Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 8731–8735, November 1990
Genetics

Filler DNA is associated with spontaneous deletions in maize

(waxy/restriction fragment length polymorphism/mutation)

SUSAN WESSLER, ALACIA TARPLEY, MICHAEL PURUGGANAN, MARION SPELL, AND RON OKAGAKI[†]

Botany Department, University of Georgia, Athens, GA 30602

Communicated by Barbara McClintock, August 21, 1990

ABSTRACT **We have determined the structure of five spontaneous deletions within the maize waxy (*Wx*) gene. Of these, four were found in spontaneous *wx* mutants (*wx-B*, *wx-B1*, *wx-B6*, *wx-C4*) and include exon sequences; the fifth is restricted to an intron and represents a restriction fragment length polymorphism of a nonmutant allele (*Wx-W23*). The deletions, which range in size from 60 to 980 base pairs (bp), cluster in a G+C-rich region of ≈1000 bp that is capable of forming stable secondary structures. Most striking is our finding that all of the alleles have DNA insertions (filler DNA) of 1–131 bp between the deletion endpoints. For three of the five deletions, the filler DNA and sequences at the deletion termini appear to be derived from sequences near one deletion endpoint. A previously reported spontaneous deletion of the maize bronze gene (*bz-R*) also contains filler DNA. The association of filler DNA with maize deletion endpoints contrasts dramatically with the rarity of similar events in animal germ-line and bacterial mutations.**

Analysis of deletion endpoint sequences has provided insight into the diversity of mechanisms underlying illegitimate recombination in prokaryotic and eukaryotic cells. In *Escherichia coli*, short direct repeats of between 5 and 10 base pairs (bp) are found in progenitor alleles flanking the sites where the majority of spontaneous deletions occur (1, 2). Pairing of one direct repeat copy with its complement at the second repeat is hypothesized to initiate deletion formation. Replication or repair of these "slipped mispaired" structures may result in deletion of the intervening DNA and one copy of the repeat (2). The direct correlation of the length of a repeat with the frequency of deletion formation underscores the role of homologous pairing between these repeats in deletion formation (1).

Deletion endpoints, derived from both somatic and germ-line cells, have been sequenced in mammals. Like *E. coli* deletions, somatic cell deletions appear to be flanked by direct repeats, although the repeat length is apparently shorter than in *E. coli* (3). For example, direct repeats of 2–5 bp flanked seven deletions of the *aprt* gene isolated from cultured (somatic) hamster cells (4). In comparison with somatic cell and prokaryotic deletions, a smaller fraction of germ-line deletions appear to be associated with direct repeats (3). Rather, a member of the highly repetitive *Alu* sequence family is frequently found near the endpoints of deletions within the α- and β-globin genes (5, 6) and the low density lipoprotein receptor of mammalian germ lines (7, 8). In addition, the finding that 16 independent deletions of the *unc-54* gene of *Caenorhabditis elegans* do not occur near obvious direct repeats or repetitive DNA elements (9) serves to underscore the diversity of illegitimate recombination mechanisms in eukaryotic germ-line cells.

The characterization of illegitimate recombination mechanisms in higher plants is of particular interest in light of apparent differences detected in mammalian somatic and germ-line deletions. In contrast with the situation in animals where the germ-line is segregated early in development, plant germ cells differentiate from apical meristems after prolonged vegetative growth. Thus, spontaneous deletions in plants might be expected to share features with mammalian somatic deletions. Unfortunately, molecular analysis of deletions in higher plants has, for the most part, been restricted to the characterization of transposable element-mediated events in *Zea mays* and *Antirrhinum majus*. Several deletions within the autonomous elements *Ac* and *Spm* have been sequenced (10–12) and aberrant excisions of the element that delete flanking host sequences have also been determined (13–15). It is unclear whether these studies are relevant to an understanding of the mechanisms underlying spontaneous deletions in plant cells because deletions within and flanking transposable elements are probably mediated by the elements' transposase (11).

