Identification of the Genomic Locations of Duplicate Nucleotide Sequences in Maize by Analysis of Restriction Fragment Length Polymorphisms

Tim Helentjaris,* David Weber† and Scott Wright*

*Molecular Biology Group, NPI, Salt Lake City, Utah 84108, and †Genetics Group, Department of Biological Sciences, Illinois State University, Normal, Illinois 61761

Manuscript received February 28, 1987
Revised copy accepted November 2, 1987

ABSTRACT

While preparing a linkage map for maize based upon loci detected through the use of restriction fragment length polymorphisms (RFLPs), it was found that 62 of the 217 cloned maize sequences tested (29%) detected more than one fragment on genomic Southern blots. Thus, more than one nucleotide sequence is present within the maize genome which is in part homologous to each of these cloned sequences. The genomic locations of these “duplicate” sequences were determined and it was found that they usually originated from different chromosomes. The process which produced them did not operate randomly as some pairs of chromosomes share many duplicate sequences while many other pairs share none. Furthermore, these shared duplicate sequences are generally arrayed in an ordered arrangement along these chromosomes. It is believed that chromosomal segments which contain several duplicate loci in a generally ordered arrangement must have had a common origin. The presence of these duplicated segments supports the idea that allopolyploidy may have been involved in the evolution of maize. Nevertheless, the duplicate loci do not primarily involve five pairs of chromosomes and thus, five pairs of homeologous chromosomes are not currently present within the maize genome. The data clearly indicate that maize is not a recent allotetraploid produced by hybridization between two individuals with similar genomic structures; however, the data are also consistent with the possibility of these shared duplicate chromosomal segments having been generated through internal duplication.

ALTHOUGH maize is generally considered to be a true diploid, it has been repeatedly suggested that it originated as an amphidiploid hybrid and has an allopolyploid genomic structure. ANDERSON (1945) noted that certain genera related to corn contain 5 chromosomes in their haploid genomes, and he suggested that corn originated as an allotetraploid hybrid of “something like a 5-chromosomed Sorghum crossed with some 5-chromosomed relative of Coix.” The reported experimental hybridization of Coix and maize (HARADA et al. 1954; NOWACKI, ANIOU and BEIBER 1972) has provided support for this possibility.

Numerous cytological observations, such as the pairing between nonhomologous chromosomes in monoploid maize (TING 1966) and formation of bivalents during diakinesis-metaphase I (MOLINA and NARANJO 1986) have also been interpreted to demonstrate the presence of large segments shared in common between chromosomes. RHOADES (1951) summarized genetic inheritance data that he felt indicated that the maize genome contained areas of duplicated sequences. More recently, “duplicated” isozymic loci have been found in maize in parallel linkages (GOODMAN et al. 1980; McMILLIN and SCANDALIOS 1980; GOODMAN and STUBER 1983; WENDEL, GOODMAN and STUBER 1985), and this also appears to demonstrate the presence of duplicated sequences and shared chromosomal segments within the maize genome.

One difficulty with each of these observations, is that they only suggest that duplicated sequences are present within the maize genome. None of these studies has actually demonstrated that common DNA sequences exist at different chromosomal locations within the maize genome. A procedure, restriction fragment length polymorphism (RFLP) analysis, has recently been developed which allows the direct detection and mapping of a nucleotide sequence within a genome. We previously applied this procedure to maize and demonstrated that loci can be detected which hybridize to a single cloned DNA sequence, and that the relative genomic locations of these loci can be determined through both linkage (HELENTJARIS et al. 1985) and monosomic (HELENTJARIS, WEBER and WRIGHT 1986) analyses. This report describes the identification and mapping of a number of duplicated DNA sequences within the maize genome by this process, and provides direct evidence as to the structure and origin of the maize genome.

MATERIALS AND METHODS

Sources of cloned sequences: Four sources of cloned sequences were used in this study. First, a cDNA library
was constructed from leaf mRNAs and used as described previously (Helentjaris et al. 1985). Seventy-nine of the 284 RFLP loci on our RFLP map were detected with these cDNA clones. Second, HindIII-generated genomic DNA fragments were sized, cloned into pUC8, and then selected for homology to sequences present in low copy number in the maize genome (Helentjaris et al. 1986). Seventy-two of the RFLP loci were detected by clones of this type. Third, a second library of genomic sequences was constructed as above except that the methylation-sensitive restriction enzyme, PstI, was used to digest the DNA. For the most part, this type of library yields a much higher percentage of unique sequence clones and abolishes the need for evaluating the probes and discarding those homologous to repetitive sequences in the maize genome. One hundred-fourteen of the RFLP loci were detected using this library. Essentially no significant differences were noted in either the distribution of these three types of clones within the maize genome or their level of duplication. The last source of probe sequences were clones of known identity supplied to us by other investigators (listed in Wright, Helentjaris and Kikuchi 1987).

