking the cutting sites, one is a bz primer, the other is *Ac* primer, to amplify fragments containing the cutting site. If the fragment can be amplify, it tells us that the enzyme digestions do not work and thus this fragment is hypermethylated; if not, it means the sequence containing cutting site is not hypermethylated. From **Figure 22** we may reach a conclusion that the inactivation of *Ac* element was due to hypermethylation. For instance, there are three *Sac*II cutting sites within the *Ac* element. Primers bz-y5 and Ds1, which flanks the *Sac*II cutting sites, were used to amplify this fragment containing *Sac*II site. From the agarose gel (**Fig. 22**), three seletions W8042-1, W8042-2 and W8042-3 all have very sharp bands with correct size, respectively, while the control *bz-m39(Ac)* does not with no doubt.

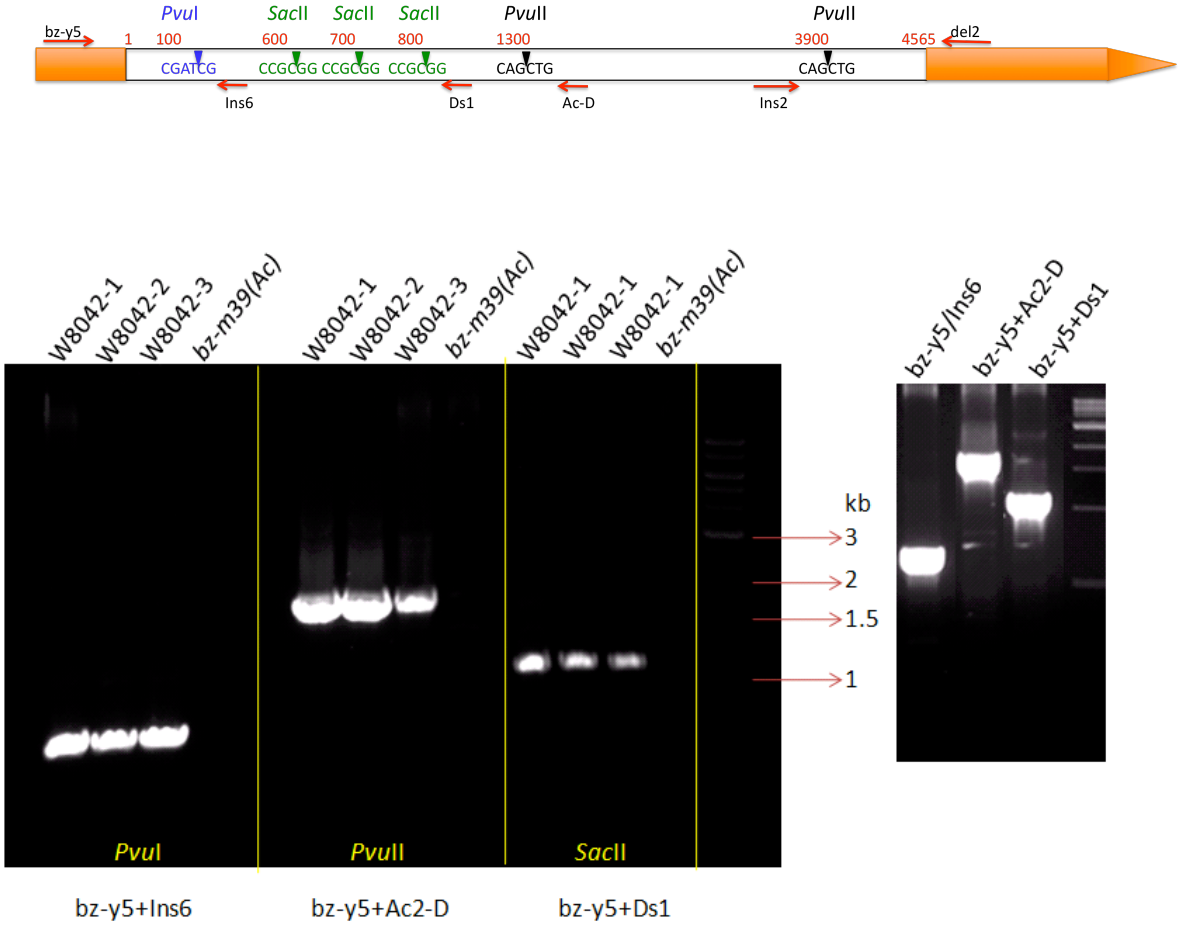


Figure 22. Schematic representation of positions of three methylation sensitive restriction enzymes PvuI PvuII SacII in *bz-m39(Ac)* (top). Primers bz-y5/Ins6, bz-y5/Ds1 and bz-y5/Ac2-D were used amplify fragments containing three methylation sensitive enzymes *Pvu*I, *Sac*II and *Pvu*II, respectively. Note that the white box represents *Ac* and orange box flanking the *Ac* represents *bz* gene. **PCR results show that all of the postulated selections have sharp bands with correct sizes while the control *bz-m39(Ac)* has not (bottom).** The agarose gel on the right side shows PCR results of amplification of undigested *bz-m39(Ac)* with the same primer sets, indicating the qualificaiton of genomic DNA from the control has no problem and no amplification of digested control was really due to enzyme digestion.

3.4 *bz* deletion adjacent to *Ac* 3' end

W8040-4

W8040-4 suffered a deletion of 183-bp adjacent *bz* sequence (**Figure 23**). It was suspected to be a single-ended *fAc* at the first beginning, since PCR with primer combination del2/Ds4 could not amplify any band (**Figure 24.B**). Then another reverse primer downstream of del2, *bz*-y2 (**Figure 2**) was used in combination with *bz*-y5 to amplify the fragment. Finally sequencing results indicate that there is no deletion or insertion in this *Ac*-derivative and 183-bp *bz* sequence adjacent to the 3' end of *Ac* element was deleted. Primer del2 is just located at the deletion junction region, and that is the reason why del2 did not work with Ds4 to amplify the 3' end of the *Ac*.

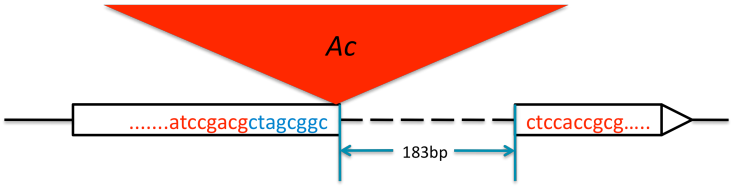


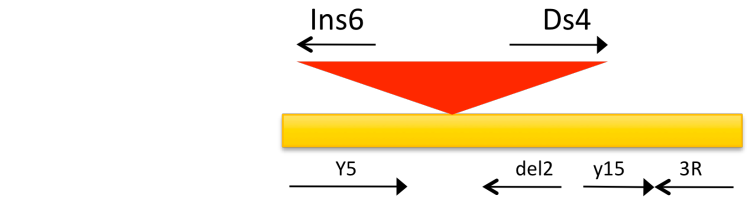
Figure 23. Schematic representation of W8040-4.