The mutant alleles of the maize waxy (*Wx*) gene provide a unique opportunity to investigate the molecular basis of spontaneous mutation in a higher plant. The *Wx* gene encodes an enzyme that is responsible for the synthesis of amylose in the endosperm, pollen, and embryo sac tissues (16). The viable and easily scored phenotypic expression of *wx* mutants has facilitated the isolation of many mutant alleles. In a previous study, we determined that 6 of 17 spontaneous mutations had sustained a deletion of part or all of the *wx* gene (17). For four of the six deletions (*wx-B*, *wx-B1*, *wx-B6*, *wx-C4*), we were able to define both deletion endpoints within or flanking the gene. We determined that these deletion termini were clustered within a 1000-bp region, suggesting that this might be a hot spot for deletion formation. The same study also revealed that one nonmutant *Wx* allele (*Wx-W23*) had sustained a small deletion within the gene. These findings, coupled with the paucity of sequenced spontaneous deletion endpoints in higher plants, prompted a detailed characterization of these mutations.

MATERIALS AND METHODS

DNA Cloning and Sequencing. The strains used in this study have been described (17). Isolation of genomic DNA from plantlets (18) and Southern blot analysis and phage screening (19) were as cited. The deletions of *wx-C4*, *wx-B*, *wx-B1*, *wx-B6*, and *Wx-W23* were previously localized by Southern blot analysis of genomic DNA (17). These genes were cloned and sequenced as follows (see Fig. 1). For *wx-B6*, the entire gene was cloned on an 11-kilobase (kb) *Eco*RI fragment after ligation of *Eco*RI-digested genomic DNA into the γ vector EMBL4 (20). The library was packaged, plated on *E. coli* strain LE392, and screened as described (19). The lesion was localized to a 0.7-kb *Sal* I/*Pst* I fragment, which was subcloned into M13mp18 (21) and sequenced by the dideoxynucleotide chain termination method (22).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

[†]Present address: Genetics Department, University of Wisconsin, Madison, WI 53706.

For *wx-C4*, the 1.9-kb *Sal* I fragment containing the deletion (17) was purified from low melting point agarose and ligated to Lambda ZAP arms (Stratagene) that had been digested with *Xho* I and dephosphorylated. The insert was subcloned into pUC119 (23) on a 1.9-kb *Sal* I/*Kpn* I fragment (the *Kpn* I site is in the γ vector) and sequenced. The *wx-B* lesion was cloned on a 5.2-kb *Sal* I fragment (17) after gel purification of the fragment and ligation into *Xho* I-digested arms of Lambda ZAP. The lesion was subcloned on a 2.0-kb *Sst* I fragment in pUC118 and unidirectional deletion clones were generated (24) and sequenced. The *wx-B1* deletion was cloned on a 2.8-kb *Sst* I fragment (17), subcloned into pUC118, and sequenced. The DNA containing the *Wx-W23* restriction fragment length polymorphism was isolated after polymerase chain reaction amplification as described (25). Genomic DNA (150 ng) and primers (700 ng) were cycled 45 times for 30 sec at 94°C and for 6 min at 72°C followed by a final incubation of 5 min at 72°C. Primers for amplification, 5′-CGCGTGTTCGTTGACCACCC-3′ (positions 1023–1042) and 3′-AGGCGCTGAACAAGGAGGCG-5′ (positions 2025–2044), were derived from the *Wx* sequence (26). Amplified DNA was digested with *Sst* I and *Pst* I and subcloned into pUC119 and sequenced.

Secondary Structure Analysis of DNA. Analysis of the *Wx* and *Bz* genes and flanking sequences utilized Zuker's folding program (27) incorporated in the Wisconsin Genetics Computer Group sequence analysis package. DNA stacking energies were obtained from Breslauer *et al.* (28). Loop energies were taken from Freier *et al.* (29). The program was implemented on the VAX/VMS 6210 system of the University of Georgia Biological Structure and Computational Facility.

RESULTS

DNA Sequencing of Deletion Endpoints. Fragments containing the deletion endpoints from *wx-B1*, *wx-B6*, *wx-C4*, and *Wx-W23* were cloned and sequenced (see *Materials and Methods*). The size and position of the five deletions relative to the *wx* introns and exons are shown in Fig. 1. Only 7 of the 14 *Wx* exons are depicted. The DNA sequences of the deletion endpoints are shown in Fig. 2 and the results are summarized in Table 1. The molecular basis of the mutant phenotypes for *wx-B1*, *wx-B6*, *wx-B*, and *wx-C4* is evident from their respective deletions. *wx-B1* is missing 655 bp