Because many of the cloned sequences hybridized to more than one genomic fragment, extensive efforts were made to confirm that these results were not due to some type of cloning artifact. We checked our cDNA clones for the presence of internal PstI sites as evidence of fortuitously ligated sequences. Only a few internal sites were found, and the number was statistically consistent with the premise that they represented naturally-occurring internal PstI sites. Our genomic DNA clones contained inserts which had been size-selected, and we later checked individual clones both for the final insert size and the presence of internal HindIII or PstI sites. No results were obtained that were inconsistent with the hypothesis that each of these clones was the result of an independent cloning event.

**Plant materials:** Populations used to map cloned sequences were as follows. Forty-seven F2 progeny from inbreds H427 × 761 and 196 F2 progeny from inbreds Tx303 × Co159 were used in linkage studies and analyzed by maximum likelihood as reported earlier (Helentjaris et al. 1986). Forty-nine progeny from a backcross of (Mangelsdorf's Multiple Tester × W22) × Mangelsdorf's Multiple Tester were also used to map cloned sequences that were uninformative in the above two F2 populations. Finally, monosomic analysis was carried out as described previously (Helentjaris, Weber and Wright 1986).

**Other methods:** All other methods for the preparation of plant DNA, Southern blots, and hybridization (aqueous with 5 × SSC at 60°C and wash (0.25 × SSC at 60°C) conditions were as described previously (Helentjaris et al. 1985, 1986).

**RESULTS**

While preparing genetic linkage maps for maize based upon loci detected by RFLPs (Helentjaris et al. 1986; Helentjaris, Wright and Weber 1986), we found that many of the maize clones detected more than one fragment on genomic Southern blots. This was despite the fact that efforts were made to utilize clones that hybridized only to unique or low-copy sequences in the maize genome. Multiple fragments were observed nowhere which restriction enzyme was used to digest the genomic DNA and this suggested that duplicated sequences were being detected in the maize genome. Of the 217 informative maize clones generated and tested to date, 62 (28.6%) appeared to detect duplicate sequences.

One important point to be considered here is that while one can be fairly certain about the relatedness of loci detected by this method, one must be more careful about drawing conclusions about loci which this study has not detected as being duplicated. For instance, the Css (sucrose synthetase-2) locus shares considerable sequence homology with the sh1 (sucrose synthetase-1) locus (McCarty, Shaw and Hannah 1986); however, under our hybridization and wash conditions, the Css probe does not hybridize with the sh1 locus and the sh1 probe does not hybridize with the Css locus. The alcohol dehydrogenase-1 and -2 loci are also known to have considerable sequence homology (Dennis et al. 1985); however, we cannot detect both duplicated loci with either clone alone using our hybridization and wash conditions, which are fairly stringent in order to simplify our analysis of hybridization patterns for mapping purposes. If less stringent conditions were used, some of the probes which detected only one fragment might detect additional fragments.

The chromosomal locations of the sequences in the maize genome which hybridized with each cloned sequence were determined by genetic linkage analysis (Helentjaris et al. 1986) and/or by monosomic analysis (Helentjaris, Weber and Wright 1986). An example of a monosomic analysis utilizing clone GID-06 used as a probe on a Southern blot is shown in Figure 1. This clone hybridizes with two DNA fragments in each of the parents (85-MT and 85/r-X1). In both parents (which are highly inbred), the hybridization signal for one fragment is more intense than for the other. Both diploid and monosomic progeny were recovered among the F1 progeny of a cross between these two lines, and as expected, all four fragments are found in confirmed diploid F1 plants. The monosomics produced by this cross contain all 10 chromosomes from the male (85-MT) parent but only 9 of the 10 chromosomes from the female parent. One of the two fragments (210B)

![Figure 1.-Monosomic analysis of a clone demonstrating duplicate loci. A clone, GID-06, was used to probe a Southern blot containing genomic DNAs prepared from two inbred lines, 85-MT and 85/r-X1, a diploid F1, and F1 individuals monosomic for specific maize chromosomes. Numbers above the monosomic individuals denote the missing chromosome in each case. The lane labeled "markers" contains molecular weight markers.](http://example.com/figure1.jpg)
Duplicated RFLP Loci in Maize

from the female parent is missing in monosomic 2 plants while the other fragment (59B) from the female parent is missing in monosomic 7 plants. Our initial interpretation of these results was that the two intense fragments observed in the two inbreds represented alleles at one genetic locus on chromosome 2 (210A and 210B) and the two less intense fragments represented alleles at a second locus on chromosome 7 (59A and 59B). We believe the observed difference in signal intensity of the two loci reflects differing rates of homology to the cloned sequence rather than differing copy numbers in the genome or some other explanation. Our assumption of two duplicate loci was later confirmed by linkage analysis, and the two duplicate loci were placed on our maize RFLP linkage map; locus 59 is located on chromosome 7 and 210 is located on chromosome 2 (Helentjaris et al. 1986).