3.5 Contamination

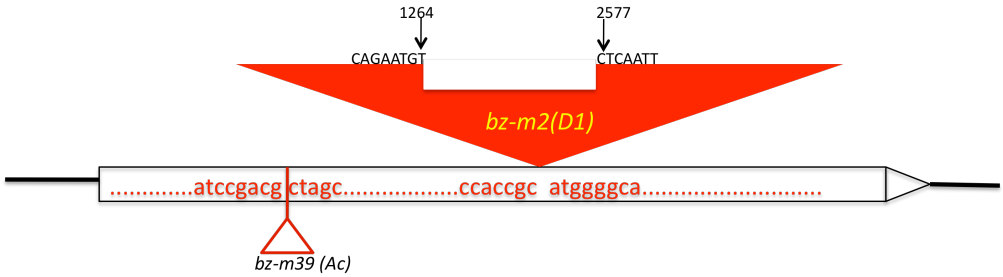
W8040-5, W8081-1, W8041-2

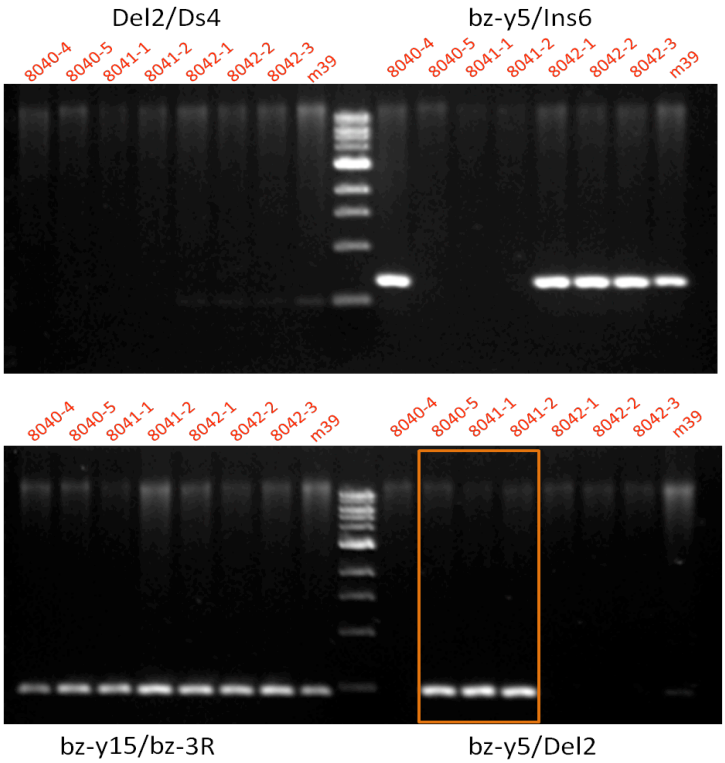
Several primer combinations were tried with selections W8040-5, W8041-1 and W8041-2 to see if the 5' and 3' ends were still present and if any transposon moved to other locations in *bz* locus (**Figure 24 (A) & Figure 25 (A)**). Long range PCR did not work for W8040-5, W8041-1, and W8041-2. So *bz*-y5/Ins6 and del2/Ds4 were used to identify if the selections' 5' and 3' ends were still present, respectively. Both 5' and 3' ends were not present in W8040-5, W8041-1, W8041-2, indicating that they had probably

transposed to other locations; W8042-1, W8042-2, W8042-3 were probably *Ds* elements since all of their 5' and 3' ends were intact.



(A)

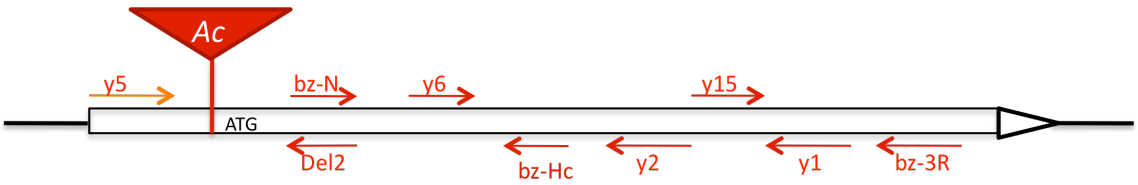




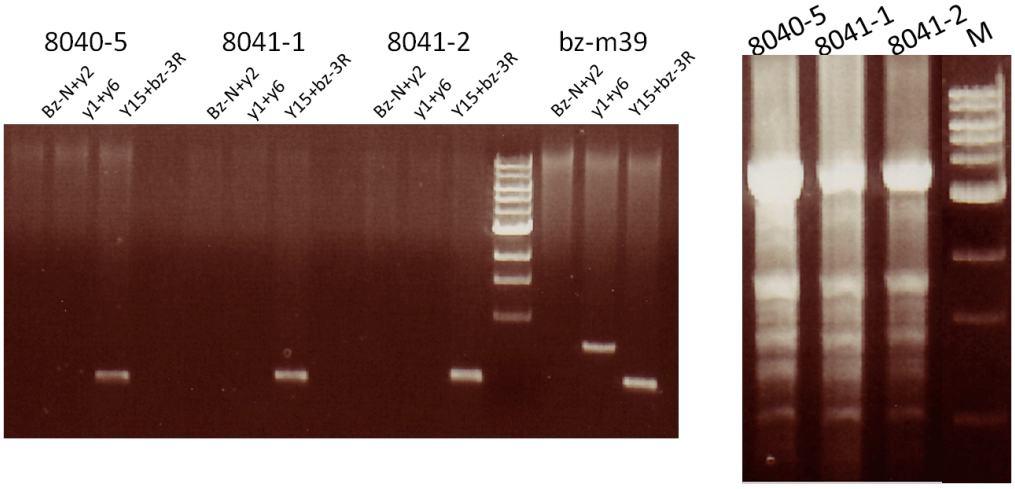
(B)

Figure 24. (A) a representation of PCR primers' relevant location in *bz* and *Ac*. (B) PCR was conducted to verify the 5' and 3' ends were intact in W8040-4, W8040-5, W8041-1, W8041-2, W8042-1, W8042-2, W8042-3.

