5-28-2019

Spontaneous mutations in maize pollen are frequent in some lines and arise mainly from retrotranspositions and deletions

Hugo K. Dooner  
*Rutgers - The State University of New Jersey, New Brunswick*

Qinghua Wang  
*Rutgers - The State University of New Jersey, New Brunswick*

Jun T. Huang  
*Global R and D*

Yubin Li  
*Rutgers - The State University of New Jersey, New Brunswick*

Limei He  
*Rutgers - The State University of New Jersey, New Brunswick*

*See next page for additional authors*

Follow this and additional works at: [https://digitalcommons.montclair.edu/biology-facpubs](https://digitalcommons.montclair.edu/biology-facpubs)

Part of the Biology Commons

**MSU Digital Commons Citation**

Dooner, Hugo K.; Wang, Qinghua; Huang, Jun T.; Li, Yubin; He, Limei; Xiong, Wenwei; and Du, Chunguang, "Spontaneous mutations in maize pollen are frequent in some lines and arise mainly from retrotranspositions and deletions" (2019). *Department of Biology Faculty Scholarship and Creative Works*. 399, [https://digitalcommons.montclair.edu/biology-facpubs/399](https://digitalcommons.montclair.edu/biology-facpubs/399)

This Article is brought to you for free and open access by the Department of Biology at Montclair State University Digital Commons. It has been accepted for inclusion in Department of Biology Faculty Scholarship and Creative Works by an authorized administrator of Montclair State University Digital Commons. For more information, please contact digitalcommons@montclair.edu.
Spontaneous mutations in maize pollen are frequent in some lines and arise mainly from retrotranspositions and deletions

Hugo K. Dooner,a,b,1 Qinghua Wang,a,2,3 Jun T. Huanga,2,4 Yubin Li,a,5 Limei He,a Wenwei Xionga and Chunguang Duc

aWaksman Institute, Rutgers University, Piscataway, NJ 08854; bDepartment of Plant Biology, Rutgers University, New Brunswick, NJ 08901; and cDepartment of Biology and Molecular Biology, Montclair State University, Montclair, NJ 07043

Contributed by Hugo K. Dooner, March 20, 2019 (sent for review March 6, 2019; reviewed by L. Curtis Hannah and Clifford Well)

While studying spontaneous mutations at the maize bronze (bz) locus, we made the unexpected discovery that specific low-copy number retrotransposons are mobile in the pollen of some maize lines, but not of others. We conducted large-scale genetic experiments to isolate new bz mutations from several bz stocks and recovered spontaneous stable mutations only in the pollen parent in reciprocal crosses. Most of the new stable bz mutations resulted from either insertions of low-copy number long terminal repeat (LTR) retrotransposons or deletions, the same two classes of mutations that predominated in a collection of spontaneous wx mutations [Wessler S (1997) The Mutants of Maize, pp 385–386]. Similar mutations were recovered at the closely linked sh locus. These events occurred with a frequency of 2–4 × 10–5 in two lines derived from W22 and in 4C063, but not at all in B73 or Mo17, two inbreds widely represented in Corn Belt hybrids. Surprisingly, the mutagenic LTR retrotransposons differed in the active lines, suggesting differences in the autonomous element make-up of the lines studied. Some active retrotransposons, like Hopscotch, Magellan, and Bz2, a Bz1 variant, were described previously; others, like Foto and Focus in 4C063, were not. By high-throughput sequencing of retrotransposon junctions, we established that retrotransposition of Hopscotch, Magellan, and Bz2 occurs genome-wide in the pollen of active lines, but not in the female germline or in somatic tissues. We discuss here the implications of these results, which shed light on the source, frequency, and nature of spontaneous mutations in maize.

Significance

Spontaneously arisen mutations provide the raw material for natural selection and evolution in all organisms, yet their frequency of occurrence is so low that special “mutation accumulation” lines are used to study them. In the course of conventional genetic experiments, we have found that spontaneous mutations arise with unexpectedly high frequency in the pollen of some maize lines, but not of others. Most of them result from insertions of retrotransposons present in a few copies in the genome or from deletions. Most interestingly, the mutagenic retrotransposons differ from line to line. The work reported here sheds light on the source, frequency, and nature of spontaneous mutations in maize.
of up to 2.5 million gametes and obtained estimates of spontaneous mutation frequencies in eight maize genes. Except for two genes, mutation frequencies were in the order of one in a million. Mutations in the two exceptions, R and I, were subsequently shown to be either due to unequal crossing over (22) or to be not heritable (23).

Stadler (20) conducted his experiments in open-pollinated corn fields, in which female rows carrying dominant factors were pollinated with plants carrying multiple recessive markers. Because of its relative ease, this experimental format has been repeated often in experiments that have looked at spontaneous mutation in different loci (23). In other words, prior experiments have measured spontaneous mutation frequencies in female gametes.

The possibility that the picture might be different in male gametes was suggested to us by results from experiments originally designed to isolate new bronze (bz) mutations from different Bz alleles. Crosses of Shrunken (Sh) Bz plants as females or males to a sh bz stock produced putative bz mutations only in crosses with the Bz parent as male. Genetic and molecular characterization of these mutations (Results) established that, surprisingly, several had resulted from insertions of low-copy number LTR retrotransposons (24).

This unexpected observation, coupled with the earlier finding by the Wessler laboratory that several spontaneous mutations had been caused by LTR retrotransposon insertions (13, 15, 25, 26), prompted us to reexamine the issue of spontaneous mutations in maize. We set up a series of large genetic experiments to isolate new bz mutations from several Bz stocks used as male or female parents. These experiments showed that spontaneous mutations in maize arise much more frequently in the pollen parent than in the ear parent, that most of them result from either insertions of low-copy number LTR retrotransposons or deletions, that the frequency of these events varies greatly among maize lines, and that the mutagenic retrotransposons differ from line to line. The results of those experiments, which shed light on the source, frequency, and nature of spontaneous mutations in maize, are discussed in this paper.

Results

Spontaneous bz Mutations. New bz mutations arising in either the female or the male parent were isolated from reciprocal crosses between Sh Bz stocks and an sh bz-R tester. The recessive sh mut mutation serves as a marker for self-pollination contamination in crosses of sh bz ear parents by Sh Bz pollen parents. The recessive bz-R allele is an internal 340-bp deletion in the bz gene (27) and serves to identify and distinguish newly arisen bz mutations.

Our first experiments were set up to isolate new mutations from Bz-B73 and Bz-McC, two alleles differing extensively in sequence and haplotype structure (28). Bz-B73 is the allele present in the B73 inbred: it was back-crossed three times to a stock carrying the sh-bz-X2 multigenic deletion (29) to preserve the B73 Sh-Bz region for studies of recombination in the interval (30, 31). The sh-bz-X2 deletion had been introduced into a W22 background by two backcrosses to Bz-W22 followed by selfing. This reduced inbreeding allowed the weak sh-bz-X2 deletion line to preserve enough vigor and fertility to function effectively as an ear and pollen parent, even under suboptimal field conditions. Bz-McC is an allele extracted from a Barbara McClintock stock and introgressed into a W22 bz-R line obtained from J. L. Kermicle (Department of Genetics, University of Wisconsin–Madison, Madison WI). So, although the two Bz lines have W22-convergent genetic backgrounds, they were derived differently. The results of the experiments with these two lines, identified as W22 Bz-B73 and W22 Bz-McC, are reported in Table 1.

Crosses of the Bz stocks to sh bz-R produced putative bz mutations, identified as rare plump bronze kernels in a purple kernel background, only in crosses with either of the two Bz parents as male. The frequency of their occurrence was similar in the two crosses: 4.3 and 3.6 per 100,000 gametes, respectively, in the Bz-B73 and Bz-McC lines. No mutations were obtained through the female parent as male or female parents. These exceptions from the two lines, identified as W22 Bz-B73 and W22 Bz-McC, are reported in Table 1.