Therefore, two distinct loci are detected by significant homology to this single cloned sequence. We have identified a number of such “duplicate” sequences which are homologous to specific probes and have placed these loci on our RFLP linkage map using both genetic linkage and monosomic analyses. Our current map is shown in Figure 2. Linkage data are available for most of the loci, and the map positions of these loci are indicated by their numerical designations above vertical marks placed along the horizontal axes representing the chromosomes. Other loci without linkage data, but still assigned to specific chromosomes, are placed along the right side of the map. The linkage groups derived from these analyses have been correlated with the conventional maize linkage map in five ways: (1) through monosomic analyses (Helentjaris, Weber and Wright 1986), (2) through linkage analysis with morphological marker loci (Helentjaris, Wright and Weber 1986), (3) through linkage analysis with isozyme loci (unpublished work with C. Stubber and M. Edwards), (4) through RFLP analysis of clones for genes with known genomic locations (Wright, Helentjaris and Kikuchi 1987), and (5) through the use of B-A translocations (which has allowed us to determine the approximate locations of centromeres, manuscript in preparation). We have denoted all loci in this figure (in bold and italics) which were detected by RFLP analyses as duplicate sequences in the maize genome.

An analysis of the distribution of these duplicate loci is presented in Table 1. The number of RFLP loci per chromosome and the number of duplicated RFLP loci per chromosome both differ significantly from a uniform distribution at $P < 0.001$ and $P < 0.05$, respectively, using a goodness of fit $\chi^2$ test. However, neither the number of RFLP loci nor the number of duplicated RFLP loci per unit of chromosome physical length (Neuffer, Jones and Zuber 1968) is significantly different from a uniform distribution at $P < 0.05$, demonstrating that when adjusted for chromosome length, the density of both unique and duplicated sequences is fairly uniform throughout the maize genome. Further, the percentages of duplicated RFLP loci per chromosome are not significantly different from each other at $P < 0.05$ using a contingency $\chi^2$ test.

The distribution of these duplicate loci shared among nonhomologous chromosomes is shown in Table 2. The nonrandom arrangement of these duplicated loci is evident with certain chromosome pairs sharing many loci and other chromosome pairs having none. For example, chromosomes 2 and 7 share 13 pairs of duplicate loci and chromosomes 3 and 8 share 10, while no pairs of duplicate loci have been found to date by this method for 16 of the 45 pairwise combinations of nonhomologous chromosomes. We detected only three intrachromosomal duplications, where the duplicate loci are both on the same chromosome at different locations. For example, loci 63 and 102 are detected by the same clone, and are 9 map units apart on chromosome 6. Thus, intrachromosomal duplications are infrequent in the maize genome; however, it should be pointed out that very closely linked duplicate loci would often be overlooked by this method of analysis. Although two hybridizing fragments might be detected, the two duplicate loci would usually cosegregate and segregant individuals might not be detected unless many progeny were analyzed.

Clearly, the mechanism that produced these pairs of duplicate loci must have operated in a nonrandom manner. This becomes especially evident when one examines the map positions of these duplicate loci along these chromosomal pairs. The map positions of pairs of duplicate loci on chromosomes 2 and 7 are presented in Figure 3. Only pairs of duplicate loci shared by chromosomes 2 and 7 are indicated, and those detected by the same cloned sequences are connected with a line. Clearly, the 13 pairs of duplicate loci shared by these chromosomes are arranged in the same approximate order within the two chromosomes. This is strong evidence that the chromosomal segments containing these sequences, well over 50 map units in both chromosomes, possess extensive regions of homology besides the small cloned sequences tested, and almost certainly had a common origin.

Most of the duplicated loci on chromosome 2 which are shared with chromosome 7 are located on the long arm of this chromosome. This largely homologous segment does not appear to extend significantly into the short arm of chromosome 2. With one exception (locus 367), the loci on the short arm of 2 examined so far are either not duplicated or are duplicated on chromosomes other than 7. Three loci (11, 294, and Ssu2) found in somewhat close proximity on 2 are also duplicated on chromosome 4.
FIGURE 2.—Our current maize RFLP linkage map. Horizontal lines denote the chromosomes and RFLP loci are represented above the lines by numerical designations. Where the exact location of either single or multiple loci is unknown, this ambiguity is denoted as an underlined segment. Loci along the right side of the figure have been assigned to chromosomes but linkage data are not available. Loci in brackets were detected as morphological markers or isozymes, while all other loci were detected by RFLP analysis. Those RFLP loci shown in bold and italics possess duplicated sequences elsewhere in the genome. Numbers below the horizontal lines are distances between loci in map units. The approximate locations of the centromeres are denoted by "-0-".
Duplicated RFLP Loci in Maize