These three selections finally turned out to be *bz-m2(D1)* pollen contaminants. This *bz-m* allele carries a *Ds* element inserted at position 755-762 in the second exon of *Bz-McC* (Dooner *et al.*, 1986). It has a 1313-bp internal deletion. First, PCR with normal Qiagen polymerase was performed with primers combination of *bz*-y5 and Del2. The bands were cut and sequenced and neither *Ds* element nor footprints were found. Then several other combinations of primers were used to identify where the *Ds* transposed to in the *bz* locus. PCR showed that the element probably was transposed to the region between y6 and y1 by comparison between the derivatives and the control *bz-m39(Ac)*. Finally the long range PCR with primers y6/y2 were carried out and sequencing results suggest that all these three selections are actually *bz-m2(D1)* contaminants. No genetic tests with these materials were carried out and this is why the pollen contaminant was not detected earlier.



(A)



(B)

Figure 25. (A) Relevant locations of PCR primers in bronze region. (B) PCR with common Qiagen Kit was carried out to identify where the transposon jumped into by several primers combinations (left). Then long range PCR was conducted to amplify the desired fragment in W8040-5, W8041-1 and W8041-2 (right).

IV. DISCUSSION

4.1 Diverse structures of *Ac* derivatives

Basically, 23 putative *Ac* mutant derivatives have been analyzed by cloning and sequencing methods. Within these 23 *Ac*-derivatives, there are 9 *Ds* elements with various sizes ranging from 2 kb to 4.5 kb. Seven *Ds* elements have no filler DNA in their deletion junction, and two have filler DNA duplicated from the *Ac* progenitor (**Figure 26; Figure 27**). Insterestingly five single-ended *fAc* elements were identified. Here the single-ended 5' *fAc* elements are different from the traditional 3' *fAc* element, since they did not transpose, still existing in their original progenitor's insertion site. They all lost adjacent *bz* sequences, and some even lost the intergenic region between *bz* and *stc1*. This is particularly obvious for *bz-s39.9* since a huge deletion (2447-bp) was identified starting from the 1st nucleotide following the transposable element to the intergenic region between *bz* and *stc1* gene.

The deletion junctions usually have two features: microhomology and filler DNA (Yan *et al.*, 1999). Yan *et al.* identified two *Ds* elements with filler DNA as long as 20 bp in the internal deletion junction; however, it has been demonstrated in our research that the filler DNA sequence may not be only limited to 20 bp. The filler can be as long as 200 bp, duplicated from some part of *Ac* itself (*bz-s39.5*, *bz-s39.7*). In our research, extra DNA was not only identified in the internal deletion junctions of *Ds* elements (*bz-s39.5*, *bz-s39.7*), but was also found in some of the *fractured Ac* element with defective 5' end (*bz-s39.1, bz-s39.4*). Conrad *et al* (2007) reported that seven *Ds* elements contain filler DNA ranging from 1 to 64 bp in the deletion junctions within the 15 derivatives they

studied; in other word, almost half of *Ds* element contain fillers. However, in our case we only found two *Ds* element have fillers in the deletion junctions within the nine *Ds* elements we have studied, which means less than half of *Ds* elements contain fillers in the deletion junctions. However, this number may not indicate anything considering the low number of total *Ds* elements we have researched. Hopefully more derivatives will be generated in the future so as to analyze whether there is any locus preference to produce fillers in deletion junctions when DSB repair takes place. Conrad *et al* (2007) summarized that fillers always originate from sequence in close poximity to the deletion. From our study on *bz-s39.5*, we may enlarge the origination source of fillers to some greater extent, that is, filler DNA can also originate from the 5' end of *Ac*. The filler is a fairly large duplication (201 bp), not in close proximity to the deletion junction. Interestingly, Conrad et al (2007) also reported a derivative *ps1-m8::Ds (Ds2)* that contains filler DNA originated from the flanking sequence adjacent to the site of *Ac* insertion at *ps1* and is missing the *Ds* 5' end sequence. This derivative resembles *bz-s39.1* and *bz-s39.4*. However, it was a bit bizzare that they still treated *ps1-m8::Ds (Ds2)* as a *Ds* element since it does not have 5' end any more, and is, therefore, a 3' *fAc* element.

Our research results once again show that footprints in maize typically have much smaller deletions or additions, usually only one or two bases; in contrast, most *Ac/Ds* footprints in yeast usually show a much larger deletions or additions of palindromic sequences (1~20 bp) (Rinehart *et al,* 1997; Scott *et al,* 1996;Weil & Kunze, 2000).

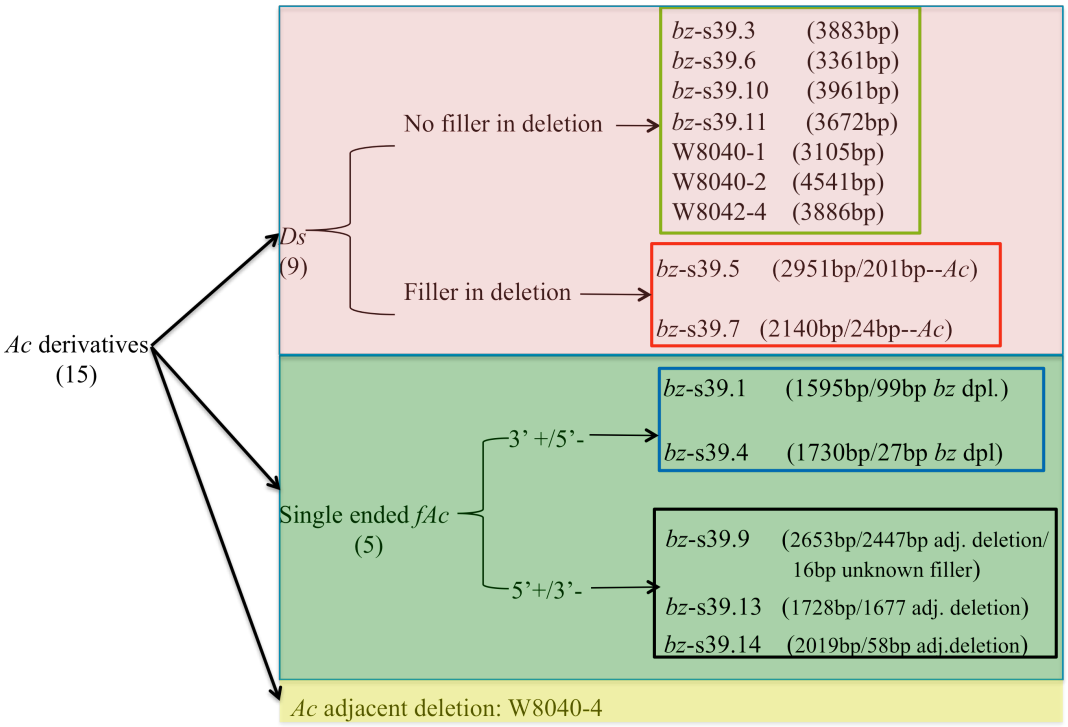


Figure 26. Summary of 15 *Ac* derivatives that have been analyzed by sequencing. Basically the derivatives can be divied by two categories: nine *Ds* elements (shown in blue rectangle) and five single-ended *fAc* (green rectangle). Nine *Ds* elements consist of seven *Ds* elements without filler DNA in internal deletion and two with filler. Five single-ended *fAc*s consist of two 3' end *fAc*s and three 5' end *fAc*s. 3' end *fAc* does not lead to adjacent genomic sequence deletion. However, 5' end *fAc* does. Each seletion was given a new *bz-s* derivative name if we still have seeds. The number shown in brackts represents the element' size and numbers following the slash represents the size of the filler or duplication.