Crosses of the Bz stocks to sh bz-R produced putative bz mutations, identified as rare plump bronze kernels in a purple kernel background, only in crosses with either of the two Bz parents as male. The frequency of their occurrence was similar in the two crosses: 4.3 and 3.6 per 100,000 gametes, respectively, in the Bz-B73 and Bz-McC lines. No mutations were obtained through the female parent as male or female parents. These exceptions from the two lines, identified as W22 Bz-B73 and W22 Bz-McC, are reported in Table 1.

The vast majority of new mutations in both lines arose from insertion of low-copy number LTR retrotransposons or from deletions. However, the proportion of LTR retrotransposon insertions and deletions differed in the two lines. Whereas 18 of 24 concordant mutations in the Bz-B73 stocks were generated as red (Bz) seedlings (nonconcordant) and those that germinated as bronze (bz) seedlings (concordant) and those that germinated as red (Bz) seedlings (nonconcordant). The latter result from postmeiotic or male gametophytic mutations in which a sperm carrying the new bz mutation fertilizes the polar nuclei in the central cell of the embryo sac and a sperm carrying the parental Bz allele fertilizes the egg (33). The red seedlings carrying a wild-type Bz allele were not studied further. DNA was amplified with different bx primers, and sequenced. The results of that analysis are summarized in the last column of Table 1. The vast majority of new mutations in both lines arose from insertion of low-copy number LTR retrotransposons or from deletions. However, the proportion of LTR retrotransposon insertions and deletions differed in the two lines. Whereas 18 of 24 concordant mutations from Bz-B73 arose from retrotransposon

<table>
<thead>
<tr>
<th>Table 1. Spontaneous bz mutations in different maize lines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bz genotype</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>W22 Bz-B73 Male</td>
</tr>
<tr>
<td>W22 Bz-McC Male</td>
</tr>
<tr>
<td>B73 Bz-B73 Male</td>
</tr>
<tr>
<td>Mo17 Bz-Mo17 Male</td>
</tr>
<tr>
<td>4Co63 Bz-4Co63 Male</td>
</tr>
</tbody>
</table>

...
insertions, only 6 of 32 mutations from Bz-McC did. On the other hand, 16 of the 32 mutations from Bz-McC arose from deletions, whereas only 3 of the 24 mutations from Bz-B73 did. Most strikingly, the mutagenic retrotransposons differ in the two lines. Insertions of the Gypsy element Magellan (26) and of Bs2, a Copia element related to Bs1 (34, 35), or of their respective solo LTRs (sLTRs), predominate among bz-B73 retrotransposition mutations (6 of 18 and 12 of 18, respectively), whereas only insertions of the Copia element Hopscotch and of a novel sLTR occur among Bz-McC mutations (2 of 6 and 4 of 6, respectively). Other classes of mutations were observed in smaller numbers: 8-bp insertions in both Bz progenitor alleles, an insertion of a 264-bp MULE element in Bz-B73, and a small duplication and single nucleotide changes in Bz-McC. All of these mutations will be discussed in detail in the next section.

The surprising discovery that low-copy number LTR retrotransposons were unusually mobile in the pollen of the two genetic lines led us to investigate whether that was also true in other maize lines. We decided to extend our analysis to three Corn Belt inbred lines that have been used extensively in either breeding or genetic work: B73, Mo17, and 4C063 (36). Although these inbreds produce colorless seeds because they carry mutations at the R and C genes required for anthocyanin pigmentation in the seed, they are genetically Sh Bz (like all other inbreds) and carry the male-sterile allele. According to maternal inheritance, a cross of Bz-sh-Bz-R female testers. The three inbreds were used as pollen parents on a large number of Sh bz-R females and the resulting ears were screened for plum, bronze kernels, which were analyzed as described earlier. The results are summarized in the bottom three entries of Table 1.

No bz mutations were detected in either B73 or Mo17 in populations of over 200,000 kernels, which would have generated 7–10 mutants if the mutation frequency in these inbreds was as high as in the two genetic lines tested. The upper 95% confidence limits for these three lines are lower than the lower 95% confidence limits of the mutation frequencies in either genetic line. The fact that the B73 inbred carried the same Bz haplotype as the introgressed Bz-B73 genetic line argues that, as one might expect, mobility of the Bs2 and Magellan retrotransposons recovered in bz-B73 mutations does not depend on a specific haplotype, but on the genetic background of the line.

In contrast, bz mutations arose in the paternal progeny of inbred 4C063 with a frequency of 2.6 × 10−5, which is not significantly different from those of the two active genetic lines. Most interestingly, several of the mutations from 4C063 resulted from insertions of a low-copy LTR retrotransposon that have not been previously found among maize mutants. We are naming them here according to the nomenclature for LTR retrotransposons adopted previously found among maize mutants. We are naming them here according to the nomenclature for LTR retrotransposons adopted, which includes an initial “M” for maize, followed by a three-letter designation, such as Bz-McC for the maize Bz-McC retrotransposon relative to “bz” mutations. We are naming them here according to the nomenclature for LTR retrotransposons adopted, which includes an initial “M” for maize, followed by a three-letter designation, such as Bz-McC for the maize Bz-McC retrotransposon relative to “bz” mutations.

Molecular Characterization of the bz Mutations. All of the mutations in this study are designated by the name of the parental allele, followed by the letter “N” to denote their natural or spontaneous origin, and a number assigned to each mutant at germination time. Thus, bz-B73:N1 stands for the first spontaneous bronze mutation from the Bz-B73 genetic line seeded in the greenhouse. Of the 70 concordant bz mutations in Table 1, the most common class resulted from insertion of an LTR retrotransposon. Table 2 presents a detailed summary of those insertion mutations, including the bz:N mutant allele designation, the name of the retrotransposon family, the size of the insertion and of its LTR, the sequence of the 5′-bp target site duplication and its location in the respective Bz allele, and the orientation of the retrotransposon relative to bz.

The retrotransposons appear to be inserted randomly in the two bz exons, roughly in proportion to exon size (10 in exon 1 and 20 in exon 2, compared with 11 and 19, respectively, expected by chance) and in no preferential transcriptional orientation relative to bz (15 same, 13 opposite, and 2 unresolved). The majority terminate in 5′ and 3′ LTRs in direct orientation flanked by 5′-bp target site duplications (TSDs), the standard configuration of an LTR retrotransposon. Complete sequences of both LTRs were obtained for most of these insertions and, as expected from a fresh LTR retroinsertion event, they are 100% identical (39). Differences in 1 or 2 bp seen occasionally between the 5′ and 3′ LTRs most likely stem from PCR amplification or sequencing errors. Exceptions to the standard LTR retrotransposon configuration include insertions with two inverted 3′ ends (bz-B73:N2 and bz-B73:N13), a simple sLTR either flanked by TSDs (bz-McC:N29, bz-4C063N1, bz-4C063N2) or not flanked by TSDs because of adjacent deletions (bz-B73:N6), a compound insertion with 5′ and 3′ LTRs from two different elements (bz-4C063N5), and fragmented elements missing one end and its flanking sequence (bz-4C063N4 and bz-4C063N14).