TABLE 1

Distribution of unique and duplicated loci in maize

<table>
<thead>
<tr>
<th>Chromosome No.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. RFLP loci**</td>
<td>47</td>
<td>43</td>
<td>30</td>
<td>28</td>
<td>23</td>
<td>20</td>
<td>30</td>
<td>33</td>
<td>23</td>
<td>13</td>
</tr>
<tr>
<td>No. duplicated RFLP loci*</td>
<td>17</td>
<td>24</td>
<td>13</td>
<td>10</td>
<td>9</td>
<td>12</td>
<td>15</td>
<td>15</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>Percent of RFLP loci duplicated</td>
<td>36.1</td>
<td>55.8</td>
<td>43.3</td>
<td>35.7</td>
<td>39.1</td>
<td>60.0</td>
<td>50.0</td>
<td>45.4</td>
<td>43.5</td>
<td>30.7</td>
</tr>
<tr>
<td>Relative length of chromosome</td>
<td>229</td>
<td>196</td>
<td>179</td>
<td>175</td>
<td>175.122</td>
<td>140</td>
<td>140</td>
<td>122</td>
<td>122</td>
<td>100</td>
</tr>
<tr>
<td>No. of RFLP loci/relative length</td>
<td>0.205</td>
<td>0.219</td>
<td>0.168</td>
<td>0.160</td>
<td>0.131</td>
<td>0.164</td>
<td>0.214</td>
<td>0.236</td>
<td>0.189</td>
<td>0.130</td>
</tr>
<tr>
<td>No. duplicated RFLP loci/relative length*</td>
<td>0.074</td>
<td>0.122</td>
<td>0.073</td>
<td>0.057</td>
<td>0.051</td>
<td>0.098</td>
<td>0.107</td>
<td>0.107</td>
<td>0.082</td>
<td>0.040</td>
</tr>
</tbody>
</table>

* Not significantly different from a uniform distribution adjusted for chromosome length.
* Values from Neuffer, Jones and Zuber (1968).
** Significantly different from a uniform distribution at \( P < 0.05 \).

** Significantly different from a uniform distribution at \( P < 0.001 \).

TABLE 2

Distribution of shared duplicate loci among maize chromosomes

<table>
<thead>
<tr>
<th>1-2</th>
<th>225-98</th>
<th>2-10</th>
<th>359-361</th>
</tr>
</thead>
<tbody>
<tr>
<td>82-32</td>
<td>404-103</td>
<td>254-350</td>
<td></td>
</tr>
<tr>
<td>109-273</td>
<td>428-427</td>
<td>320-321</td>
<td>4-7</td>
</tr>
<tr>
<td>243-242</td>
<td>365-222</td>
<td>95-316</td>
<td></td>
</tr>
<tr>
<td>243-244</td>
<td>2-2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>246-340</td>
<td>243-244</td>
<td>5-5</td>
<td></td>
</tr>
<tr>
<td>246-340</td>
<td>3-5</td>
<td>4-9</td>
<td></td>
</tr>
<tr>
<td>1-3</td>
<td>2-3</td>
<td>249-341</td>
<td>4-10</td>
</tr>
<tr>
<td>109-358</td>
<td>273-358</td>
<td>3-6</td>
<td>208-366</td>
</tr>
<tr>
<td>1-4</td>
<td>2-4</td>
<td>Me1-Me2</td>
<td></td>
</tr>
<tr>
<td>Adh1-Adh2</td>
<td>11-317</td>
<td>5-6</td>
<td></td>
</tr>
<tr>
<td>96-95</td>
<td>294-359</td>
<td>3-7</td>
<td>53-302</td>
</tr>
<tr>
<td>84-36</td>
<td>Ssu2-Ssu1</td>
<td>358-355</td>
<td></td>
</tr>
<tr>
<td>1-5</td>
<td>2-5</td>
<td>3-8</td>
<td>Tpi4-Tpi3</td>
</tr>
<tr>
<td>226-75</td>
<td>320-115</td>
<td>52-328</td>
<td></td>
</tr>
<tr>
<td>Phy1-Phy2</td>
<td>70-69</td>
<td>6-6</td>
<td></td>
</tr>
<tr>
<td>357-360</td>
<td>2-6</td>
<td>108-204</td>
<td>63-102</td>
</tr>
<tr>
<td>357-360</td>
<td>1-3</td>
<td>294-361</td>
<td>201-3</td>
</tr>
<tr>
<td>1-6</td>
<td>340-245</td>
<td>202-114</td>
<td>6-8</td>
</tr>
<tr>
<td>205-63</td>
<td>219-218</td>
<td>2-1</td>
<td></td>
</tr>
<tr>
<td>246-245</td>
<td>2-7</td>
<td>338-339</td>
<td>9-206</td>
</tr>
<tr>
<td>205-102</td>
<td>4-5</td>
<td>364-276</td>
<td>63-301</td>
</tr>
<tr>
<td>46-35</td>
<td>425-426</td>
<td>101-103</td>
<td></td>
</tr>
<tr>
<td>1-7</td>
<td>49-44</td>
<td>Tpi4-Tpi3</td>
<td>102-301</td>
</tr>
<tr>
<td>96-316</td>
<td>118-113</td>
<td>Pdk1-Pdk2</td>
<td></td>
</tr>
<tr>
<td>109-335</td>
<td>123-47</td>
<td>3-9</td>
<td></td>
</tr>
<tr>
<td>Sod4-Sod2</td>
<td>210-59</td>
<td>89-14</td>
<td>6-9</td>
</tr>
<tr>
<td>1-8</td>
<td>221-216</td>
<td>418-416</td>
<td></td>
</tr>
<tr>
<td>205-301</td>
<td>337-45</td>
<td>203-116</td>
<td>9-9</td>
</tr>
<tr>
<td>352-353</td>
<td>292-363</td>
<td>Shi-Csi1</td>
<td></td>
</tr>
<tr>
<td>1-9</td>
<td>367-30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>93-299</td>
<td>405-277</td>
<td>4-6</td>
<td>9-10</td>
</tr>
<tr>
<td>99-97</td>
<td>A.lr-Alr2</td>
<td>6-7</td>
<td>25-366</td>
</tr>
</tbody>
</table>