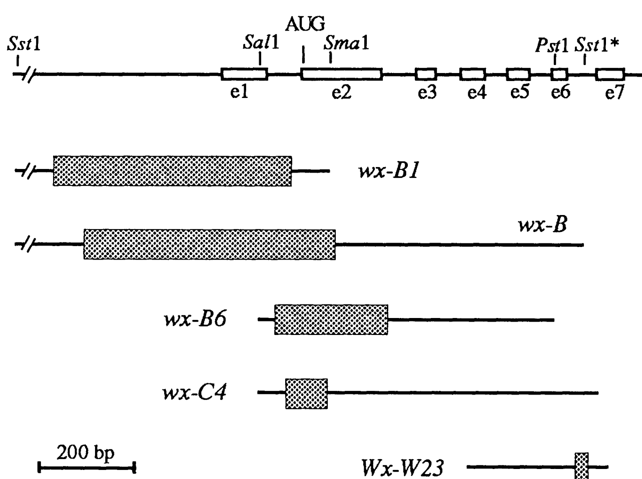

FIG. 1. Positions of deletions relative to the *Wx* gene. The positions of exons 1–7 (of the 14 *Wx* exons) and the translational start are noted (26). Shaded boxes indicate the extent and position of each deletion. Only those restriction sites mentioned in the text are depicted. The *Sst* I* site is deleted in *Wx-W23* and its derivative *wx-B1*.

upstream of the transcription start, all of exon 1, and part of intron 1. *wx-B* is also missing sequences upstream of the transcription start site in addition to exon 1, intron 1, and the AUG start in exon 2. *wx-B* and *wx-C4* are missing the amino terminus, which encodes the transit peptide (18, 26) that is required for transport of the protein into amyloplasts, the site of starch biosynthesis.

Unlike the mutant alleles, the deletion in *Wx-W23* is restricted to intron sequences. Although the deletion begins 11 bp from the 5′ end of the intron, RNA splicing is apparently normal; no quantitative or qualitative difference in Wx mRNA was detected when *Wx-W23* was compared to other nonmutant *Wx* alleles (M. Varagona, R.O., G. Baran, and S.W., unpublished observation).

One striking feature of all the lesions is the presence of from 1 to 131 bp between the deletion endpoints (Fig. 2, F′, lowercase letters). Extra DNA between the endpoints of DNA rearrangements, translocations, and deletions has been observed in other eukaryotes and has been referred to as either filler DNA or insertion junctions (30). For *wx-B1*, *wx-C4*, and *Wx-W23*, the filler DNAs are derived from sequences within 50 bp of one deletion endpoint (Fig. 2, F, underlined boldface letters). Interestingly, this nearby sequence also contains both the 5′ and 3′ deletion endpoints (Fig. 2, A, A′, B, and B′; underlined with the filler DNA). The extent of sequence identity between the adjacent sequence and the endpoint region is dramatic: 18/18 bp for *wx-C4* and *Wx-W23* and 17/17 bp for *wx-B1*. In contrast, the origin of the filler DNA of *wx-B6* is not as obvious because it is only 1 bp. Although the sequence that bridges the deletion, TGATG (Fig. 2, underlined), is found within 43 bp (Fig. 2, underlined boldface letters), the 5/5-bp match is not extensive and may simply reflect a random occurrence. The filler DNA in *wx-B* is large (131 bp) and complex. This DNA also appears to be derived from sequences near the deletion endpoints and from reiterations of these sequences. Unlike the other deletions, there is no simple way to bridge the deletion endpoints with flanking sequences.

There is only one other spontaneous deletion mutant in maize for which the sequence of endpoints is available. The *bz-R* allele has sustained a 340-bp deletion that includes part of exons 1 and 2 of the bronze gene (31). In describing this mutation, Ralston *et al.* (31) noted that there were 5 bp between the deletion endpoints (Fig. 2, X′). Interestingly, this filler DNA, together with the adjacent 4 bp (Y′) share 8/9 bp identity with an upstream sequence (Fig. 2, XY, underlined boldface letters).

Secondary Structure Analysis of *Wx* DNA. The endpoints of the large *wx* deletions (*wx-B*, *wx-B1*, *wx-C4*, and *wx-B6*) fall in a region of ≈1 kb including ≈600 bp of 5′ flanking sequences and *Wx* exons 1 and 2 (Fig. 1). To understand why the deletions were clustering within the *Wx* gene, the sequence of the entire gene was analyzed for the potential of forming stabilizing secondary structures that might serve to bring the deletion endpoints together. Examination of the *wx* deletions revealed that the deleted sequences were rich in G and C residues when compared to the rest of the gene; 66–71% (Table 1) for the deleted DNA versus 56% for the rest of the gene. In addition, we determined that stable secondary structure is possible in sequences spanning most of the deleted region. We found stabilization free energies ranging from −45.8 to −353.7 kcal/mol (1 cal = 4.184 J). These values are significantly less than the −15 kcal/mol that is reported as the cut-off for stable secondary structure formation under physiological conditions (32). Finally, our calculations show that stem and loop structures can sequester at least 94% of the bases that comprise each deletion. This would bring the deletion endpoints in spatial proximity to one another.