Magellan Insertions. Six mutations in Bz-B73 arose from insertions of an RLG Magellan retrotransposon. In four of them (bz-B73:N1, bz-B73:N3, bz-B73:N7, and bz-B73:N16), the insertion is the same 5.6-kb element ending in identical 341-bp LTRs that is found in the Maize TE database (36, 38). In two of them, the insertion is much shorter. bz-B73:N6 has a 342-bp sLTR insertion beginning with the terminal AA of the polypurine tract flanking the 3′ LTR and ending with a deletion of the 3′ terminal A of the Magellan LTR. Also missing are 10 bp of the bz first exon corresponding to bases 435–446 in Bz-B73. Consequently, it is not possible to identify the five bases of the TSD. No microhomologies between bz and Magellan sequences at the deletion junction could be identified that would explain the origin of this unusual insertion by a microhomology-mediated rearrangement (40, 41) at or shortly after transposition. bz-B73:N2 has an unusual −0.7-kb insertion with two inverted 3′ ends. The insertion is probably palindromic as our cloned amplification products failed to produce a complete sequence and its size was estimated by Southern blot (SI Appendix, Fig. S1). Our clone of the 5′-bretroelement junction contained an insertion of Magellan’s 3′ terminal 238 bp flanked by the proximal 5′ sLTR sequence on the 5′ side, and 23-bp deletion of bz exon 1 sequence. Including the distal TSD, on the 3′ side. Our clone of the 3′-retroelement-bz junction contained an insertion of Magellan’s 3′ terminal 255 bp flanked by the distal TSD sequence on the 3′ side and an 18-bp of bz exon 1 sequence, including the proximal TSD, on the 5′ side.

Bs2 Insertions. Twelve mutations in Bz-B73 arose from insertions of an RLC retrotransposon closely related to Bsl that we have termed Bs2 (34, 35). Bs2 is 3,269-bp long and has 330-bp LTRs that share only 61% sequence identity with those of Bsl over their entire length, but 90% identity over the first 100 bp. More importantly, the non-LTR sequences of Bs2 and Bsl are 96% identical over their entire length and encode a partial polypeptide with a reverse-transcriptase domain and a segment of a vacuolar ATPase (42). Therefore, they belong to the same maize LTR retrotransposon family, even though their LTRs do not meet the 80% sequence identity family criterion (37). Seven mutations harbor Bs2 elements with 330-bp LTRs: bz-B73:N8, bz-B73:N11, bz-B73:N18, bz-B73:N23, bz-B73:N31, bz-B73:N32, and bz-B73:N39. Four others harbor a closely related version of Bs2 with shorter 275-bp LTRs that we have termed Bs2-short: bz-B73:N21, bz-B73:N25, bz-B73:N33, and bz-B73:N38. Although the 275-bp Bs2s LTR appears to have arisen by a simple deletion close to the 5′ end of the Bs2 LTR, there is no obvious microhomology at the deletion junction to suggest an origin of the former by a DNA replication or repair mechanism.
Table 2. Independent retrotransposon insertions among the spontaneous bz mutations arisen in this study

<table>
<thead>
<tr>
<th>bz-873:N mutants</th>
<th>Retrotransposon insertion</th>
<th>Insertion size (kb)</th>
<th>LTR (bp)</th>
<th>TSD</th>
<th>Insertion site in bz allele</th>
<th>Orientation (relative to bz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>bz-873:N1</td>
<td>Magellan</td>
<td>5.6</td>
<td>341</td>
<td>GCCTC</td>
<td>154–158 exon 1</td>
<td>Opposite</td>
</tr>
<tr>
<td>bz-873:N2</td>
<td>Magellan with two inverted 3’ ends</td>
<td>0.7</td>
<td>341</td>
<td>CCTTC</td>
<td>53–57 exon 1</td>
<td>Cannot tell</td>
</tr>
<tr>
<td>bz-873:N3</td>
<td>Magellan</td>
<td>5.6</td>
<td>341</td>
<td>ACCTG</td>
<td>218–222 exon 1</td>
<td>Opposite</td>
</tr>
<tr>
<td>bz-873:N6</td>
<td>Magellan LTR + deletion bz exon 1</td>
<td>0.3</td>
<td>341</td>
<td>Deleted</td>
<td>Exon 1</td>
<td>Same</td>
</tr>
<tr>
<td>bz-873:N7</td>
<td>Magellan</td>
<td>5.6</td>
<td>341</td>
<td>CACGG</td>
<td>202–206 exon 1</td>
<td>Opposite</td>
</tr>
<tr>
<td>bz-873:N8</td>
<td>Bz2</td>
<td>3.3</td>
<td>330</td>
<td>GTGCC</td>
<td>1002–06 exon 2</td>
<td>Same</td>
</tr>
<tr>
<td>bz-873:N11</td>
<td>Bz2</td>
<td>3.3</td>
<td>330</td>
<td>CGGCC</td>
<td>1140–44 exon 2</td>
<td>Same</td>
</tr>
<tr>
<td>bz-873:N13</td>
<td>Bz2 with two inverted 3’ ends</td>
<td>0.5</td>
<td>232</td>
<td>CGCGG</td>
<td>1298–1302 exon 2</td>
<td>Cannot tell</td>
</tr>
<tr>
<td>bz-873:N16</td>
<td>Magellan</td>
<td>5.6</td>
<td>341</td>
<td>GCCTC</td>
<td>460–464 exon 1</td>
<td>Opposite</td>
</tr>
<tr>
<td>bz-873:N18</td>
<td>Bz2</td>
<td>3.3</td>
<td>330</td>
<td>CCACC</td>
<td>660–664 exon 2</td>
<td>Opposite</td>
</tr>
<tr>
<td>bz-873:N21</td>
<td>Bz2-s</td>
<td>3.2</td>
<td>275</td>
<td>CCTGC</td>
<td>374–378 exon 1</td>
<td>Opposite</td>
</tr>
<tr>
<td>bz-873:N23</td>
<td>Bz2</td>
<td>3.3</td>
<td>330</td>
<td>CCTTC</td>
<td>1186–90 exon 2</td>
<td>Same</td>
</tr>
<tr>
<td>bz-873:N25</td>
<td>Bz2-s</td>
<td>3.2</td>
<td>275</td>
<td>ACGAG</td>
<td>644–648 exon 2</td>
<td>Same</td>
</tr>
<tr>
<td>bz-873:N31</td>
<td>Bz2</td>
<td>3.3</td>
<td>330</td>
<td>TCCTG</td>
<td>1072–76 exon 2</td>
<td>Same</td>
</tr>
<tr>
<td>bz-873:N32</td>
<td>Bz2-s</td>
<td>3.2</td>
<td>275</td>
<td>GGGTG</td>
<td>1250–1254 exon 2</td>
<td>Opposite</td>
</tr>
<tr>
<td>bz-873:N33</td>
<td>Bz2-s</td>
<td>3.2</td>
<td>275</td>
<td>CGCTG</td>
<td>1172–76 exon 2</td>
<td>Opposite</td>
</tr>
<tr>
<td>bz-873:N38</td>
<td>Bz2-s</td>
<td>3.2</td>
<td>275</td>
<td>CGACG</td>
<td>900–904 exon 2</td>
<td>Same</td>
</tr>
<tr>
<td>bz-873:N39</td>
<td>Bz2-s</td>
<td>3.3</td>
<td>330</td>
<td>ACCTC</td>
<td>881–885 exon 2</td>
<td>Opposite</td>
</tr>
<tr>
<td>bz-McC:N mutants</td>
<td>Hopscotch</td>
<td>4.8</td>
<td>231</td>
<td>GCCCT</td>
<td>933–937 exon 2</td>
<td>Same</td>
</tr>
<tr>
<td>bz-McC:N14</td>
<td>Hopscotch</td>
<td>4.8</td>
<td>231</td>
<td>ACCCT</td>
<td>218–222 exon 1</td>
<td>Same</td>
</tr>
<tr>
<td>bz-McC:N29*</td>
<td>RLX-Macpre sLTR</td>
<td>0.5</td>
<td>479</td>
<td>GACGT</td>
<td>818–822 exon 2</td>
<td>Same</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>bz-4Co63:N mutants</th>
<th>Retrotransposon insertion</th>
<th>Insertion size (kb)</th>
<th>LTR (bp)</th>
<th>TSD</th>
<th>Insertion site in bz allele</th>
<th>Orientation (relative to bz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>bz-4Co63:N1</td>
<td>RLX-Fonty sLTR</td>
<td>0.3</td>
<td>278</td>
<td>CTGAC</td>
<td>1016–20 exon 2</td>
<td>Opposite</td>
</tr>
<tr>
<td>bz-4Co63:N2</td>
<td>RLX-Jeli sLTR</td>
<td>0.2</td>
<td>225</td>
<td>CGGTG</td>
<td>1175–79 exon 2</td>
<td>Same</td>
</tr>
<tr>
<td>bz-4Co63:N3</td>
<td>RLX-Fonty</td>
<td>5.3</td>
<td>297</td>
<td>GTGTO</td>
<td>1318–22 exon 2</td>
<td>Same</td>
</tr>
<tr>
<td>bz-4Co63:N4</td>
<td>Fractured RLX-Focou + deletion bz exon 1</td>
<td>2.4</td>
<td>299</td>
<td>CGCTC</td>
<td>51–55 exon 1</td>
<td>Opposite</td>
</tr>
<tr>
<td>bz-4Co63:N5</td>
<td>Compound RE (5’, RLX-Focou; 3’, RLX-Jeli)</td>
<td>4.2</td>
<td>299/224</td>
<td>CGCGG</td>
<td>1371–75 exon 2</td>
<td>Same</td>
</tr>
<tr>
<td>bz-4Co63:N7</td>
<td>RLX-Jeli</td>
<td>3.2</td>
<td>226</td>
<td>CCTGC</td>
<td>777–781 exon 2</td>
<td>Same</td>
</tr>
<tr>
<td>bz-4Co63:N8</td>
<td>RLX-Jeli</td>
<td>3.2</td>
<td>226</td>
<td>CTGGA</td>
<td>1045–49 exon 2</td>
<td>Same</td>
</tr>
<tr>
<td>bz-4Co63:N14</td>
<td>Fractured RLX-Jeli + deletion bz exon 1</td>
<td>2.4</td>
<td>226</td>
<td>TGCTC</td>
<td>83–87 exon 1</td>
<td>Same</td>
</tr>
<tr>
<td>bz-4Co63:N15</td>
<td>RLX-Jeli</td>
<td>3.2</td>
<td>226</td>
<td>CCACC</td>
<td>1185–89 exon 2</td>
<td>Opposite</td>
</tr>
</tbody>
</table>