Duplicate loci detected by single cloned sequences and shared between chromosomes are listed in order of the chromosome pairs. Chromosome pairs with no shared duplicate loci are not listed.

while three other loci (320, 365, and 254) are similarly duplicated on chromosome 10. Thus, at least small portions of these latter chromosomes also may share a common origin with segments of chromosome 2.

The centromere on chromosome 7 is not precisely located on our RFLP map, and although most of these loci are known to be in the long arm, it is uncertain how many if any of these duplicate loci region are located in the short arm of this chromosome.

One other interesting observation is that our most distal RFLP locus on the long arm of chromosome 2, locus 32, is duplicated on chromosome 1 at locus 82. This does not seem to fit the overall pattern that most of the long arm of 2 shares homology with chromosome 7. We have observed that this locus, 32, is deleted in some maize inbred lines, while all of the remaining proximal loci on the long arm of 2 are still present. Perhaps this represents a transposition from chromosome 1 onto 2 which is a more recent event in the evolution of maize and may also be correlated with the observation that there are two cytological forms of chromosome 2 in maize (Neuffer, Jones and Zuber 1968).

Not all of the duplicated segments shared by pairs of chromosomes exhibit an order as consistent as that found in chromosomes 2 and 7. Chromosomes 3, 6 and 8 also share several duplicate loci; however, the order is considerably more complex. Figure 4 shows that the ten pairs of duplicate loci found on chromosomes 3 and 8 can be grouped into two clusters by their relative genomic locations. Five of the loci are situated in the long arm of chromosome 3 and the corresponding duplicate loci lie in the long arm of chromosome 8. The other five loci are located primarily in the short arm of chromosome 3 near the centromere and share duplicate loci at the left end of our chromosome 8 linkage group (near in not in the short arm). The exact relation of these sequences to the centromere position on chromosome 8 is unclear because the centromere on chromosome 8 has not been precisely located. In addition to the 10 loci shared with chromosome 3, chromosome 8 also possesses six duplicate loci with counterparts on chromosome 6. These are found in one cluster toward the center of the chromosome 8 linkage group,
surrounded by the two 3-8 clusters with one overlap. Interestingly, the duplicate loci on chromosome 6 are widely spaced, but at least two pairs, 1-2 and 103-101, are located extremely close together on both chromosomes. Only one pair of duplicate loci was found on chromosomes 3 and 6, Me1 and Me2 (NADP malic enzyme).

While the chromosome pair, 2 and 7, and the group involving chromosomes 3, 6 and 8, contain the largest number of pairs of duplicate loci detected in this study, other regions of homology were also detected. Five loci (97, 98, 209, 403 and 427) located in a relatively short segment of the long arm of chromosome 9 are duplicated on chromosome 1; however, the corresponding loci on chromosome 1 are widely separated into two distinct groups. Three pairs of closely linked duplicate loci are also located on chromosomes 1 and 5. Most of the other duplicated RFLP loci are found as smaller blocks or isolated cases.

DISCUSSION

Numerous data of several types have been previously obtained that suggest that the maize genome contains at least some duplicated sequences and possibly much larger segments shared between chromosomes. Since Anderson's (1945) supposition that maize originated as an amphiploid, based upon chromosome numbers of five and ten for maize and some of its relatives, other indirect evidence such as the experimental hybridization of Coix and maize (Harada et al. 1954; Nowacki, Aniou and Beiber 1972) has been interpreted to support this theory. While suggestive, these types of results may just reflect similar evolutionary backgrounds among some of the Gramineae.
Pairing between nonhomologous chromosomes lacking homologs has also been considered as evidence for large scale interchromosomal homology in maize. Ting (1966) analyzed diakinesis in monoploid maize and reported that 50.9% of the cells contained one or more bivalents or multivalents. He indicated that this was "convincing evidence that modern maize is an allotetraploid with perhaps a base number of 5." However, Ford (1952), Snipe (1967), and Weber and Alexander (1972) observed lower frequencies of pairs (29.6, 15.3 and 15.4%, respectively) at diakinesis in maize monoploids.