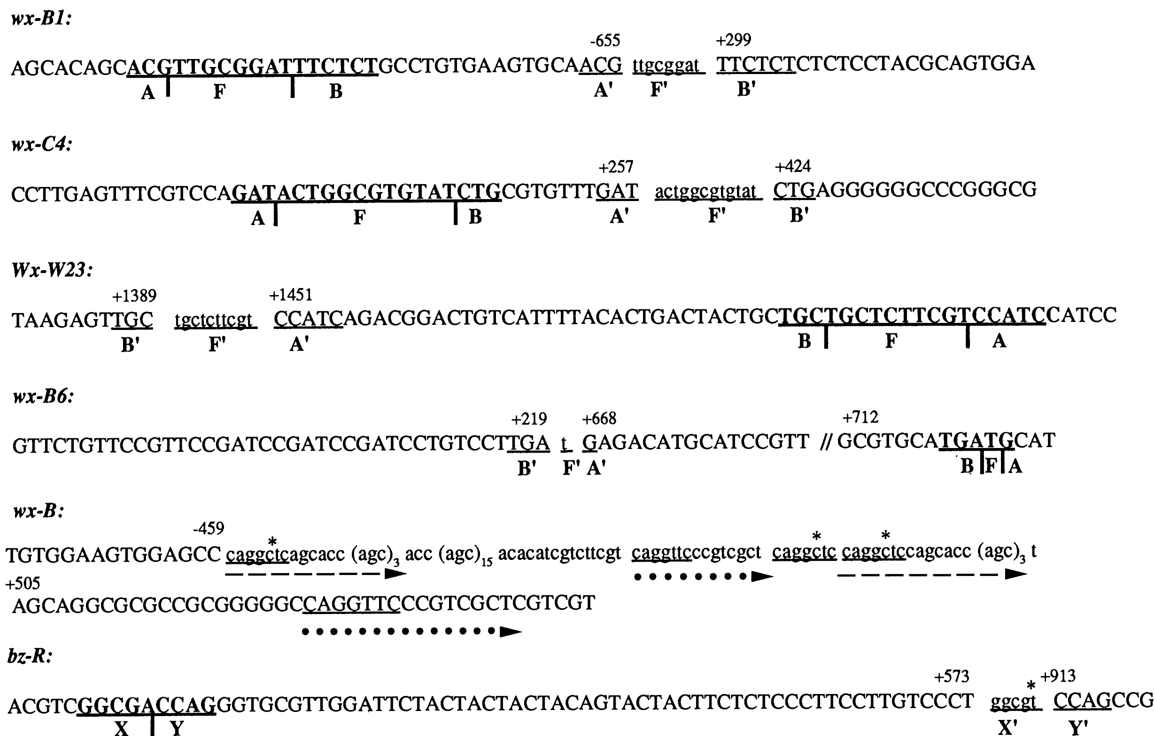

FIG. 2. DNA sequence of the deletion endpoints. The numerical positions of the endpoints relative to the start of *Wx* transcription (+1) are shown. Filler DNAs are in lowercase letters and are underlined along with sequences at the deletion endpoints that are similar to nearby sequences that are boldface and underlined. Asterisks denote sequence mismatches with homologous DNA near the deletion endpoints. The letters are referred to in the text and in Fig. 4. For *wx-B*, the solid lines represent a homologous sequence and the dotted and dashed arrows indicate direct repeats.

To determine whether potential secondary structures are common throughout the *Wx* gene or are restricted to the 1-kb region described above, we constructed a profile of the gene by folding a 500-nucleotide sequence window, beginning with position −900, and recording its stability. This window was then moved 250 bp downstream and the process was repeated until the end of the *Wx* gene. This analysis indicated that all of the large deletions occurred in a region capable of forming particularly stable secondary structures (Fig. 3). The mean free energies in this region are 150% that of the mean for the rest of the gene.