*Three other mutations were identical to bz-McC:N29.

One mutation (bz-873:N13) harbors a 0.5-kb insertion with two inverted Bz2 3’ ends. As with its counterpart bz-873:N2, we were unable to generate a complete sequence of this insertion and its size was estimated from a Southern blot (SI Appendix, Fig. S2). In both the proximal (5’) and distal (3’) bz-retroelement junctions, the 5-bp TSD flanked a Bz2 3’ terminal sequence, suggesting a palindromic structure of the insertion.

Hopscotch Insertions. Two mutations in bz-McC arose from insertions of an RL Retrotransposon. The insertions in these two mutants (bz-McC:N1 and bz-McC:N18) are identical to each other and to the Hopscotch retrotransposon insertion in wx-K (25). They all encode an apparently complete 1,438-aa gag/pol polyprotein and are, most likely, autonomous retrotransposable elements.

Insertions of New Retrotransposons. Several spontaneous bz mutations were caused by insertion of retrotransposons not previously found among maize mutants. Most of them arose in the inbred 4Co63, whose genome is not yet characterized and which has been used principally in studies of variegated pericarp (43, 44). Six of the nine 4Co63 mutations resulted from insertions of retrotransposons belonging to the RLX-Jeli family (37). Of these, three (bz-4Co63:N7, bz-4Co63:N8, and bz-4Co63:N15) were identical 3.2-kb insertions flanked by 226-bp LTRs and encoding proteins with homology to a pyruvate decarboxylase and a PAP-specific phosphatase, rather than a gag/pol protein. Therefore, these are not insertions of an autonomous retroelement, but of an element that has captured host gene sequences. One mutation (bz-4Co63:N2) was caused by insertion of a 225-bp RLX-Jeli sLTR and is flanked by a 5-bp TSD. Most likely, this sLTR arose by intrachromosomal recombination between the two flanking LTRs of an intact element at the time of or shortly after chromosomal integration. The other two mutations were caused by insertions of more complex RLX-Jeli elements. The insertion in bz-4Co63:N14 is a 2.2-kb fractured RLX-Jeli with an adjacent bz deletion: missing are the 5’ terminal 1,009 bp of the RLX-Jeli element found in the previous mutants, including the entire 5’ LTR, and 35 bp of adjacent bz sequences, including the 5’ TSD. A similar structure consisting of a fractured retrotransposon adjacent to a deletion of the host gene was described for the Magellan-induced wx-M mutation in maize (26). The insertion in bz-4Co63:N5 is the most unusual: a 4.2-kb compound element consisting of 2852-bp from a new RLX element that we have named RLX-Focou at the 5’ end and 1,383-bp of an RLX-Jeli retrotransposon at the 3’ end. This compound insertion is flanked by the same 5-bp TSD, but ends in different LTRs: a 299-bp 5’ LTR from RLX-Focou and a 224-bp 3’ LTR from RLX-Jeli. RLX-Focou encodes fragments of a gag/pol and a PHD protein and is, thus, also nonautonomous.

None of the insertions in the other three bz-4Co63:N mutations are in the database. The 278-bp insertion in bz-4Co63:N1 corresponds to the sLTR from a retrotransposon that we have named RLX-Fonty. The sLTR is flanked by a 5-bp TSD and probably arose by a similar mechanism to that of the sLTR in bz-4Co63:N2. The 5.3-kb insertion in bz-4Co63:N3 is an apparently intact, Gypsy retrotransposon that we have termed RLX-Foto. Its 297-bp LTRs show no significant similarity to any sequences in the maize TE database and its body encodes a 1,468-aa, potentially active, polyprotein, so this may be an autonomous element. Finally, the
insertion in bz-4Co63:N4 corresponds to a 2.4-kb fractured RLX-Focou retroelement accompanied by an adjacent bz deletion. Missing from the element are the terminal 500 bp from the 5’ end, including its 5’ LTR, and missing from bz are 77 bp of adjacent sequence, including the 3’ or distal TSD (the insertion is in opposite orientation to the gene), bz and RLX-Focou share 4 bp of homology at the deletion junction, so the deletion could have arisen as a microhomology-mediated rearrangement at or just after integration.

The last new insertion arose in the Bz-McC genetic line and is unique because of its distinct premeiotic origin. Mutant bz-McC:N29 harbors the insertion of a unique unusually long, 479-bp sLTR from a retrotransposon that we have designated RLX-Macpre. Three other bz-McC:N mutations had exactly the same 479-bp insertion at the same location in exon 2. The simplest explanation for this series of four identical mutations is that they arose in micro-sporogenesis from a premeiotic Macpre retrotransposition event that produced a cluster of pollen grains carrying the same mutation. Other than this set, all of the other retrotransposon mutations in Table 2 were unique and most likely arose at meiosis.

**Other Mutations.** The remaining concordant bz mutations briefly summarized in Table 1 consisted mostly of deletions of various sizes. A few were caused by single-base pair changes, insertions of a MULE element or of 8 bp, and small duplications. Table 3 presents the molecular details of these mutations.