Molina and Naranjo (1986) analyzed diploid maize plants and reported that 10 bivalents were regularly formed; however, "secondary associations were observed and a maximum of five groups of two bivalents was frequent at diakinesis-metaphase I." Such associations of bivalents have not been observed in diploid maize analyzed by one of us (D. F. Weber, unpublished results). Molina (1982) examined Zea perennis (2n = 40) and reported that 10 bivalents and 5 quadravalents were present in 55% of the cells. They also examined F1 hybrids involving Zea mays, Zea perennis and Zea diploperennis, and from observations made on meiotic cells in these F1 hybrids, they concluded that there was "strong evidence in favor of n = 5 for the genus Zea" (Molina and Naranjo 1986).

Exchanges between unpaired nonhomologous maize chromosomes have also identified the locations of presumptive redundant sequences within the maize genome. Weber and Alexander (1972) recovered 22 reciprocal translocations from monoploid diploid crosses in maize, and they found that twelve of the translocations involved chromosomes 6 and 7 and two involved chromosomes 2 and 6 at cytologically indistinguishable breakpoints. Because these translocations were found repeatedly, the breakpoints of these translocations (and probably the other eight translocations recovered from these crosses) presumably identify sites of interchromosomal homology.

The most specific prediction regarding the organization of the maize genome has been made by Bennett (1983). He attempted to predict the mean spatial order of the 10 haploid chromosomes in maize and noted that his model suggested that maize is an allotetraploid. He predicted that chromosomes 1 and 5, 2 and 4, 3 and 9, 6 and 7, and 8 and 10 were putative homoeologs in the maize genome. Our current study has mapped 81 pairs of duplicate RFLP loci among the 45 possible nonhomologous chromosome pairs (1.8 per pair) while Bennett's predicted homoeologous chromosome pairs had seven loci (1.4 per pair). Thus, his predicted homoeologous chromosome pairs did not contain more of these than would be expected by chance.

Correlation of our observations with the cytological data and evidence for alloploidy is confusing. This is due somewhat to the fact that although many of these studies predict large duplicated areas within the maize genome, these studies taken together with our results rarely predict the same areas. For instance in Weber and Alexander (1972), twelve 6-7 translocations and two 2-6 translocations were recovered. However, none of the duplicate loci identified in the current study involved chromosomes 6 and 7, and only two involved chromosomes 2 and 6. Also, they did not recover any 2-7, 3-8, or 6-8 translocations in their study, while these were the most common duplications found in our study. In their type of study, recombination may have taken place primarily between highly homologous, but relatively small regions in nonhomologous chromosomes, regions that were not detected by the limited number of probes utilized in the current study. Alternatively, the translocations recovered in their study may have been produced by recombination between segments in two chromosomes which each contained the same sequence repeated many times. It is possible that the same type of small but highly repeated sequences is responsible for certain cytological observations reported earlier but they do not correspond to the areas of duplication we have detected. Sequences repeated numerous times in restricted regions of nonhomologous chromosomes are known to exist in the knobs of maize chromosomes (Peacock et al. 1981) and in heterochromatin. We would not detect RFLP loci in such regions unless they were interspersed with low copy number sequences, because we are only able to utilize clones that are present in very low copy number. Perhaps this might also explain why some regions of our RFLP map are not well populated with RFLP loci, as these areas may primarily contain sequences present in the genome in high copy number and few, if any, sequences in low copy number. For instance, our RFLP map currently has only one RFLP locus on the short arm of chromosome 10.

There are other lines of evidence which point to at least very small areas of duplication in the maize genome but that may also support the idea of larger areas of shared chromosomal segments. Rhoades (1951) pointed out that at least 14 cases of duplicate, two of triplicate, and one of quadruplicate factor inheritance were known at that time. All of the alleles had to be in the recessive condition for the mutant phenotype to be expressed, a fact suggestive of duplicate loci. While certainly suggestive, these types of data do not demonstrate that these duplicate loci specify related gene products. Rhoades concluded "That the architecture of the germ plasm of maize contains duplicated regions can hardly be doubted, but whether or not they represent vestiges reflecting
an ancient amphidiploid origin or represent later occurring duplications cannot be decided at this time."

RHOADES (1951, 1955) also pointed out that the long arm of chromosome 9, the short arm of 5, the distal region of the long arm of 4, and 8 contained few known mutants and might contain redundant regions. The numerous duplicated loci we have detected on chromosome 8, and its extensive homology with chromosomes 3 and 6 are in agreement with his prediction concerning chromosome 8. The duplicate RFLP and isozyme loci on chromosomes 1 and 5 is in agreement with his prediction regarding chromosome 5. Also, the current conventional maize genetics map contains few known mutants in the long arm of chromosome 2, and this is consistent with our observation that the long arm of chromosome 2 is largely homologous with portions of chromosome 7. RHOADES (1951) also noted that many of the duplicate loci then known were on chromosome 9, including duplicate loci in both the long and short arm. We have found ten duplicated loci on chromosome 9, and the duplicate loci of five of these are on chromosome 7. Another pair of duplicate loci, sucrose synthetase-1 (sh1) and sucrose synthetase-2 (Css), are both located on chromosome 9, with sh1 in the short arm and Css in the long arm (McCARTY et al. 1986).