DISCUSSION

We have reported the first detailed molecular characterization of spontaneous deletions in a plant. Two aspects of our findings are of particular interest. First, we have shown that four of the deletions cluster in a G+C-rich region that is capable of forming stable secondary structures when single stranded. Similar clusters have been noted in other studies of spontaneous deletions in eukaryotes. For example, deletion endpoints in the hamster *aprt* locus have been shown to correlate with regions containing dyad symmetry and stabilizing stem and loop structures (4). It has been hypothesized that the presence of features such as dyad symmetries (4) and short direct repeats (1, 2, 4) or repetitive DNA (5–8) may increase the probability of recombinational events.

Whereas the clustering of deletions in both plant and animal genes may indicate similarities in certain aspects of deletion formation, the association of filler DNA with all of the *wx* deletions analyzed in this study may signify an important difference. Due to the paucity of spontaneous mutations characterized at the molecular level in higher plants, there are very few deletion endpoint sequences available for comparison with the *wx* deletions. However, another spontaneous mutation in maize, the *bz-R* allele, also contains a deletion with filler DNA that appears to be derived from sequences near one of the deletion endpoints (Fig. 2) (31). The existence of filler DNA among the mutations of two different maize genes argues that these types of deletions are

Table 1. Summary of maize waxy deletions

Allele	Location, bp	Size, bp	G+C content, %	$\Delta G°$, kcal/mol
wx-B	−458 to +521	980	69.4	−350
wx-B1	−654 to +298	953	67.2	−305
wx-B6	+220 to +667	448	71.5	−178
wx-C4	+258 to +423	166	66.1	−40
wx-W23	+1390 to +1450	59	44.0	0

$\Delta G°$, stabilization free energy.

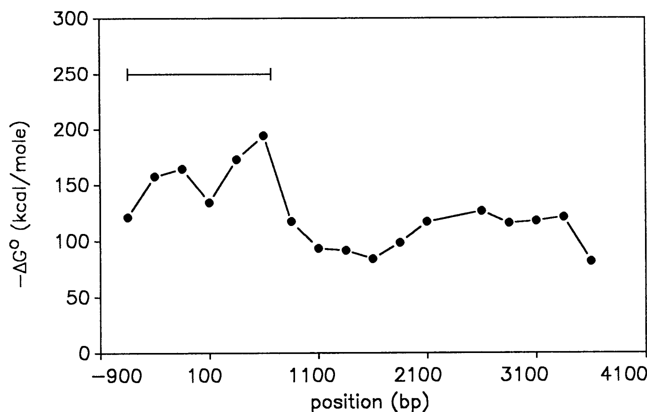

FIG. 3. Secondary structure stability across the *Wx* gene. The stabilization free energies ($\Delta G°$) of predicted secondary structures for 500-nucleotide sequence windows are shown across the gene. Bar delimits the region where the large deletions occur.

common in maize rather than reflecting some peculiarity of the *Wx* gene.

Filler DNA, though uncommon, has been associated with a few natural mutations in eukaryotes. Deletions of the *C. elegans* unc-54 gene (9), mouse β-globin gene (33), and human α-globin gene (5) also contain insertions of 18, 66, and 131 bp, respectively, and, in each case, the insertion is derived from sequences that are at least 600 bp from the deletion endpoints. A human β-globin deletion has more features in common with the maize deletions (34); it has a heptanucleotide insertion between the deletion endpoints that is identical to a sequence located 4 bp from one deletion terminus.

Filler DNA has also been encountered in mammalian somatic cells at the junctions of DNA rearrangements in both lymphoid (35, 36) and nonlymphoid (30, 37) cells, and at the site of integration of DNA transformed into mouse cells (38, 39). Immune system rearrangements contain filler DNA called N regions, which are generated by the untemplated addition of nucleotides to free DNA ends by terminal deoxynucleotidyltransferase (40). In contrast, the origin of filler DNA in cells without detectable terminal deoxynucleotidyltransferase, such as nonlymphoid cells, is unknown. For the larger insertions, it has been suggested that a pool of preformed oligonucleotides serves as a source of molecules that can be added to free DNA ends (30). In a study designed to understand DNA end joining and the origin of filler DNA in nonlymphoid cells, Roth *et al.* (37) transfected linear simian virus 40 genomes into cultured monkey cells and analyzed the DNA sequence at the recircularization joint. Of 200 junctions examined, 18 contained deletions of the terminal sequences and extra nucleotides at the site of recircularization. Whereas most additions involved a single base pair, 4 of the 18 contained filler DNA of 3–13 bp that was, in each case, derived from a sequence within 5 bp from the deletion endpoint (analogous to F and F′ in Fig. 2). More significantly, this nearby sequence was also homologous with the sequences at the 5′ and 3′ endpoints (analogous to A, B, A′, and B′ in Fig. 2). Roth *et al.* (37) called these insertions duplication junctions and proposed that slipped mispairing and DNA repair were involved in their formation.