Deletions ranged in size from 19 bp within bz to several kilobases, often including the entire bz gene and sequences from the adjacent stk1 and stk7 genes (45). Because some of the mutations failed to yield a PCR product with any primer pair tested, we do not know the exact size of the larger deletions, but they must stretch for at least 3-kb based on the location of the adjacent gene primers. In many of the other deletions, microhomologies of 4–7 bp could be detected at the deletion junction, suggesting that they could have arisen by a mechanism involving microhomology-mediated repair of a DNA double-strand break (40, 41). The bz-B73:N12 mutation arose by the insertion in exon 1 of a short, and clearly non-autonomous, 264-bp MULE, which produced the 9-bp TSD typical of this transposon superfamily. This short element must have been mobilized by a larger autonomous MULE in the parental line. Three mutations carried 8-bp insertions consisting of either a TCA or TAA trinucleotide plus five bases identical to the five bases preceding the trinucleotide. Possibly, these originated by the failed insertion of Bz2 and Hopscotch retrotransposons, respectively, that left behind the 5-bp TSD separated by three bases from the 5’ and 3’ ends of the LTRs. Only four mutations arose from single-base pair changes, which resulted in either an amino acid change in the highly conserved glucosyltransferase domain of the Bz-encoded protein or in the formation of a premature stop codon.

In a few instances, mutants from the same parent were found to carry the same mutation event, suggesting a premeiotic origin. Thus, bz-B73:N28 and bz-B73:N29 carried the same 19-bp deletion in exon 1; bz-McC:N17, bz-McC:N23, and bz-McC:N25 had the same G-A single-base pair change in exon 2, and bz-McC:N24 and bz-McC:N26 had the same 8-bp insertion in exon 1.

**Spontaneous sh Mutations.** Having found that mutations at the bz locus by either retrotransposition or deletion occurred frequently in the pollen of our two genetic lines, we next asked whether similar mutations could also be recovered at the closely linked sh locus. To identify and subsequently PCR-isolate new sh mutations, we took advantage of the sh-bz-X2 deletion stock, where the entire sh-bz region has been deleted (29), so any new mutations can be unambiguously characterized in the hemizygous condition. We polledinated sh-bz-X2 females with the W22 Sh-B73

<table>
<thead>
<tr>
<th>Table 3. Spontaneous bz mutations not caused by a retrotransposon insertion</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>bz mutant allele</strong></td>
</tr>
<tr>
<td>bz-B73:N12</td>
</tr>
<tr>
<td>bz-B73:N14</td>
</tr>
<tr>
<td>bz-B73:N27</td>
</tr>
<tr>
<td>bz-B73:N28*</td>
</tr>
<tr>
<td>bz-B73:N29*</td>
</tr>
<tr>
<td>bz-McC:N3</td>
</tr>
<tr>
<td>bz-McC:N4</td>
</tr>
<tr>
<td>bz-McC:N5</td>
</tr>
<tr>
<td>bz-McC:N6</td>
</tr>
<tr>
<td>bz-McC:N7</td>
</tr>
<tr>
<td>bz-McC:N9</td>
</tr>
<tr>
<td>bz-McC:N10</td>
</tr>
<tr>
<td>bz-McC:N11</td>
</tr>
<tr>
<td>bz-McC:N12</td>
</tr>
<tr>
<td>bz-McC:N13</td>
</tr>
<tr>
<td>bz-McC:N14</td>
</tr>
<tr>
<td>bz-McC:N171</td>
</tr>
<tr>
<td>bz-McC:N231</td>
</tr>
<tr>
<td>bz-McC:N241</td>
</tr>
<tr>
<td>bz-McC:N251</td>
</tr>
<tr>
<td>bz-McC:N261</td>
</tr>
<tr>
<td>bz-McC:N27</td>
</tr>
<tr>
<td>bz-McC:N28</td>
</tr>
<tr>
<td>bz-McC:N30</td>
</tr>
<tr>
<td>bz-McC:N31</td>
</tr>
<tr>
<td>bz-McC:N32</td>
</tr>
<tr>
<td>bz-McC:N35</td>
</tr>
<tr>
<td>bz-4Co63:N6</td>
</tr>
<tr>
<td>bz-4Co63:N9</td>
</tr>
</tbody>
</table>

*These two mutant alleles had identical mutations.
†These three mutant alleles had identical mutations.
‡These two mutant alleles had identical mutations.
Bz-B73 and W22 Sh-McC Bz-McC lines and isolated shrunken purple kernels against a background of plump purple kernels. Any self-contaminant in the sh-bz-X2 stock would produce a shrunken bronze kernel and thus be ignored. The shrunken phenotype is much more variable than the bronze phenotype, so we applied a lax criterion in the selection of putative mutants. Consequently, many candidates turned out to be false-positives and we cannot distinguish in this experiment true nonconcordant mutations from false-positive selections. The results of the genetic analysis of putative sh mutants are presented in Table 4.

Heritable sh mutations were recovered in both crosses, although less frequently than bz mutations, which is not surprising given the greater variability and line-dependent penetrance of the shrunken phenotype. The mutations were analyzed with a set of eight primer pair combinations covering the entire length of the ~7-kb sh gene (46). Similar types of mutations at sh and bz were observed in each line: a deletion plus two MULE insertions in Sh-B73 and two deletions plus five Hopscotch insertions in Sh-McC. The MULE element recovered in the two sh mutations was the same as that recovered in the bz mutation of the W22 Sh-B73 Bz-B73 line. Similarly, the same retrotransposon, Hopscotch, was recovered in sh and bz mutations of the W22 Sh-McC Bz-McC line. We conclude, therefore, that high pollen mutability is a line-dependent, rather than locus-specific, phenomenon.

The molecular characterization of sh mutants is reported in SI Appendix, Table S1. Two of the sh mutants in the W22 Sh-B73 Bz-B73 genetic line, sh-B73:N2 and sh-B73:N4, arose from insertions of the same MULE element in exons 12 and 13 of the sh gene, respectively. This element is identical to the one causing the bz-B73:N12 mutation in the same line. Therefore, this line must carry an autonomous MULE element encoding the transposase responsible for mobilization of the short elements. Curiously, of the five sh-McC mutations carrying Hopscotch insertions, three have insertion in the 225-bp exon 12. These data suggest that the comparatively short sh exon 12 may be particularly susceptible to transposon insertions.

Generality of Retrotransposition in Pollen. To determine the extent of pollen retrotransposition of Hopscotch, Bs2, and Magellan on a genome-wide scale, we used a combination of high-throughput sequencing and computational techniques. Because we had established genetically that retrotransposition occurs exclusively in the male of W22 Bz-McC and W22 Bz-B73 and not at all in the B73 inbred, we created four different F1 populations by reciprocal crosses: (i) Bz-B73 × W22, (ii) Bz-B73 × W22 Bz-McC, (iii) Bz-B73 × W22 Bz-B73, and (iv) W22 Bz-B73 × Bz-B73. Crosses (ii) and (iv) serve as negative controls because we do not expect new transpositions from the inbred B73 used as male. Cross (i) should identify new transpositions of Hopscotch, the retroelement shown to move in W22 Bz-McC: new junctions should be present only in cross (i) and absent from the other three. Similarly, cross (iii) should identify new transpositions of Bs2 and Magellan, the two retroelements shown to move in W22 Bz-B73: new junctions should be present only in cross (iii) and absent from the other three.

We deep-sequenced retrotransposition junctions in pools of 1,000 individuals from each of these F1 populations with the objective of identifying endogenous and new junctions on the basis of their relative abundance, the new ones expected to be about 1/1,000th as abundant on average as the endogenous ones. The details of this experiment are described in SI Appendix, SI Materials and Methods. In brief, an equal amount of young leaf tissue from a total of 1,000 individuals from each F1 population was harvested and pooled. A modified Splinkerette-PCR method (47) was used to isolate the insertion junctions of the three different retroelements. Sequencing libraries were constructed and sequenced on an Illumina Miseq system. The sequence output was analyzed using InsertionMapper modules (48).