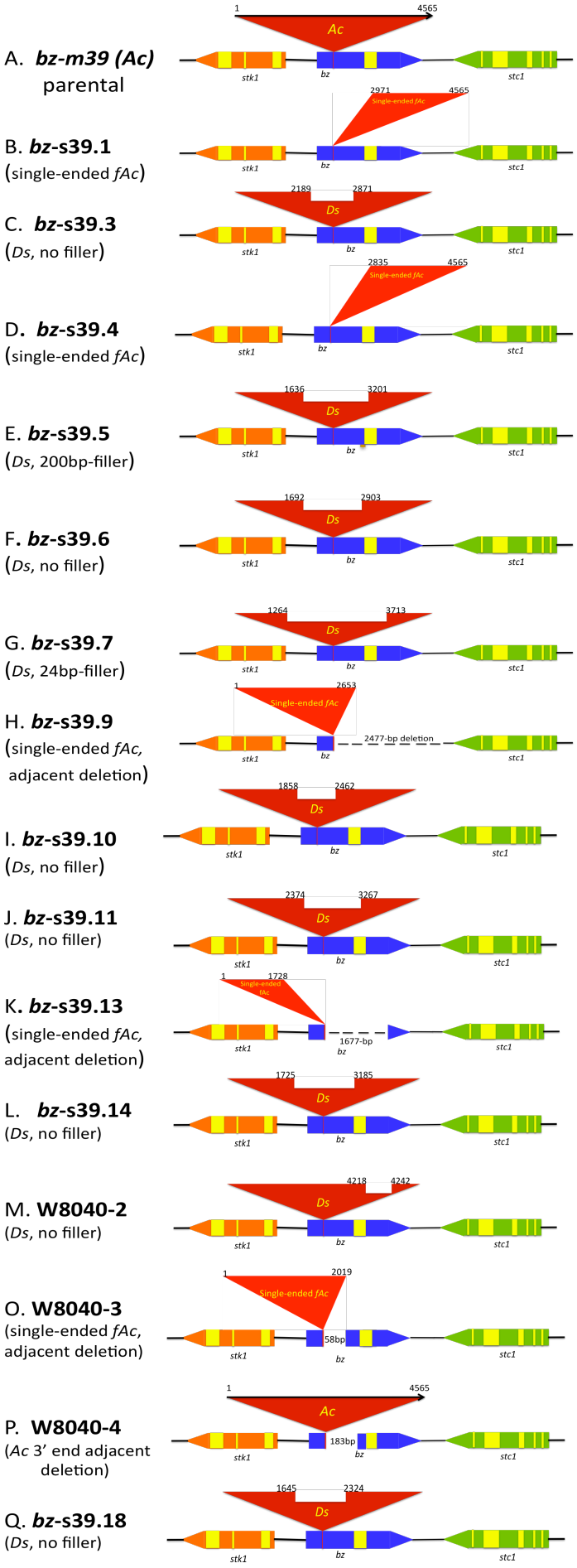


Figure27. Structure of the *bz-m39(Ac)* chromosome and some *bz*-s derivatives. The *Ac* is shown as a red triangle. The numbers identify the last retained nucleotide positions of the deletion junctions relative to the sequence of the 4565-bp *Ac*. Genes are represented as pentagons pointing in the direction of transcription and are not drawn to scale: introns are in yellow; exons of *stk1, bz* and *stc1* genes are in orange, blue and green, respectively.

(A). *bz-m39(Ac)* parent;

(B). *bz-s* derivative *39.1*, a 1595-bp 3' single-ended fAc plus a 99-bp duplicated sequence from *bz* adjacent to the *Ac* 3' end.

(C). *bz-s* derivative *39.3*, a 3883-bp *Ds* element.

(D). *bz-s* derivative *39.4*, a 1730-bp 3' single-ended *fAc* plus a 27-bp duplicated *bz* sequence adjacent to the *Ac* 3' end.

(E). *bz-s* derivative *39.5*, a 2951-bp *Ds* element with 201-bp filler duplicated from *Ac* in the deletion junction.

(F). *bz-s* derivative *39.6*, a 3361-bp *Ds* element.

(G). *bz-s* derivative *39.7*, a 2140-bp *Ds* element with 24-bp filler duplicated from *Ac* in the deletion junction.

(H). *bz-s* derivative *39.9*, a 2653-bp 5' single-ended *fAc* plus a 16-bp filler from an unknown source; a 2477-bp *bz* and intergenic sequence between *bz* and *stc1* was deleted together with the *Ac* 3' end.

(I). *bz-s* derivative *39.10*, a 3961-bp *Ds* element.

(J). *bz-s* derivative *39.11*, a 3672-bp *Ds* element.

(K).*bz-s* derivative *39.13*, a 1728-bp 5' single-ended *fAc* plus a 1677-bp deletion of adjacent *bz* sequence.

(L). *bz-s* derivative *39.14*, a 3105-bp *Ds* element.

(M). *bz-s* derivative W8040-2, a 4541-bp *Ds* element with three nucleotides at 4216, 4217 and 4243-bp positions changed from G, T and A to A, A and G, respectively.

(O). *bz-s* derivative W8040-3, a 5' 2019-bp single-ended *fAc* plus a 58-bp deletion of adjacent *bz* sequence.

(P). *bz-s* derivative W8040-4, a 183-bp deletion of *bz* sequences adjacent to the *Ac* 3' end.

(Q). *bz-s* derivative *39.18*, a 3886-bp *Ds* element.

4.2 Microhomology: evidence for repair synthesis at the site of *Ac* excision

Microhomologies are short repeat sequences found in the single-stranded overhangs on the ends of double-stranded breaks. When no homologous template is present, NHEJ is the mechanism that utilizes microhomologies to repair the DSB. Except for three contaminant selections (W8040-5, W8041-1, W8041-2) and five postulated hypermethylated *Ac* derivatives (*bz-s*39.2, *bz-s*39.12, W8042-1, W8042-2 and W8042-3), 15 derivatives have been analyzed by searching for microhomology in their deletion junctions. Among the 15 derivatives, microhomologies have been found in most junctions (10 in total), ranging from 2 to 5 bp (**Figure 28**). These results are consistent with previous studies in maize and *Saccharomyces cerevisiae* (Yan *et al.*, 1999; Yu *et al.,* 2004) It was not possible to find evidence for microhomology in *bz-s39.9* since the source of the filler DNA cannot be identified. Conrad *et al* (2007) indicate that repeat sequences at deletion junctions are not prerequisite for formation of state II *Ds* element in maize based on their study that at least 5/15 of the deletion do not have direct repeats flanking filler DNA or deletion junctions. However, we have found that 10/15 have direct repeats flanking fillers or deletion junctions, which means that only 5/15 do not. Anyway, it is still early to make the statement that microhomologies are not the prerequisite for state II *Ds* element formation since the total number of *Ds* element we currently have identified is statistically too low.

