Inferences on the Genome Structure of Progenitor Maize Through Comparative Analysis of Rice, Maize and the Domesticated Panicoids

William A. Wilson,*,[1] Sandra E. Harrington,* Wendy L. Woodman,† Michael Lee,† Mark E. Sorrells* and Susan R. McCouch*

**Department of Plant Breeding, Cornell University, Ithaca, New York 14853-1901 and †Department of Agronomy, Iowa State University, Ames, Iowa 50011*

Manuscript received June 8, 1998
Accepted for publication June 8, 1999

ABSTRACT

Corn and rice genetic linkage map alignments were extended and refined by the addition of 262 new, reciprocally mapped maize cDNA loci. Twenty chromosomal rearrangements were identified in maize relative to rice and these included telomeric fusions between rice linkage groups, nested insertion of rice linkage groups, intrachromosomal inversions, and a nonreciprocal translocation. Maize genome evolution was inferred relative to other species within the Panicoideae and a progenitor maize genome with eight linkage groups was proposed. Conservation of composite linkage groups indicates that the tetrasomic state arose during maize evolution either from duplication of one progenitor corn genome (autoploidy) or from a cross between species that shared the composite linkages observed in modern maize (alloploidy). New evidence of a quadruplicated homeologous segment on maize chromosomes 2 and 10, and 3 and 4, corresponded to the internally duplicated region on rice chromosomes 11 and 12 and suggested that this duplication in the rice genome predated the divergence of the Panicoideae and Oryzoideae subfamilies. Charting of the macroevolutionary steps leading to the modern maize genome clarifies the interpretation of intercladal comparative maps and facilitates alignments and genomic cross-referencing of genes and phenotypes among grass family members.

THE family Gramineae is a diverse group of widely adapted species that have been classified into two major clades and a series of smaller groups based on molecular phylogenetic studies (Clark *et al.* 1995; Soreng and Davis 1998). One clade contains the Panicoideae subfamily (among others) including maize (*Zea mays*), sugarcane (Saccharum), sorghum (Sorghum), and millet (Pennisetum) and the other clade contains the Pooideae subfamily including wheat (Triticum), barley (Hordeum), rye (Secale), and oat (Avena). The subfamily Oryzoideae, including rice (Oryza) and wild rice (Zizania), and Bambusoideae (the woody bamboos) are recognized as early diverging lineages. Despite the evolutionary distance between these crop species, molecular mapping of the nuclear genomes using restriction fragment length polymorphism (RFLP) has allowed the development of comparative chromosome maps (for graphical displays of Rice-1 through Rice-12 see http://genome.cornell.edu/rice/quickqueries/; for review see Devos and Gale 1997).

The genome of cultivated rice is considered to resemble an ancestral grass genome with a high base chromosome number ($x = 12$), relatively few large-scale duplications (Causse *et al.* 1994; Harushima *et al.* 1998), and a relatively small genome size of 430 Mb (Arumuganathan and Earle 1991). Intensive research efforts in rice have produced high-density genetic maps (Causse *et al.* 1994; Kurata *et al.* 1994; Harushima *et al.* 1998), a library of expressed sequences (Yamamoto and Sasaki 1997), and physical maps based on ordered arrays of artificial chromosome clones (Umehara *et al.* 1994; Wang *et al.* 1995; Kurata *et al.* 1997; Zhang and Wing 1997). The ability to locate phenotypic effects on molecular linkage maps has provided the necessary tools to associate genes with agronomically important traits and positionally clone several of the underlying loci (Song *et al.* 1995; Yoshimura *et al.* 1998; Wang *et al.* 1999).

The maize genome offers a marked contrast to rice with a genome size six times larger (2500 Mb), a high proportion (60–80%) of rapidly evolving repetitive DNA (Flavell *et al.* 1974), and intragenomic gene duplication estimated to be equivalent to 1.5 rice genomes (Helentjaris *et al.* 1988; Ahn and Tanksley 1993). Intergenic repetitive DNA appears to be organized as a series of nested insertions of retrotransposons (San Miguel *et al.* 1996), accumulating rapidly due to numerous site-specific insertion events that expand the repetitive fraction of DNA between conserved gene sequences (Chen *et al.* 1997). Despite this rapid divergence of sequence in the repetitive fraction of the maize genome, linkage segments can be defined that are clearly syntenic

Corresponding author: Susan McCouch, Plant Breeding Department, 418 Bradfield Hall, Cornell University, Ithaca, NY 14853-1901. E-mail: srm4@cornell.edu

[1] *Present address:* Pioneer Hi-Bred International, Inc., Research and Product Development Windfall Research Center, Windfall, IN 46076.

with rice and other grasses, based on mapping RFLP detected by cDNAs. Most of these segments are duplicated within the maize genome and reflect a polyploidization event in the maize lineage (Helentjaris *et al.* 1988; Ahn and Tanksley 1993) that appears to have occurred after the divergence of maize from its near relatives, sugarcane (Da Silva *et al.* 1993; D'Hont *et al.* 1996; Ming *et al.* 1998) and Sorghum (Chittenden *et al.* 1994; Pereira *et al.* 1994), as both of these Panicoids lack the extensive intragenomic duplication observed in maize.