Read number data for each of the three retrotransposition junctions in each of the four F1 populations are given in SI Appendix, Tables S2–S4, which also show the location of the junction sequences in the B73 genome (49). The experiment successfully resolved endogenous and new retroelement junctions on the basis of abundance. The presence of most endogenous elements in the respective parental line was confirmed by PCR, as indicated under the column “PCR results” in each table. In some cases, they were also confirmed in the recently sequenced version of W22 carrying the Bz-W22 haplotype (50). New junctions were present only in the expected F1 population: that is, cross (i) for Hopscotch (SI Appendix, Table S2) and cross (iii) for both Bs2 (SI Appendix, Table S3) and Magellan (SI Appendix, Table S4). The new junctions varied in abundance from 0.01 to 0.0001 of the average endogenous elements, so we adopted an empirical cut-off to accept a retrotransposition junction as representing a new germinal transposition or not. We ran PCRs on low-abundance candidates and chose as a cut-off limit the lowest abundance junction that could be successfully amplified in the respective DNA pool (SI Appendix, Tables S2–S4). This empirical approach resulted in the acceptance of practically all of the low-read candidates that met the initial selection criterion of being present only in the expected cross. The results of the next-generation sequencing (NGS) analysis are summarized in Tables 5 and 6.

The number of endogenous insertions in the active lines varies from five to seven, most of them being inserted in nonrepetitive DNA (Table 5). By comparing the presence or absence of each endogenous element in the parental lines with the known retrotransposon activity of each line, we were able to identify potentially autonomous elements. Thus, a unique Hopscotch insertion in chromosome 2 of W22 Bz-McC (SI Appendix, Table S2) and a unique Magellan insertion in chromosome 5 of W22 Bz-B73 (SI Appendix, Table S4) may be autonomous elements. Interestingly, Bs2, the most active retroelement in our study, has four potentially autonomous insertions: one each in chromosomes 1, 2, 5, and 8 of W22 Bz-B73 (SI Appendix, Table S3).

The NGS data for Magellan are supported by Southern blot data (SI Appendix, Fig. S3). The restriction enzyme HindIII cuts within Magellan, separating the two LTRs, so a HindIII digest of the various lines should produce twice as many LTR-hybridizing bands as the number of endogenous insertions. NGS uncovered

Table 4. Spontaneous sh mutations from Sh lines used as male parents

<table>
<thead>
<tr>
<th>Sh genotype</th>
<th>Effective kernel population</th>
<th>Shrunken selections</th>
<th>Shrunken mutations</th>
<th>Mutation frequency (×10^-5)</th>
<th>Nature of sh mutants</th>
</tr>
</thead>
<tbody>
<tr>
<td>W22 Sh-B73</td>
<td>439,943</td>
<td>24</td>
<td>5</td>
<td>1.1 (0.4–2.6)*</td>
<td>One deletion</td>
</tr>
<tr>
<td>W22 Sh-McC</td>
<td>408,468</td>
<td>14</td>
<td>8</td>
<td>2.0 (0.7–3.4)*</td>
<td>Two undefined</td>
</tr>
</tbody>
</table>

*95% Confidence limits.
7 endogenous Magellan insertions in W22 Bz-B73 and 6 in W22 Bz-McC, and the Southern blot detected 13 bands in the former and 11 in the latter. Interestingly, the hybridization patterns of the two W22-converted genetic lines are very similar, supporting their largely common genetic background. An ~11-kb band present in most W22 lines, including the recently sequenced Bz-W22 version of W22 (50) but absent from Bz-McC, may correspond to the active autonomous Magellan element, putatively assigned to chromosome 5 by the NGS analysis. Not surprisingly, Magellan insertions have been found in at least two other mutations arisen in W22 lines: pl1-987 (GenBank AF015269.1) and bz-s39.71 (51).

The number of new insertions in the active lines was more variable (Table 6). Among the 1,000 progeny from W22 Bz-McC males, we found 18 new insertions of Hopschotch, that is, 1.8% of the progeny have a new Hopschotch transposition event. Among the 1,000 progeny from W22 Bz-B73 males, we found 300 new insertions of Bz2 and 91 new insertions of Magellan, for overall transposition percentages of 30% and 9.1%, respectively. Our observed frequencies of Bz mutations (Table 1) are in line with these numbers. The single gene and genome-wide analyses establish that Bz2 is the most active retroelement of the three, possibly because of the larger number of putative autonomous elements in W22 Bz-B73 (SI Appendix, Table S5). Significantly, the vast majority of new insertions are in non-repetitive sequences. Of the 409 new insertions identified in this study, 370 are in unique sequences with ≥85% identity in the B73 genome and of them, 264 or 71% fall within 500 bp of gene models.

**Somatic Retrotransposition.** A possible explanation for the low-coverage retrotransposon junctions detected in our NGS experiments, particularly with Bz2 and Magellan, is that they arise from somatic retrotransposition events in the individual seedlings. To test this possibility, we looked for novel junctions at five different developmental stages of two individual plants from cross (iii) B73 × W22 Bz-B73. The developmental stages included the 3rd, 7th, and 11th leaf of plant no. 2 (Table S7). Among the 1,000 progeny seen with Bz2 and 91 new insertions of Magellan, for overall transposition percentages of 30% and 9.1%, respectively. Our observed frequencies of Bz mutations (Table 1) are in line with these numbers. The single gene and genome-wide analyses establish that Bz2 is the most active retroelement of the three, possibly because of the larger number of putative autonomous elements in W22 Bz-B73 (SI Appendix, Table S5). Significantly, the vast majority of new insertions are in non-repetitive sequences. Of the 409 new insertions identified in this study, 370 are in unique sequences with ≥85% identity in the B73 genome and of them, 264 or 71% fall within 500 bp of gene models.

<table>
<thead>
<tr>
<th>Retro-transposon</th>
<th>Line*</th>
<th>No. endogenous sites</th>
<th>No. sites with B73 hits†</th>
<th>No. sites in genes‡</th>
<th>No. sites in IGRs§</th>
<th>No. sites in repetitive DNA¶</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hopschotch</td>
<td>W22 Bz-McC</td>
<td>5</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Bz2</td>
<td>W22 Bz-B73</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Magellan</td>
<td>W22 Bz-B73</td>
<td>7</td>
<td>6</td>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

*Line where retrotransposon is active.
†≥85% identity.
‡≤500 bp.
§Intergenic regions.
¶Five copies.

As expected from the earlier genetic and genome-wide studies, new Hopschotch junctions were observed in either of the two B73 × W22 Bz-B73 plants (SI Appendix, Table S5). New Bz2 and Magellan junctions were observed in the immature tassel samples from both plants (SI Appendix, Tables S6 and S7). These probably represent independent transposition events in different pollen grains, as the immature tassels had already gone through meiosis when collected. In agreement with the earlier NGS study, more new junctions were seen with Bz2 than with Magellan (23 vs. 5), evenly distributed between the two plants. Peiccularly, one new Bz2 junction was detected in the immature ear of plant no. 2, suggesting that Bz2 retrotransposition may occur infrequently in the female germline (SI Appendix, Table S6). The only evidence for a somatic retrotransposition event in a new Magellan junction in the 11th leaf of plant no. 2 (SI Appendix, Table S7). We conclude that somatic retrotransposition in maize occurs rarely, if at all.