More recently, "duplicate" maize isozymic loci (isozymes with common enzymatic activities) have been identified. Several maize isozymes, including alcohol dehydrogenase (Adh), catalase (Cat), isocitrate dehydrogenase (Idh), and malate dehydrogenase (Mdh), have been shown to be produced by two or more loci, and many of these have been mapped within the maize genome (MCMLIN and SCANALDOS 1980; GOODMAN and STUBER 1983; WENDEL, GOODMAN and STUBER 1985). The presence of duplicated isozymes have been interpreted as providing strong evidence for internal duplication and possibly polyploidy in a number of species (GOTTLEIB 1982). Several examples of parallel linkages of duplicate isozyme loci have been reported in maize (GOODMAN et al. 1980; WENDEL et al. 1986). For example, Idh2 and Mdh2 are located on chromosome 6 and Idh1 and Mdh1 are both located on chromosome 8. Mdh4, Amp1, Amp2, Pgm1 and Cat2 are all located on chromosome 1, while Mdh5, Amp3, Pgm2 and Cat1 are found on chromosome 5. These parallel linkages are certainly suggestive of common derivations for segments of these pairs of chromosomes.

Our current RFLP results are also very consistent with information on these "duplicate" isozymic loci. Because duplicate isozymic loci produce enzymes with the same catalytic activity, it has been assumed that the templates of these loci may possess similar nucleotide sequences. However, this assumption in maize has only been shown to be true for one pair of duplicate isozymic loci, the Adh1 and Adh2 loci (DENNIS et al. 1985). Previously, chromosomes 3 and 6, 6 and 8, and 1 and 5 were all reported to contain duplicate and linked pairs of isozymic loci. In the current study, we have found that there indeed are homologous DNA sequences that are shared among chromosomes 3, 6 and 8, and 1 and 5 and they are clustered into distinct regions. The Idh1 and Idh2 (on chromosomes 8 and 6, respectively); Mdh1, Mdh2, and Mdh3 (on chromosomes 8, 6 and 3); Tpiβ and Tpiα (on chromosomes 8 and 3); and Hex1 and Hex2 (on chromosomes 3 and 6) duplicated isozyme loci are located on three chromosomes where we have detected significant duplicated segments (GOODMAN et al. 1980; WENDEL et al. 1986). A number of 1–5 duplicated RFLP loci are also closely located to regions where duplicate isozyme loci had previously been mapped. Thus, our current data are very consistent with data from duplicate isozymic loci, and this provides very strong support for the concept that duplicate isozymic loci often do indeed identify regions of similar nucleotide sequences and possibly also infer much larger related chromosomal segments.

The duplications we have detected and mapped in the maize genome could have been produced by several mechanisms: (1) formation of tandem duplications with subsequent dispersal, (2) reverse transcription of mRNA or hnRNA with subsequent reinsertion into the genome, (3) fusion of related genomes, or (4) internal duplication of chromosome segments. Examination of the individual characteristics of each model may provide some indication as to their relative roles in the generation of the duplications we have detected.

Tandem duplications have previously been shown to be present in the maize genome in the nucleolar organizing region (MCCLIINTOCK 1933), centromeres (MCCLIINTOCK 1938), the A locus (LAUGHAN 1949), the R locus (STADLER and NEUFFER 1953), and in maize knobs (PEACOCK et al. 1981). The methods used here would usually not detect exact or inexact tandem duplications with a low copy number (less than five) because they would be detected as either comigrating or non-comigrating fragments that cosegregate to a very high degree. Hence, our data, provides little information about this type of duplication or whether it transposes to different genomic locations.

The second mechanism has been proposed to explain some observations with duplicated animal genes. Gene sequences can be duplicated by reverse transcription of hnRNAs or mRNAs into DNA, and the DNA copy is then inserted into the genome (STEIN et al. 1983). This process was detected through the identification of genes which lack certain introns present in other homologous genes. It is unlikely that either this mechanism or the previously described
one played a significant role in generating the duplications detected in this study as there is little reason to expect that copies of segments would be inserted non-randomly. In other words, there is little reason to expect that segments duplicated from the long arm of chromosome 2 would be preferentially inserted into chromosome 7 in an ordered arrangement and not into other regions of the genome.

As indicated earlier, it has repeatedly been suggested that two genomes with \( N = 5 \) fused to produce an ancient amphidiploid which evolved into the modern maize genome, in a manner analogous to that three genomes with \( N = 7 \) are believed to have fused to produce the modern wheat genome (Sears 1975). The regions of presumptive homoeology involving chromosomes 2 and 7, 3 and 8, 6 and 8, and others identified in the current study demonstrate that large segments of considerable homology exist within the maize genome. Such regions would be present if fusion of related genomes occurred at some time during the evolution of maize. On the other hand, this study clearly demonstrates that five pairs of homoeologous chromosomes are not currently present within the maize genome. The data simply do not identify five pairs of chromosomes which share large numbers of duplicate loci over most of their length. In fact, the current study identified pairs of duplicate loci involving chromosome 1 with eight of the nine other maize chromosomes, and chromosomes 2, 3, 4 and 6 each had duplicate loci on seven nonhomologous chromosomes. We conclude that maize is not the result of a recent allopolyploid event because it does not now contain homoeologous chromosomes. However, the extensive largely homologous segments and numerous duplicated loci identified in the current study suggest that an allopolyploidization event may have been involved in its evolution. If such an event occurred, the genome has been extensively restructured after it occurred. One disconcerting observation connected with this model is that the events necessary for this restructuring (chromosome aberrations) are currently infrequent in maize as evidenced from the fact that no detectable translocations were found in 55 strains of maize from Latin America (Cooper and Brink 1937), and neither translocations nor inversions were found in 90 strains of corn from Latin America (Rhoaedes and Dempsey 1953).