The model of Roth *et al.* (37) explains the origins of duplication junctions that arise in simian virus 40 genomes transfected into mouse cells. The similarities between the filler DNA in *wx-B1*, *wx-C4*, and *Wx-W23* and duplication junctions led us to modify their model to explain our mutations. We propose that the maize deletions probably occurred during lagging strand DNA replication when discontinuous DNA synthesis provides free ends and long stretches of single-stranded DNA. For *wx-B1*, *wx-C4*, and *Wx-W23*, misalignment of newly replicated B′ with B (Fig. 4; see Fig. 2 for the identity of A, B, F, etc.) may be facilitated by higher-order structures that would serve to bring the direct repeats into spatial proximity (6, 41). For the *wx* deletions, the DNA between A′ and B′ is capable of assuming stable secondary structure when single stranded (Table 1 and Fig. 3). If this mispaired structure is resolved by DNA replication, as is shown in Fig. 4 (*Left*), the newly synthesized strand will be deleted for the DNA between the direct repeat B and B′ and for one copy of the repeat. This class of deletion is very common in both prokaryotes (1, 2) and eukaryotes (4, 42). To explain the origin of the filler DNA (F′), a second slip mispairing is postulated to occur (Fig. 4 *Right*); continued replication would result in the deletion of all DNA between A′ and B′ and the addition of F′. Although this model can account for the origin of the insertion junctions, it does not address the more important question of why a second slip mispairing is favored over the resolution of a single mispaired structure.

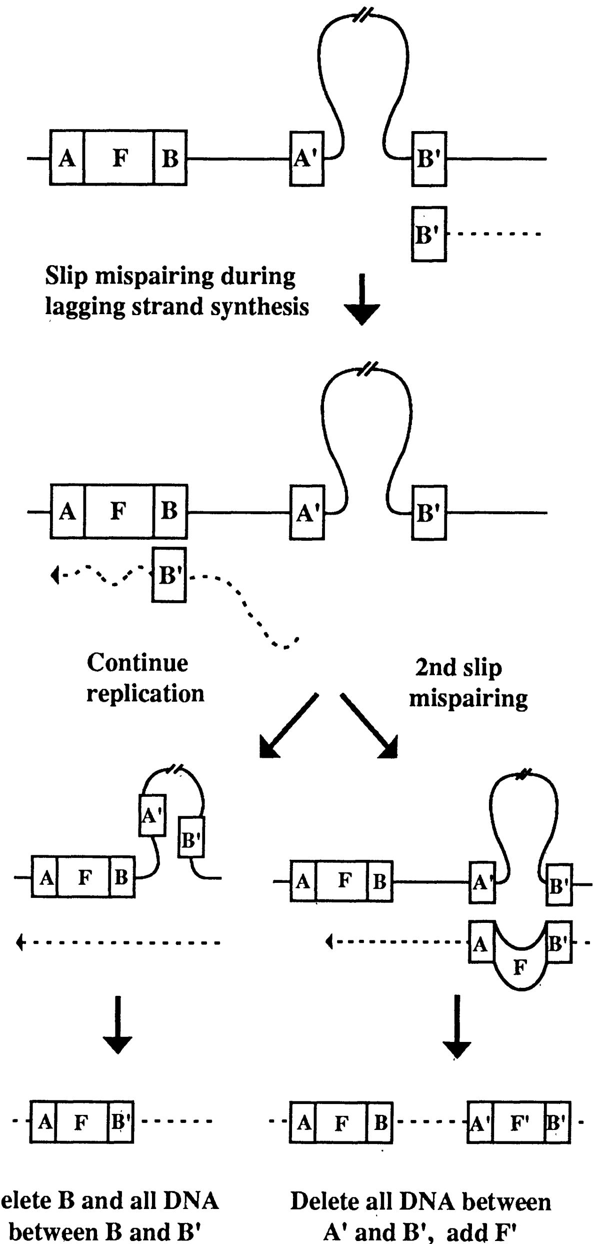

FIG. 4. Model to explain the origin of filler DNA during DNA replication. The DNA sequences are represented by the letters defined in Fig. 2. Dotted lines represent newly replicated DNA strands.