Generally the microhomology that has been found in the 10 derivatives can be divided into two categories: *Ds*-microhomology and single-ended *fAc*-microhomology. The *Ds*-microhomology group consists of 6 members (D6232, D6234, D6235, D6236, D5078, W8040-1). In this group, the microhomology is based on the internal deletion junction of

Ac. In the second group, which consists of 4 members (D6242, D6233, D5080, W8040-3), all the microhomology is based on the sequence of *bronze* and the single ended fractured element. For instance, D6242 has 99-bp duplicated *bz* sequence between the footprint (CTAGCGG) and the single ended *fAc*. By BLAST search it has been demonstrated that the 99-bp was duplicated from the bronze sequence right downstream of the fractured element. The first four nucleotides (TTCA) following the 99-bp bronze sequence was found homologous with the first four nucleotides of the fractured element (TTCA).

A model modified from Yan *et al.* (1999) is proposed to explain how internal deletion occurs in *Ac* (**Figure 29**). Figure 28 (A) shows how internal deletion without filler DNA forms in *Ds*. A double strand break is generated when *Ac* excises from the lower sister chromatid. The DSB is repaired with the upper sister chromatid as the template. During repair, slip mispairing takes place between the two direct repeats (black rectangles) leading to the loss of one repeat and the sequence between the two repeats deleted from the newly synthesized strand. **Figure 29 (B)** shows how internal deletion with filler DNA forms in *Ds*. Different from (A), a second slip-mispairing occurs between different direct repeats (purple rectangles) leading to an internally deleted *Ds* element with filler DNA (yellow rectangles) instead of the deleted sequence (green rectangles). This model can still be used to explain how deletion occurs in *bz-s39.13*, which has no filler DNA in the deletion junction. The only difference between deletion of *bz-s39.13* and other regular *Ds* internal deletions is that, in *bz-s39.13*, slip-mispairing occurs in the *Ac* adjacent *bz* sequence, rather than *Ac* internal sequence as in other *Ds* element. However it is not possible to explain *bz-s39.9* because of the filler DNA of unknown-source. For *bz-s39.1*

and *bz-s39.4*, a possible model was proposed to explain how *Ac* 3' end adjacent *bz* sequence duplicates and how 5' end of *Ac* is missing (**Figure 30**). F and G are joined together and adjacent sequence GE is replicated. E combines with C and *Ac* 3' end (triangle DCB) thus is replicated accompanied with the deletion of *Ac* 5' end (triangle DAC). After that sequence HK, which is the *Ac* 3' end adjacent sequence is replicated. Consequently the 3' end single-ended *fAc* will have a duplicated *Ac*-3'-end-adjacent host sequence and a fractured *Ac* element.

	derivatives	deletion junctions	
Ds	*bz-s39.3*	CCTTGTCTTTGGATAAT CCTTGTCT----------	AGAAATATTGGAAAGTGT ----------TTGGAAAGTGT
Ds	*bz-s39.5*	CCGTTTTTAATCGGGATGATCCCGT CCGTTTTTAATCGGGA----------	CCCCTTTTGAAGCATAGT ----------TGAAGCATAGT
Ds	*bz-s39.6*	TTTAGTTAAAGGTCA TTTAGTTA----------	CTAGCTGTAGCATGCTT ----------GCATGCTT
Ds	*bz-s39.7*	CAGAATGTACGTGCACGT CAGAATGT----------	CAGCCGTGTTATACTCCCT ----------TATACTCCCT
Ds	*bz-s39.11*	TCAAAGCGATTGTTCTTG TCAAAGCGAT----------	AAGGGATGTGCTAGC ----------GTGCTAGC
Ds	*bz_s39.14*	GTGAAAAGGATCATGGCAAA GTGAAAAGGAT----------	GGATAAATATATGTCTGAAC ----------ATGTCTGAAC
Single-ended *fAc*	*bz-s39.1*	GGTCGCCTTCCCGTTCAGCTCCC GGTCGCCTTCCCG----------	CATGGTGATTCATACAAAGTT ----------TTCATACAAAGTT
Single-ended *fAc*	*bz-s39.4*	GAACAGCCTGAGCGCGCGAAGATG GAACAGCCTGAGCGCGC----------	CATTAGGAGAATGGCCGT ----------GAATGGCCGT
Single-ended *fAc*	*bz-s39.13*	AAGGATCATGGCAAAGA AAGGATCATG----------	ATACATGAGTTGCAGCA ----------AGTTGCAGCA
Single-ended *fAc*	W8040-3	CATCTTGTCAAAATAAGTCA CATCTTGTCAAAATA---------	CCGACGGCGAGTCCTCCCCG ----------AGTCCTCCCCG

Figure 28. Sequences of the deletion junction in the *bz-s* derivatives from *bz-m39(Ac)*. The sequence of the 5' and 3' ends of the deletion and the corresponding sequence in *Ac* are shown for each derivatives. Microhomology is shown in red. The *Ac* sequence is shown in the same orientation as its transposase transcript and the numbered carats refer to positions in that sequence (1-4565). The derivatives sequences are in the same orientation as *Ac*.