**Discussion**

**Spontaneous Mutations.** We have shown here in large population experiments that stable mutations at the maize bz locus can arise spontaneously with the surprisingly high frequency of ~4 × 10⁻³ in the pollen, but not ear, progeny of some lines. Previous large-scale studies of spontaneous mutation in maize had been largely restricted to the female parent because of crossing convenience, so a high spontaneous mutation frequency in the pollen was unexpected. This high-mutation frequency was found to be a property of some lines, but not others. Specifically, two genetic lines, derived recently by introgressing the Bz-McC and Bz-B73 haplotypes into a W22 genetic background, and the long-standing inbred 4Co63 (36) showed this property, but B73 and Mo17 Corn Belt inbreds representing the two heterotic groups commonly used in hybrid corn production (52) did not. Thus, spontaneous mutation frequency in the pollen is a polymorphic trait in maize. The high-pollen mutation frequency of the two genetic lines was not restricted to bz, as it could also be detected at the closely linked sh locus.

An analysis of the bz and sh mutations revealed that they resulted mostly from either insertions of low-copy number LTR retrotransposons or deletions (Tables 1 and 2). These are the same two main classes of mutations found in the collection of spontaneous wx mutations by the Wessler laboratory (13, 15, 25, 26), suggesting that those mutations also arose in microsporogenesis and that perhaps most spontaneous stable mutations in maize do. Oliver Nelson had

| Retrotransposon insertion sites identified by NGS in pollen progeny of active lines |
|----------------------------------|----------------|----------------|----------------|----------------|
| Retrotransposon | Line* | No. new sites | No. sites with B73 hits† | No. sites in genes‡ |
| Hopschotch       | W22 Bz-McC | 18             | 16                | 12             |
| Bz2              | W22 Bz-B73 | 300            | 270               | 187            |
| Magellan         | W22 Bz-B73 | 91             | 84                | 65             |

*Line where retrotransposon is active.
†≥85% identity.
‡≤500 bp.
§Intergenic regions.
¶Five copies.
collected the majority of these mutations from genicists and breeders, who would send them to him on a regular basis, but without a record of whether they had arisen in the male or the female germline. However, surprisingly few other retrotransposon-caused mutations have been described in maize, perhaps because most mutations have been isolated from the female germline. An interesting case is adhl1-S5446, which harbors Bs1, the first LTR retrotransposon described in maize (34). This mutation occurred as a single kernel in the paternal progeny of a barley stripe mosaic virus-treated plant whose pollen was exposed to allyl alcohol selection. Other examples of spontaneous mutations with retrotransposon insertions are: brown midrib3-1 (bm3-1) with a B5 element of undetermined size (53); d-3;4, with a 327-bp Sleepy solo-LTR (54); pl1-987, with a 1.2-kb Magellan element (GenBank AF015269); and i-o, with a Hopscotch insertion of undetermined size (55). One common feature of all these retroelements is that they are present in a low number of copies (<10) in the maize genome. In the genetic lines used in this study, W22 Bz-McC and W22 Bz-B73, these same elements were shown to be mobile: Hopscotch in W22 Bz-McC and Magellan and a close relative of Bs1 that we have dubbed Bs2 in W22 Bz-B73. A few mutations were unrelated to retrotransposon insertions or deletions. The preeminent single base pair missense bz mutation detected in W22 Bz-McC came from a screen of over a million gametes. The single base pair mutation detected in the inbred 4Co63 came from a screen of over half a million gametates. In contrast, the gamete populations screened for B73 and Mo17 were slightly larger than 200,000, so the failure to detect similar single base pair mutations in these inbreds may be purely a frequency issue. Similarly, nonautonomous MULE insertions in bz and sh were detected only in W22 Bz-B73, most likely because only this line carries the corresponding autonomous MULE transposon. LTR Retrotransposon Mobility. The most plausible interpretation for the retrotransposon mobility in these lines is that the active line carries the autonomous element encoding the retrotransposition functions required for movement of elements in the same family (56). Thus, W22 Bz-McC should carry at least one autonomous Hopscotch element, but the B73 inbred and W22 Bz-B73 should not. Similarly, W22 Bz-B73 should carry at least one autonomous element of both Magellan and Bs2, but B73 and W22 Bz-McC should not. Our genome-wide survey of endogenous insertion sites of these elements (SI Appendix, Tables S2–S4) supports this. Unique retrotransposon–host junctions were detected for Hopscotch in chromosome 2 of W22 Bz-McC (SI Appendix, Table S2) and for Magellan in chromosome 5 of W22 Bz-B73 (SI Appendix, Table S4). For Bs2, the most active retroelement in our study, four unique retrotransposon–host junctions in different chromosomes were detected in W22 Bz-B73 (SI Appendix, Table S3). These unique retrotransposons junctions, or at least one of them in the latter case, most likely correspond to the autonomous elements. It is interesting that the B73 and Mo17 inbreds widely used by breeders lack retrotransposon activity. One is tempted to speculate that, in addition to their desirable agronomic properties, these inbreds might have been selected for their genetic stability and against the low-level spontaneous variation reported in some long-time inbred lines (57, 58).

The analysis of new retrotransposon junctions in the male progeny of W22 Bz-McC and W22 Bz-B73 (SI Appendix, Tables S2–S4) allowed us to conclude that high retrotransposition in pollen was a genome wide phenomenon, and not just restricted to the loci that had been assayed genetically. Among the pollen progeny from W22 Bz-McC, 1.8% carried a new Hopscotch transposition event and among those from W22 Bz-B73, 30% carried a new Bs2 and 9.1% a novel Magellan transposition event, respectively. Most new retrotranspositions are in nonrepetitive sequences: 90% are in unique sequences and of them, 71% are within 500 bp of gene models. Clearly, like the HIV and murine leukemia virus human LTR retroviruses (59, 60)—and the stress-activated Tos17, LORE1, and Osen retrotransposons in rice, Lotus, and Arabidopsis, respectively (61–63)—these low-copy maize retrotransposons integrate preferentially into genomic regions.

Our genetic study with the inbred 4Co63 identified LTR retrotransposons not previously described in the B73 genome (37). This is not surprising and, most likely, countless new mobile retrotransposons will be identified in other maize genomes, as they are analyzed, just as they have been identified in Arabidopsis ecotypes other than the Col-0 reference (64). We have named the new mobile elements by combining parts of the inbred’s singular name (Four County Sixty-three). Photo is a 5.3-kb, apparently intact and possibly autonomous Gypsy retrotransposon encoding a 1,468-amino acid polypeptide, Focou is a 2.9-kb RLX element encoding fragments of a gappol and a PHD protein, so it resembles Bs1 and Bs2 in having captured a host gene fragment (42), and Fonty corresponds to a novel 228-bp sLTR.

We can only speculate as to why low-copy number maize LTR retrotransposons are mobile in the male, but not the female, germline. The simplest explanation is that the promoters in the LTRs of these retrotransposons are pollen-specific, as is the LORE1 promoter from Lotus japonicus (65). Alternatively, these LTR retrotransposons may be less methylated in the male germ line than in the female germline. In Arabidopsis, CHH methylation is lost from retrotransposons in microspores and sperm cells and restored in the embryo after fertilization (66).

Structure of Recently Mobilized Retrotransposons. The majority of the LTR elements that transposed into the bz locus (21 of 30) end in two LTRs in direct orientation, the standard LTR retrotransposon configuration, and are flanked by 5-bp TSDs. The two LTRs are identical in sequence, as expected of newly transposed elements (39). However, a sizable minority of new insertions had a different structure. The most common exception was a solo LTR flanked by a 5-bp TSD (three of nine: bz-McC:N29, bz-4Co63:N1, bz-4Co63:N2). These most likely originated by an intrachromosomal recombination event between the two retrotransposon LTRs at the time of integration or shortly thereafter. Their occurrence as new, and not historical, products indicates that sLTRs can originate at the time of transposition, rather than in subsequent generations following the integration of a complete retrotransposon (67–69).