The *wx-B6* deletion is a poorer fit with the model proposed in Fig. 4 because F and A, as drawn, are only 1 bp each (Fig. 3). After slip mispairing of B′ with B, replication would be more likely to continue than to form a second slip mispaired structure. Similarly, for *bz-R*, slip mispairing of Y′ with Y can be envisioned; however, a second slip mispairing is unlikely because there is no homology at the other deletion endpoint. The filler DNA of *wx-B6* and *bz-R* may not be derived from these flanking sequences and may arise by the untemplated addition of a nucleotide (*wx-B6*) or a short oligonucleotide (*bz-R*) to a free DNA end.

The origin of the 131-bp insertion in the *wx-B* allele is much more complex than the other filler DNAs and cannot be explained by the model in Fig. 4 or by the untemplated addition of oligonucleotides. The termini of the insertion are composed of an imperfect 22-bp direct repeat (Fig. 2, dashed arrow). In addition, there is a second direct repeat of 15 bp; one copy is in the insert and the other is in flanking sequences

(dotted arrow). Finally, the sequence AGC is reiterated 21 times; this sequence is found at the 3′ endpoint and 1 bp away from the 5′ endpoint. There is no simple way to account for this complex structure. Perhaps this sequence is derived from an unstable intermediate structure. Alternatively, the presence of long stretches of G and C residues at the termini of the deleted DNA (data not shown) may have facilitated intrastrand base pairing and led to strand switching and polymerase stuttering during DNA replication.

In summary, we find that filler DNA is associated with five deletions in the maize *wx* gene. Although filler DNA is a common feature of all of these deletions, examination of the DNA sequence has led us to suggest that more than one mechanism is responsible for its insertion. In contrast with our findings in maize, filler DNA is rarely associated with germ-line mutations in animals. This apparent discrepancy may result from a fundamental difference in the nature of germ cells in plants and animals. Roth *et al.* (30) noted that filler DNA are frequently associated with mammalian somatic cell DNA rearrangements but are rare in bacteria and yeast. To explain this apparent discrepancy, they suggest that a certain degree of sloppiness may be acceptable in somatic cells because the evolutionary cost to the organism is relatively minor. Thus, DNA repair and replication processes that generate filler DNA may be unacceptable in germ cells or in organisms that are their own germ cells, such as bacteria and yeast. Plant germ cells differentiate from apical meristems from prolonged vegetative growth, unlike animal germ cells, which are set aside early in development and protected. Given this situation, it may not be possible to shield plant germ cells from the harmful effects that may accompany division and growth prior to their differentiation. Our results suggest that one by-product of a plant's life style may be a higher frequency of mutational events that are more reminiscent of somatic than germ cell rearrangements in animals.

We thank Drs. Cliff Weil, Glenn Galau, and Steve Ludwig for critical reading of the manuscript. This study was supported by National Institutes of Health Grant GM32528 to S.R.W.