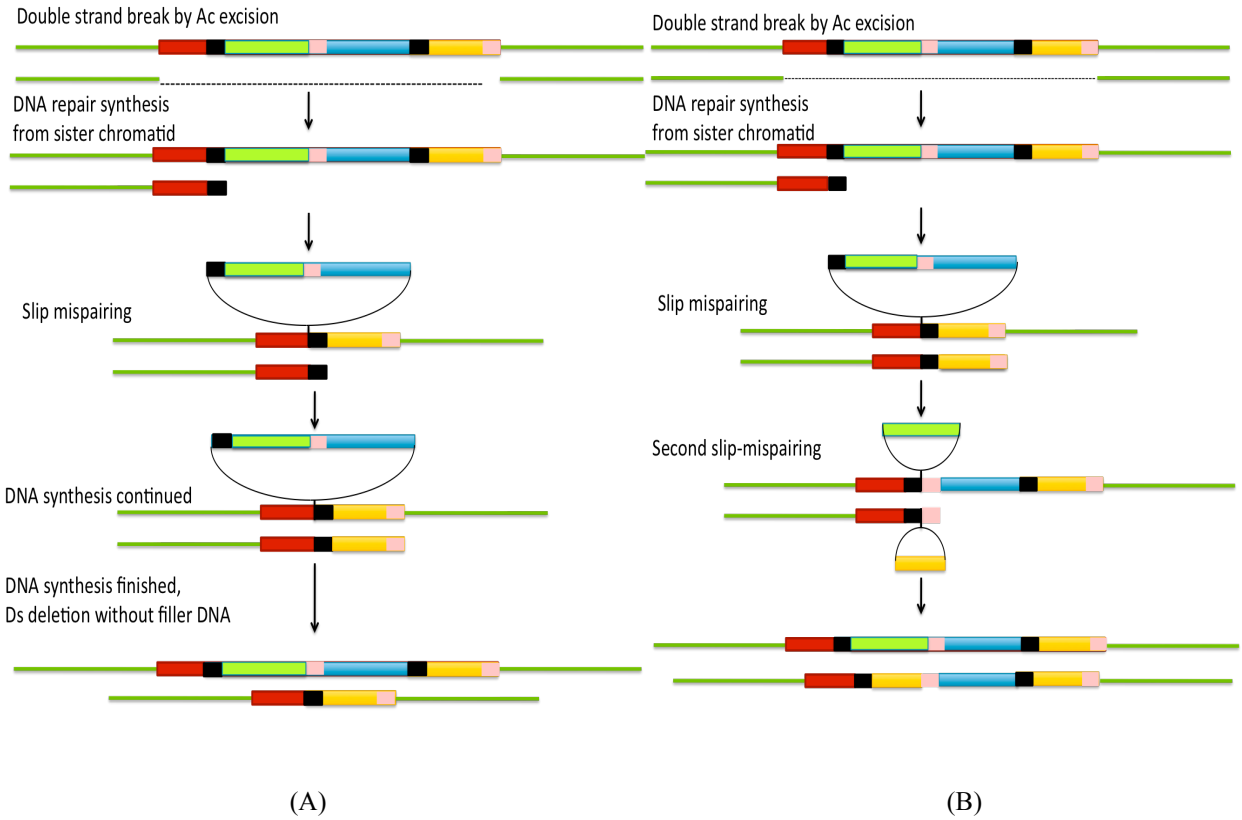


Figure 29. Model for the origin of internal deletion in *Ds* from *Ac*. (A) shows how internal deletion without filler DNA forms in *Ds*. A double strand break is produced when *Ac* excised from the lower sister chromatid. The DSB is repaired using the upper sister chromatid as the template. During repair, slip mispairing takes place between the two direct repeats (black rectangles) leading to deletion of one repeat and the sequence between the two repeats from the newly synthesized strand. (B) shows how internal deletion with filler DNA forms in *Ds*. Different from (A), a second slip-mispairing occurs between different direct repeats (pink rectangles) leading to the internally deleted *Ds* element with filler DNA

(yellow rectangles) instead of the deleted sequence (green rectangle). The black and pink rectangles represent two different direct repeats, respectively.

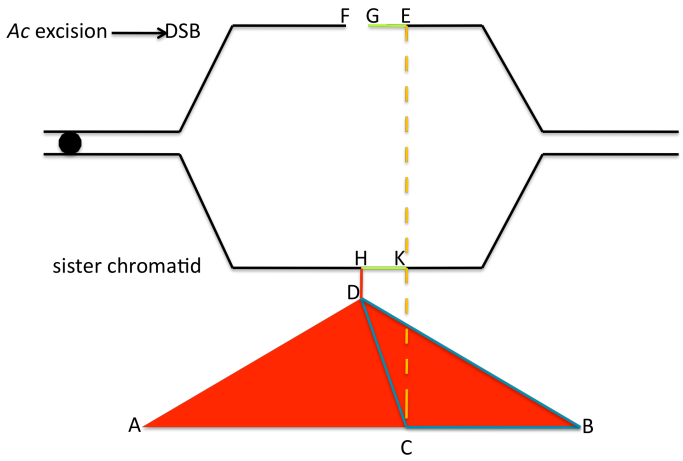


Figure 30. Model for origination of 5' single-ended *fAc* element. F and G are joined together and adjacent sequence GE is replicated. E combines with C and *Ac* 3' end (triangle DCB) thus is replicated accompanied with the deletion of *Ac* 5' end (triangle DAC). After that sequence HK, which is the *Ac* 3' end adjacent sequence is replicated. Consequently the 3' end single-ended *fAc* will have a duplicated *Ac* 3' end adjacent host sequence and a fractured *Ac* element.

4.3 Adjacent deletion led by *Ac* abortive transposition

We found that, in W8040-4, *Ac* is retained in the original insertion site but 183-bp *bz* adjacent to the *Ac* 3' end was deleted. Because of the deletion of *bz* sequence, W8040-4 has a stable *bronze* phenotype even though *Ac* remains intact. It has been widely found that adjacent deletions take place in prokaryotic and eukaryotic transposons. Dooner et al (1988) found that a derivative *bz-s:2114 (Ac)* retains *Ac* at the original insertion site but has lost a 789 bp upstream *bz* sequence adjacent to insertion. Adjacent deletion can be explained as abortive transposition reactions involving only one element end (**Figure 31**).

Ac 3' end (a) can either join to receptor site d to form new junctions a-d and b-c, thus leading to *Ac* 3' end adjacent deletion, or can join to receptor site c and b ligate with d, to form new junctions a-c and b-d, thus leading to inversion of *Ac* 3' end adjacent sequence.

Besides the adjacent deletion caused by abortive transposition of single TE, there are some other types of adjacent deletion caused by TE pairs, and the deletion of adjacent host sequences may not be limited to *Ac* 3' end (Huang and Dooner, 2008). The frequency of adjacent deletions is 10 fold higher when two TEs are close to each other in the chromosome than when a single TE is present.

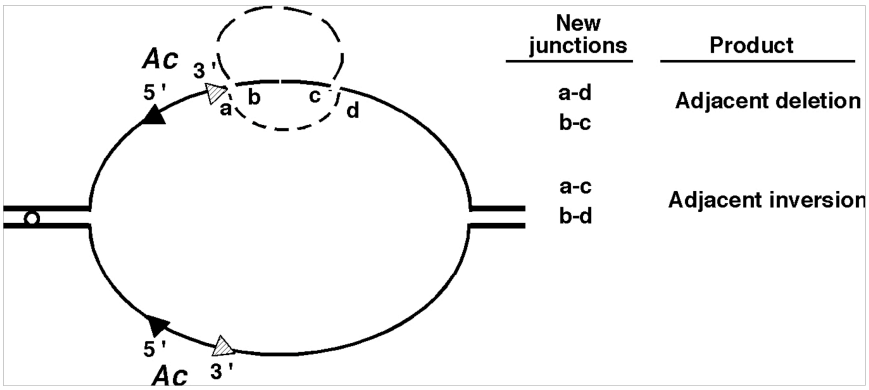


Figure 31. Shematic representation of *Ac* 3' end adjacent deletion. *Ac* 3' end (a) can either join to receptor site d to form new junctions a-d and b-c, thus leading to *Ac* 3' end adjacent deletion, or can join to receptor site c and b ligate with d, to form new junctions a-c and b-d, thus leading to inversion of *Ac* 3' end adjacent sequence. This figure is generously dedicated by H.K.Dooner.