Within the concept that the grass family can be considered a single genetic system (Bennetzen and Freeling 1993; Moore *et al.* 1995), rice has emerged as a model for comparing genome structure across family members. Indeed, the genetic content of each of the 12 rice chromosomes has been partially identified within the base chromosome complements of other domesticated grasses (Ahn and Tanksley 1993; Van Deynze *et al.* 1995a,b; Devos and Gale 1997). In these studies, linkage groups of the Pooideae and Panicoideae contained segments homeologous to rice linkage groups in compound-composite arrangements. The component-composite relationships between rice linkage groups and those of other domesticated grasses demonstrated homeology among grass genomes; this suggests that analysis of rice may provide insight into the basic organization of the ancestral grass genome.

Efforts to physically align orthologous loci, identify evolutionary events involving chromosome rearrangement, and ultimately identify genes underlying common phenotypic effects (Paterson *et al.* 1995) require increased resolution of existing genetic and physical maps. We mapped loci detected by maize cDNA probes in both rice and maize to extend and refine this genome comparison and used this information to make inferences about the genome structure of a progenitor from which modern maize may have evolved. Understanding the structure of such a progenitor will facilitate alignments and genomic cross-referencing of genes and phenotypes and clarify the interpretation of intercladal map comparisons.

MATERIALS AND METHODS

Plant materials: The previously described interspecific *Oryza sativa* (*cv. BS125*) / *O. longistaminata* // *O. sativa* BC_1F_1 rice population (Causse *et al.* 1994) was utilized to map loci detected by a selection of maize cDNA probes. The DNA extracted for this study was obtained from BC_1 plants that have been maintained by vegetative propagation for 10 years at Cornell University. Two recombinant inbred maize populations, T232/CM37 (48 RILs) and CO159/Tx303 (41 RILs) developed at Brookhaven National Laboratory (BNL) (Burr *et al.* 1988), were utilized to map loci detected by the Iowa State University (ISU) maize cDNA library. Both populations have been utilized by maize researchers to place molecular markers and to produce a database of mapping information (Burr *et al.* 1993; Matz *et al.* 1994). The T232/CM37 population was utilized by Ahn and Tanksley (1993) to map 250 loci in maize corresponding to 145 single-copy loci in rice.

Probes: Maize clones from three cDNA libraries, Colorado State University (CSU; provided by C. Baysdorfer via E. Coe), University of Arizona (UAZ; provided by T. Helentjaris), and Iowa State University (ISU; provided by M. Lee), were selected for mapping in rice. The cDNA libraries were constructed from maize endosperm and etiolated seedlings (UAZ), leaf (CSU), and root (ISU) tissues as described in Shen *et al.* (1994), Chao *et al.* (1994), and Pereira *et al.* (1994), respectively.

In this study, probes from the ISU cDNA library were mapped for the first time in both maize and rice. CSU and UAZ probes that had been previously utilized in maize were mapped only in rice and were aligned with previously reported map positions that were available from public databases (Burr *et al.* 1993; Matz *et al.* 1994; Davis *et al.* 1996). A portion of the cDNA probes selected from the UAZ and CSU libraries have been end-sequenced and are homologous to known genes (Chao *et al.* 1994; Shen *et al.* 1994). In addition to cDNAs with known function, three maize genes were mapped in rice: phytoene desaturase (*pds*; provided by K. Oishi), phosphohexose isomerase (*phi*; provided by M. Sachs), and enolase 2 (*eno2*; M. Sachs).

All probes were surveyed on *grass garden* blots consisting of *Eco*RI-digested genomic DNA from oat, wheat, barley, rice, bamboo, sugarcane, pearl millet, sorghum, and maize, to evaluate signal and copy number across cultivated species of the Poaceae, as described in Van Deynze *et al.* (1998). In addition, genomic DNA from *Joinvilleae ascendens* (leaf tissue provided by K. Wood of the National Tropical Botanical Garden, Lawai, Kauai, Hawaii) and *Flagelleria* sp. (Cornell University Conservatory), organisms closely related to the grass family (Doyle *et al.* 1992; Duvall *et al.* 1993; Davis 1995), were included to broaden the survey of sequence divergence in the monocots based upon Southern hybridization intensity.