Finally, it is possible that the related sequences we have detected were generated through internal duplications of chromosome segments within the maize genome. Although our observations are certainly consistent with such a derivation, it is difficult to distinguish this model from the model that maize had an allopolyploid origin followed by extensive restructuring of the genome and sequence divergence. However, certain facts are of interest. First, our study detected extensive regions of duplication in only certain segments of the chromosomes. If these were the result of fusion of two genomes, then some mechanism must have existed(s) which either promoted or prevented sequence divergence differentially over large chromosomal segments so that we are now unable to detect all loci as being duplicated elsewhere in the maize genome. Alternatively, some of these duplicated sequences may have been lost during the evolution of maize, and this would also explain why we were unable to detect them. Secondly, because many segments exist as dispersed multiple copies and not just "duplicated" sequences, we feel this also argues against fused genomes as the only source of duplications. For instance, the data for triose phosphate isomerase, Tpi, obtained through both isozyme analysis and RFLP analysis with a gene clone, confirm that there are at least three dispersed sequences for this enzyme. The same was also seen for some of the randomly chosen sequences of unknown function we used.

Obviously, it is not clear which of the latter two mechanisms played the most important role in determining the current structure of the modern maize genome. Possibly all four mechanisms, tandem duplication, reverse transcription and reinsertion, genome fusion, and partial duplication, have operated to some degree, with one or both of the latter two mechanisms playing a significant role in determining the current "global" genomic structure. However, the relative importance of these mechanisms is not clear at this time. Perhaps a similar search for duplications and parallel linkages in related genera such as Teosinte, Tripsacum, Sorghum, or even some of the \( N = 5 \) species using the maize probes which identify duplicated sequences would be most informative. If, for instance, the duplicate maize clones showed more homology to each other than they did to any of the genes from the putative precursors, this might suggest that they were derived from internal duplication of a common ancestor genome as opposed to having been generated during an ancient fusion event of two separate genomes. If there appears to be some organization to the process of duplication, that is some duplications are found in some but not all relatives, then this also might shed light on the ancestry of maize.

It is also important to consider the differences we have found when comparing the genomes of maize and tomato plants. As described in this paper, 62 of the 217 informative maize probes tested (28.6%) appeared to detect duplicate sequences. This contrasted sharply with our own results in tomato where, of the 104 informative clones tested, only two appeared to detect duplicate sequences under our standard hybridization conditions. Thus, we found that duplications of low copy number sequences are over
fourteen times as numerous in maize as in tomato. Bernatzky and Tanksley (1986b) also found that the majority of their cloned tomato sequences represented "unique" sequences, but found a higher percentage of duplicated sequences, a result that may be due to ancillary factors such as differences in hybridization conditions. Another difference between maize and tomato is the much higher level of polymorphism in maize. If we hybridize a labeled maize probe with DNAs from unrelated maize lines which were digested with three different restriction enzymes, we almost invariably identify a polymorphism in maize. If we hybridize a labeled maize probe with DNAs from unrelated maize lines which were digested with three different restriction enzymes, we almost invariably identify a polymorphism in maize. If we hybridize a labeled maize probe with DNAs from unrelated maize lines which were digested with three different restriction enzymes, we almost invariably identify a polymorphism in maize.

In helping to add more duplicated RFLP loci to the maize RFLP linkage map. Additionally we would like to acknowledge the significant efforts of C. Stubber and M. Edwards in helping us to place the isozyme loci on our linkage map. We would also like to acknowledge the contributions of Alan Katz for carrying out the statistical analyses in Table 1 and Susan Patterson and Jan Tvang for preparing the figures for this manuscript. We would like to thank Sandor Ltd. (Basel, Switzerland) and its subsidiary company, Northrup King Co., whose funding supported the initial work reported herein as well as Funk Seeds International for providing nursery facilities in which part of the material utilized in this study was grown. This work was supported in part by grant 6-CR-CR-1-2213 from the United States Department of Agriculture.

LITERATURE CITED


Duplicated RFLP Loci in Maize


Stadler, L. J., and M. G. Neuffer, 1953 Problems of gene structure. II. Separation of R' elements (S) and (P) by unequal crossing over. Science 117: 471–472.


Communicating editor: S. D. Tanksley