4.4 Recombination experiment:

The long term goal of this project is to determine if two transposable elements in different homologues (in trans) can recombine at meiosis. To test this, we could

reconstitute the 4565-bp *Ac* element at *bz* by recombination between two *Ds* or single ended *fAc* elements with nonoverlapping deletions, since we have several different *Ac* derivatives from the same *bz-m*(*Ac*) allele. Besides *Ds* elements, several single-ended *fAc* elements can still be suitable for the reconstruction purpose. It should be mentioned that several derivatives seeds were lost in the greenhouse this past winter, so some derivatives like W8040-2, W8040-3, W8040-4, W8042-2 and W8042-4 are not considered here for recombination purposes. It is a shame, that we lost W8040-4 and W8040-2, since W8040-4 has an intact *Ac* element accompanied by a 183-bp adjacent *bz* sequence deletion and it would be able to recombine with any other *Ds* element or *fAc* element, and W8040-2 contains only 24-bp deletion in *Ac* and it would also be able to recombine with many other *Ds* or *fAc* elements.

It should be noted that in some derivatives, *bronze* sequences adjacent to the 3' end of the transposable elements were lost; anyhow it can still be considered for recombination with other derivatives since an intact *Ac* element can still be reconstructed by recombination eventually as long as other derivatives possess intact *bronze* sequence. In addition *bz-s39.1* and *bz-s39.4*, which have some filler DNA in the *Ac* excision site, can be considered for recombination purposes as well, because those filler will not impact the *bronze* gene function by recombination with other derivatives. Considering that we lost several derivatives in the greenhouse, we have very few derivatives left that are suitable for reconstructing the intact *Ac* element. Finally, two derivatives, *bz-s39.9* and *bz-s39.10*, are considered as appropriate derivatives to reconstruct the 4.6-kb *Ac* (**Figure 32**). First the deletions in *bz-s39.9* and *bz-s39.10* do not overlap; second, the *bz-s*39.9 2477-bp *fAc*

adjacent deletion in *bz* does not matter since it can recombine with *bz-s39.10* to give the full *bz* sequence.

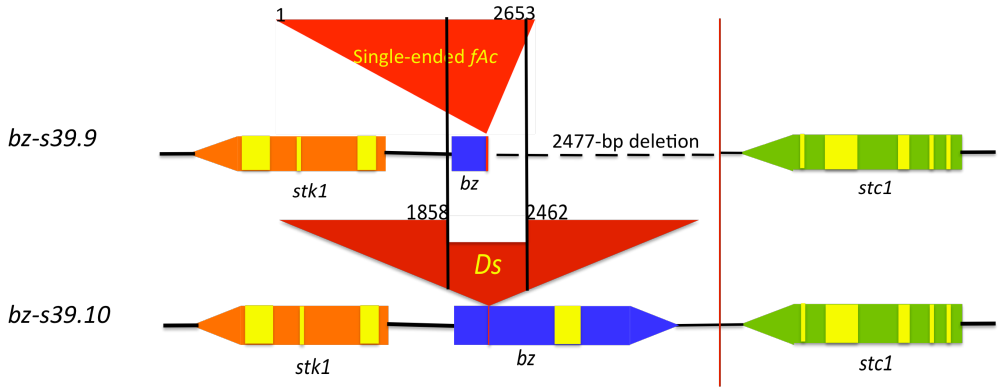


Figure 32. *bz-s39.9* and *bz-s39.10* are two possible candidate derivatives to reconstruct 4.6-kb *Ac* element by recombination experiment. The deletions occurred in bz-s39.9 and bz-s39.10 are non-overlapped thus they can be candidates for rescontruction experiment.

In conclusion, all of the 23 *Ac* derivatives have been characterized by PCR amplification and sequencing. The results indicate that not all the deletion occurred in the internal part of *Ac* element. Five derivatives suffered either 5' or 3' end deletions. Three of them lost the 3' end and adjacent *bz* sequences of various length, ranging from 60 to 2500 bp. Microhomologies were found in most of the derivatives, indicating that microhomologies play an important role in the repair of DSBs. Unfortunately, we lost several suitable derivatives for our next-step recombination experiment. We still have two derivatives that may serve as candidates for future recombination experiments, and the isolation of additional derivatives is being planned for this summer.

V. LITERATURE

Behrens, U., N. Fedoroff, A. Laird, M. Muller-Neumann, P. Starlinger (1984). Cloning of the *Zea mays* controlling element *Ac* from the *wx-m7* allele. *Mol. Gen. Genet.* 194: 346–347

Britt, A. B., and V. Walbot. (1991). Germinal and somatic products of *Mu* excision from the *Bronze-1* gene of *Zea mays*. *Mol. Gen. Genet.* **227:**267-276.

Brutnell, T. P., and S. L. Dellaporta. (1994) Somatic inactivation and reactivation of *Ac* associated with changes in cytosine methylation and transposase expression. *Genetics* 138: 213–225.

Brutnell, T. P., B. P. May and S. L. Dellaporta. (1997) The *Ac-st2* element of maize exhibits a positive dosage effect and epigenetic regulation. *Genetics* 147: 823–834.

Chomet, P. S., S. Wessler and S. L. Dellaporta. (1987). Inactivation of the maize transposable element *Activator* (*Ac*) is associated with its DNA modification. *EMBO J.* 6: 295–302.

Conrad, L.J., Bai, L., Ahern, K., Dusinberre, K., Kane, D.P., and Brutnell, T.P. (2007). State II Dissociation (*Ds*) Element Formation Following *Activator* (*Ac*) Excision in Maize. *Genetics* 177.

Dooner, H.K., and Belachew, A. (1989). Transposition pattern of the maize element *Ac* from the *bz*-m2(*Ac*) allele. *Genetics* **122**, 447-457.

Dooner H.K., He L. (2008) Maize Genome Structure Variation: Interplay between Retrotransposon Polymorphisms and Genic Recombination. *The Plant cell.* **20**(2):249-58.

Dooner, H.K., and Martinez-Ferez, I.M. (1997). Recombination occurs uniformly within the bronze locus, a meiotic recombination hotspot in the maize genome. *Plant Cell* **9**, 1633-1646.