1. Albertini, A. M., Hofer, M., Calos, M. P. & Miller, J. H. (1982) *Cell* **29**, 319–328.
2. Farabough, P. J., Schmeissner, U., Hofer, M. & Miller, J. H. (1978) *J. Mol. Biol.* **126**, 847–857.
3. Meuth, M. (1989) in *Mobile DNA*, eds. Berg, D. E. & Howe, M. M. (Am. Soc. Microbiol., Washington, DC), pp. 833–860.
4. Nalbantoglu, J., Hartley, D., Phaer, G., Tear, G. & Meuth, M. (1986) *EMBO J.* **5**, 1199–1204.
5. Nicholls, R. D., Fischel-Ghodsian, N. & Higgs, D. R. (1987) *Cell* **49**, 369–378.
6. Vanin, E. F., Henthorn, P. S., Kioussis, D., Grosveld, F. & Smithies, O. (1983) *Cell* **35**, 701–709.
7. Lehrman, M. A., Russell, D. W., Goldstein, J. L. & Brown, M. S. (1987) *J. Biol. Chem.* **272**, 3354–3361.
8. Lehrman, M. A., Schneider, W. L., Sudhof, T. C., Brown, M. S., Goldstein, J. L. & Russell, D. W. (1985) *Science* **227**, 140–146.
9. Pulak, R. A. & Anderson, P. (1988) *Mol. Cell. Biol.* **8**, 3748–3754.
10. Pohlman, R., Fedoroff, N. V. & Messing, J. (1984) *Cell* **37**, 635–643.
11. Masson, P., Surosky, R., Kingsbury, J. A. & Fedoroff, N. V. (1987) *Genetics* **177**, 117–137.
12. Dooner, H., English, J., Ralston, E. & Weck, E. (1986) *Science* **234**, 210–211.
13. Taylor, L. P. & Walbot, V. (1985) *EMBO J.* **4**, 369–376.
14. Sommer, H., Bonas, U. & Saedler, H. (1988) *Mol. Gen. Genet.* **211**, 49–55.
15. Roth, D. B. & Wilson, J. H. (1986) *Mol. Cell. Biol.* **6**, 4295–4304.
16. Nelson, O. E. & Rines, H. W. (1962) *Biochem. Biophys. Res. Commun.* **9**, 297–300.
17. Wessler, S. R. & Varagona, M. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 4177–4181.
18. Shure, M., Wessler, S. & Fedoroff, N. (1983) *Cell* **35**, 225–233.
19. Wessler, S. R., Baran, G., Varagona, M. & Dellaporta, S. L. (1986) *EMBO J.* **5**, 2427–2432.
20. Frischauf, A., Lehrach, H., Poustka, A. & Murray, N. (1983) *J. Mol. Biol.* **170**, 827–842.
21. Yanisch-Perron, C., Vieira, J. & Messing, J. (1985) *Gene* **33**, 103–109.
22. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
23. Vieira, J. & Messing, J. (1987) *Methods Enzymol.* **153**, 3–11.
24. Dale, R. M., McClure, B. A. & Houchins, J. P. (1985) *Plasmid* **13**, 31–40.
25. Kim, H.-S. & Smithies, O. (1988) *Nucleic Acids Res.* **16**, 8887–8903.
26. Klosgen, R. B., Gierl, A., Schwarz-Sommer, Zs. & Saedler, H. (1986) *Mol. Gen. Genet.* **203**, 237–244.
27. Zuker, M. & Stiegler, P. (1981) *Nucleic Acids Res.* **9**, 133–148.
28. Breslauer, K., Frank, R., Blocker, H. & Marky, L. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 3746–3750.
29. Freier, S., Kierzek, R., Jaeger, J., Sugimoto, N., Caruthers, M., Neilson, T. & Turner, D. H. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 9373–9377.
30. Roth, D. B., Chang, X.-B. & Wilson, J. H. (1989) *Mol. Cell. Biol.* **9**, 3049–3057.
31. Ralston, E. J., English, J. J. & Dooner, H. K. (1988) *Genetics* **119**, 185–197.
32. Streisinger, G., Ikada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, E. & Inouye, M. (1966) *Cold Spring Harbor Symp. Quant. Biol.* **31**, 77–84.
33. Goldberg, S. Z., Kuebbing, D., Trauber, D., Schafer, M. P., Lewis, S. E., Popp, R. A. & Anderson, W. F. (1986) *J. Biol. Chem.* **261**, 12368–12374.
34. Spritz, R. A. & Orkin, S. H. (1982) *Nucleic Acids Res.* **10**, 8025–8029.
35. Concannon, P., Pickering, L. A., Kung, P. & Hood, L. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 6598–6602.
36. Alt, F. W. & Baltimore, D. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 4118–4122.
37. Roth, D. B., Porter, T. N. & Wilson J. H. (1985) *Mol. Cell. Biol.* **5**, 2599–2607.
38. Anderson, R. A., Kato, S. & Camerini-Otero, R. D. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 206–210.
39. Williams, T. J. & Fried, M. (1986) *Mol. Cell. Biol.* **6**, 2179–2184.
40. Landau, N. R., Schatz, D. G., Rosa, M. & Baltimore, D. (1987) *Mol. Cell. Biol.* **7**, 3237–3243.
41. Glickman, B. W. & Ripley, L. S. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 512–516.
42. Efstratiadis, A., Posakony, J. W., Maniatis, T., Lawn, R. M., O'Connell, C., Sprizt, R. A., Deriel, J. K., Forget, B. G., Weissman, S. M., Slightom, J. L., Blechl, A. E., Smithies, O., Baralle, F. E., Shoulders, C. C. & Proudfoot, N. J. (1980) *Cell* **21**, 653–668.