Three insertions were accompanied by deletion of adjacent bz sequences. bz-B73:N6 has an insertion of what appears to be a Magellan sLTR, but the insertion is not simple: its first two bases, AA, are remnants of the Magellan polypurine tract immediately preceding the 3′ LTR and terminal A of the LTR is missing, together with 10 bp of the 5′ first exon. Two mutants have insertions of fractured elements missing one end and its flanking sequence. bz-4Co63:N4 has an insertion of a 2.4-kb truncated RLX-Focou element missing its 5′ end, including the entire 5′ LTR, adjacent to a 77-bp deletion of the bz first exon, including the TSD. Focou and bz share 4 bp of homology at the deletion junction, so the deletion presumably arose by microhomology-mediated recombination at or just after integration. bz-4Co63:N14 has an insertion of a 2.2-kb truncated RLX Jelli missing 1 kb from its 5′ end relative to other RLX Jelli insertions (Table 2) adjacent to a 35-bp deletion of the bz first exon, including the TSD. Jelli and bz share 5 bp of homology at the deletion junction, so this deletion also probably arose by a microhomology-mediated recombination event around the time of integration.

Two of the most unusual mutations had insertions with inverted 3′ ends flanking 5-bp TSDs. bz-B73:N2 had a 0.7-kb insertion consisting of the inverted 3′ ends of two Magellan LTRs and bz-B73:N13 had a 0.5-bk insertin consisting of the inverted 3′ ends of two Bs2 LTRs. For these two insertions, we were able to obtain the sequences of the junctions, but not of the entire element, and estimated their sizes from Southern blots (SI Appendix, Figs. S1 and S2). Not having the entire sequence of the insertions, it is difficult to envision how these complex structures arose during the retrotransposition process. Perhaps the mutation with the most complex insertion is bz-4Co63:N5. It carries a compound insertion with 5′ and 3′ LTRs from two different elements. The 4.2-kb element consists of 2.8-kb from a Focou element at the 5′ end and 1.4-kb of a Jelli element at
the 3′ end. The compound insertion is flanked by the same bz 5-bp TSD, but ends in different LTRs: a 299-bp 5′ LTR from Focou and a 224-bp 3′ LTR from Jeli. This hybrid element probably arose by recombination between a Focou and a Jeli retrotransposon during transposition, by a process similar to retrotransposon insertion where two distinct RNA templates copackaged in a single virus-like particle can lead to new combinations of sequences (70). A high level of pairwise recombination between members of an LTR retrotransposon family was also reported recently in Arabidopsis (71).

The simple 8-bp insertions detected in three mutants—bz-B73: N27, bz-McC:N24, and bz-McC:N26—can also be considered to be the products of abortive retrotransposition events. They consist of either a TCA or TAA trinucleotide flanked on either side by a 5′-bp direct repeat. We infer that these retroelements arose by the failure of Bz and Hopscotch retrotransposons, respectively, that left behind three bases from their 5′ and 3′ ends separated by a 5-bp TSD.

Timing of Retrotransposition. About one-third of all new insertions have noncanonical LTR retrotransposon structures and many of them can be explained by recombination events during transposition, suggesting that, as in fission yeast (72), retrotransposition occurs at the time of DNA replication, probably at meiosis I of the maize male germline. The observation that all but one of the 30 mutations listed in Table 2 were unique supports this inference. The one exception was mutant bz-McC:N29, which harbors a novel 479-bp Macro sLTR insertion in exon 2. Three other bz-McC:N mutations had the same insertion at the same location, suggesting that premeiotic retrotransposition events that produce sectors with the same mutation can also occur during microsporogenesis. Like its maize counterparts, the Lotus LÖREI retrotransposon also transposes preferentially in the male and does so late in development because all new insertions are unique (65). The situation in Arabidopsis is slightly different. Retrotranspositions of Evadé in epirogenetic cells were found to be nonunique, although they appear to arise in the female reproductive lineage (73), and the heat-induced retrotranspositions of Onsen in siRNA-deficient plants also occur in sporogenous tissues, but are more common premeiotically (74).

Germline vs. Somatic Retrotransposition. Our genetic and genomic data indicate that the mobile LTR retrotransposons Hopscotch, Bz2, and Magellan move in the male germline but not the female germline. To examine if these retrotransposons move in somatic tissues, we looked for new retroelement–host junctions for the three elements at five different developmental stages, including seedling, juvenile, and mature leaf, immature tassel, and immature ear. The experiment (SI Appendix, Table S7) shows that somatic retrotransposition in maize is a very rare event and that the LTR retrotransposons studied here move mainly in the male germline.

Deletions and Retrotranspositions: A Common Mechanism? Retrotranspositions and deletions combined were, by far, the predominant bz mutations that arose in the pollen of W22 Bz-B73, W22 Bz-McC, and 4Co63, although deletions outnumbered retrotranspositions in W22 Bz-McC. However, no mutations of either kind were recovered in either B73 or Mo17 in populations large enough to have recovered several mutants if the mutation frequency in those inbreds was as high as in the lines tested. The absence of both mutant classes in these inbreds prompts us to speculate that deletions and retrotranspositions share a common origin.

LTR retrotransposition is clearly a sloppy process, as one-third of the insertions recovered at the bz locus lacked the standard structure of an LTR retrotransposon and showed evidence of recombination within the retroelement or between the retroelement and adjacent bz sequences. Others, like the 8-bp insertions, lacked a clear retroelement vestige, but could be explained by abortive transpositions that left behind three nucleotides from the 5′ and 3′ ends of the LTRs.

The deletions ranged in size from 19 bp within bz to at least 3 kb outside bz, based on the position of the primers used in PCR amplification. In many of the smaller deletions, microhomologies of 4−7 bp could be detected at the deletion junction, suggesting an origin by microhomology-mediated repair of a DNA double-strand break (40, 41). Similar microhomologies at deletion junctions were found among spontaneous wx deletions (14) and newly originated dmIDR, Ds, and d TED elements (75−77) in maize. We propose that these deletions arise from the cell’s DNA repair machinery in the process of repairing the break created by the retrotransposon endonuclease and move at a relatively high frequency in the pollen of certain maize lines. Unlike previous reported cases of mobile plant retrotransposons, they are not stimulated to move by any stress or known epigenetic factors. Rather, in analogy to the well-characterized DNA transposons, they appear to move in lines that most probably carry an autonomous element expressing the proper retrotransposition functions. A prediction of this work is that mobilization of potentially mobile low-copy nonautonomous retrotransposons remaining in nonrepetitive—and most likely hypomethylated—DNA could be transferred to an inactive maize line from an active one simply by crossing. This is clearly a controlled way of mobilizing retrotransposons, an attractive and largely untapped source of genetic variation (81). It remains to be seen whether similar retrotransposon mobility polymorphisms will be found in other crops.

Materials and Methods

Because of their importance to this study, the genetic lines used are described in Results. Detailed description of other plant materials and molecular methods can be found in SI Appendix. Included in the latter are the molecular characterization of mutations and the high-throughput sequencing experiments to identify and map retroelement insertion sites. Sequences reported in this paper have been deposited in GenBank with accession nos. MK532510–MK532536 (82).

ACKNOWLEDGMENTS. We thank Brian Gelfand and Dibyendu Kumar in the Waksman Genome Facility for help with next-generation sequencing of DNA; Daja O'Brien for assistance with computer network issues; and Marc Probasco and Harrison Wang for field management and greenhouse plant care. This research was supported by National Science Foundation Grants MCB 09-20218 (to H.K.D.) and IOS 13-39238 (to H.K.D. and C.D.). J.T.H. was a recipient of a Busch−Waksman Postdoctoral fellowship.


The structure of inducible germline elements in maize: Evidence for rare repair synthesis at the site of 40:327


Plant Cell