Dooner, H.K., J. English, E. Ralston and E.Weck. (1986). A single genetic unit specifies two transposition functions in the maize element *Activator*. *Science* **234**:210-211.

Doring, H.P., Tillmann, E., and Starlinger, P. (1984). DNA sequence of the maize transposable element *Dissociation*. *Nature* **307**, 127-130.

Doring, H. P., M. Freeling, S. Hake, M. A. Johns, R. Kunze. (1984). A *Ds* mutation of the *Adh1* gene in *Zea mays* L. *Mol. Gen. Genet*. 193: 199–204.

Doring, H. P., B. Nelson-Salz, R. Garber and E. Tillmann. (1989). *Double Ds* elements are involved in specific chromosome breakage. *Mol. Gen. Genet*. 219: 299–305.

Doring, H. P., I. Pahl and M. Durany. (1990). Chromosomal rearrangements caused by the aberrant transposition of *double Ds* elements are formed by *Ds* and adjacent non-*Ds* sequences. *Mol. Gen. Genet.* 224: 40–48.

Doseff, A., R. Martienssen, and V. Sundaresan. 1991. Somatic excision of the *Mu*l transposable element of maize. *Nucleic Acids Res.* **19:**579-584.

English, J., Harrison K., and Jones, J.D.G. (1993). A genetic analysis of DNA sequence requirements for *Dissociation* state I activity in tobacco. *The Plant Cell* **5**, 501-514.

Fedoroff, N., S. Wessler and M. Shure. (1983). Isolation of the transposable maize controlling elements *Ac* and *Ds. Cell* 35: 235–242

Fu, H., Zheng, Z., and Dooner, H.K. (2002). Recombination rates between adjacent genic and retrotransposon regions differ by two orders of magnitude. *Proc. Natl. Acad. Sci. USA* **99,** 1082-1087.

Fusswinkel, H., S. Schein, U. Courage, P. Starlinger and R. Kunze. (1991). Detection and abundance of mRNA and protein encoded by transposable element *Activator* (*Ac*) in maize. *Mol. Gen. Genet.* 225: 186–192.

He, L., and Dooner, H.K. (2007). Recombination in a 100-kb genic interval containing Helitrons and retrotransposons. In 49th Annual Maize Genet. Conf. Abstracts (St. Charles, IL), pp. 99.

Huang, J., and Dooner, H.K. (2008). Chromosome restructuring by interacting transposon pairs. *The Plant Cell* **20,** in press.

Kunze, R., H. Saedler and W.-E. Lonnig. (1997). Plant transposable elements, pp. 332–469 in *Advances in Botanical Research*, edited by J. A. CALLOW. *Academic Press*, London.

Kunze, R., P. Starlinger and D. Schwartz (1988). DNA methylation of the maize transposable element *Ac* interferes with its transcription. *Mol. Gen. Genet.* 214: 325–327.

Kunze, R., and C. F. Weil. (2002). The *hAT* and *CACTA* superfamily of plant transposons, pp. 565–610 in Mobile DNA, edited by N. L. CRAIG. ASM, Washington, DC.

McClintock, B. (1949). Mutable loci in maize. *Carnegie Inst. Wash. Yearbook* **48:** 142-154.

McClintock, B. (1955). Controlled mutation in maize. *Carnegie Inst. Wash. Yearbook* **54**:245-255.

McClintock, B. (1956a). Controlling elements and the gene. *Cold spring Harbor Symp. Quant. Biol.* **21**:197-216.

McClintock, B. (1956b). Mutation in maize. *Carnegie Inst. Wash. Yearbook* **55**: 323-332.

McClintock, B. (1963). Further studies of gene control systems in maize. *Carnegie Inst. Wash. Yearbook.* **62**, 486-493.

Muller-Neumann, M., J. I. Yoder and P. Starlinger. (1984). The DNA sequence of the transposable element *Ac* of *Zea mays* L. *Mol. Gen. Genet.* 198: 9–24.

Pohlman, R. F., N. V. Fedoroff and J. Messing. (1984). The nucleotide sequence of the maize controlling element *Activator*. *Cell* 37: 635–643.

Ralston, E. J., J. English, and H. K. Dooner. (1987). Stability of deletion, insertion and point mutations at the *bronze* locus in maize. *Theor. Appl. Genet.* **74**:471-475.

Ralston, E. J., J. English, and H. K. Dooner. (1989). Chromosome-breaking structure in maize involving a fractured *Ac* element. *Proc. Natl. Acad. Sci. USA* **86**:9451-9455.

Rhoades, M. M. (1952). The effect of the *bronze* locus on anthocyanin formation in maize. *Am. Nat.* **86**:105-108.

Rinehart, T. A., C. Dean, and C. F. Weil. (1997). Comparative analysis of non-random DNA repair following *Ac* transposon excision in maize and *Arabidopsis*. *Plant J.* **12:**1419-1427.

Schwartz, D., and E. Dennis. (1986). Transposase activity of the *Ac* controlling element in maize is regulated by its degree of methylation. *Mol. Gen. Genet.* 205: 476–482.

Scott, L., D. LaFoe, and C. F. Weil. (1996). Adjacent sequences influence DNA repair accompanying transposon excision in maize. *Genetics* **142:**237-246.

Weck, E., U. Courage, H. P. Doring, N. Fedoroff and P. Starlinger. (1984). Analysis of *sh-m6233*, a mutation induced by the transposable element *Ds* in the sucrose synthase gene of *Zea mays*. EMBO J. 3: 1713–1716.

Weil, C.F. and R. Kunze. (2000). Transposition of maize *Ac/Ds* transposable elements in the yeast *Saccharomyces cerevisiae. Nature Genet.* **26:** 187-190

Weil, C.F. and Wessler, S.R. (1993). Molecular evidence that chromosome breakage by *Ds* elements is caused by aberrant transposition. *The Plant Cell* **5**, 515-522.

Yan, X., Martínez-Férez, I.M., Kavchok, S., and Dooner, H.K., (1999). Origination of *Ds* Elements From *Ac* Elements in Maize: Evidence for Rare Repair Synthesis at the Site of *Ac* Excision. *Genetics.* **152**: 1733-1740.

Yu, J., Marshall, K., Yamaguchi. M., Haber, J.E., and Weil, C.F., (2004). Microhomology- dependent end joining by host factors in *Saccharomyces cerevisiae. Molecular & Cellular Biology*.**24.3**.1351-1364

Zhang, J. and Peterson, T. (1999). Genome Rearrangements by Nonlinear Transposons in Maize. Genetics. 153, 1403-1410.

Zhang, J. and Peterson, T. (2004). Transposition of Reversed *Ac* Element Ends Generates Chromosome Rearrangements in Maize. *Genetics*. 167, 1929-1937.