RFLP protocols and data analysis: DNA digestion, Southern blotting, probe preparation, and hybridization were similar to the methods reported by McCouch *et al.* (1988) for rice and Veldboom *et al.* (1994) for maize. For rice, hybridizations were performed at 65° for at least 20 hr and subsequent washes at 65° at 2×, 1×, and 0.5× SSC. Probes were surveyed on restriction-digested genomic DNA from the recurrent parent *O. sativa* (cv. BS125) and the *O. sativa*/*O. longistaminata* F_1 hybrid. Five restriction enzymes (*Eco*RV, *Eco*RI, *Xba*, *Hin*dIII, and *Sca*I) were sufficient to detect at least one polymorphic locus for >95% of probes giving strong signal. Probes were then hybridized to filters containing genomic DNA from the rice BC_1F_1 individuals to obtain segregation data for the polymorphic restriction fragments.

Copy number was estimated in rice on the basis of the number of fragments detected by each probe with five enzymes on parental polymorphism surveys. Probes detecting single-copy loci in rice were classified on the basis of detection of a single band in the inbred BS125 parent and one or two copies in the interspecific F_1 for at least two restriction enzymes. Low-copy probes detected more than one but less than four fragments in the inbred parent. Higher-copy probes that detected more than four fragments were not utilized for mapping. Multiple-copy loci were designated A, B, C, or D on the basis of the molecular weight (descending order) of the mappable fragments from the BS125 parent. If only one fragment of a multiple copy probe was mapped, a designation of X was added to indicate the possibility of additional loci that were not mapped due to monomorphism or poor signal.

Marker placement on rice and maize framework maps: The LOD 2.5 rice framework map presented in Causse *et al.* (1994) was used as the basis for integrating markers detected by an-

chor probes (Ahn and Tanksley 1993) and maize cDNAs mapped in this study. Selection of framework markers was based on screening the primary data for suspect double crossovers represented by a single data point (marker) and each was reviewed for accuracy of the initial scoring. Subsequent to this screening, 3 individuals, SL-35, SL-40, and SL-147, were identified as confounded genotypes due to a high frequency of double crossovers scattered along all linkage groups. These individuals were eliminated from the primary data set and the remaining 110 individuals formed the basis of a revised mapping population. Genetic linkage maps were constructed for maize and rice using Mapmaker v2.0 (Lander *et al.* 1987). Genetic distance in rice was estimated from recombination frequencies using the Kosambi function. Loci in rice detected by maize probes were added to the revised framework map using the *try* command. Subsequent *compare* and *ripple* commands were used to verify positioning at LOD = 2.0. Rice centromeric intervals were based on the secondary and telotrisomic centromeric placement of Singh *et al.* (1996).

A maize map was constructed using selected markers from the T232/CM37 RI population framework marker data from Matz *et al.* (1994), comparatively mapped markers from Ahn and Tanksley (1993), and ISU markers mapped in this study. ISU markers were mapped either directly based on segregation data in the T232/CM37 population or aligned from data obtained from the Tx303/CO159 RI map. A LOD 2 framework map was constructed based on selection for comparatively mapped markers, while maintaining the BNL framework. Genetic distance was estimated from recombination frequencies using the Kosambi function. Markers mapped in the Tx303/CO159 RI (Burr *et al.* 1988; Matz *et al.* 1994) and CO159/Tx303 immortalized F_2 population (Davis *et al.* 1996) were positioned to intervals based on markers anchoring the maps (providing links among likely orthologs at the resolution RFLP analysis).

RESULTS

Placement of polymorphic loci detected by maize probes on the rice map: A total of 210 maize cDNA probes were surveyed on garden blots and parental polymorphism surveys for hybridization signal and RFLP, respectively. Of these, 23 probes (~10%) detected loci in maize and gave no signal in rice, 5 detected monomorphic loci in rice, 11 were high-copy probes, 10 detected smears, 14 were not chosen for mapping based upon position or weak signal, and 147 maize probes (35 CSU, 38 UAZ, 71 ISU, and three known genes) were mapped in rice. These probes identified 182 new loci that serve as common reference points between the rice and maize genome maps (Figures 1 and 2).

The 182 loci were combined with 423 of the loci previously mapped in rice (Causse *et al.* 1994), including 146 loci comparatively mapped between maize and rice (Ahn and Tanksley 1993), to produce a 623-marker rice map that facilitates genomic comparisons with other grasses (Figure 1). Rice chromosomes were divided into long and short arms based on the secondary and telotrisomic centromeric approximations (Singh *et al.* 1996).

Placement of polymorphic loci detected by maize probes on the maize T232/CM RI map: Marker loci from several maize RFLP genetic maps were combined to facilitate comparison of the maize and rice genomes. The BNL 96 T232/CM RI genetic map (Burr *et al.* 1993) was selected to serve as a framework for the addition and positioning of loci necessary for the comparison. Newly mapped ISU markers, previously mapped markers from Ahn and Tanksley (1993), and markers retrieved from the maize genome database (http://www.agron.missouri.edu) were positioned relative to the framework at LOD 2.0. Conflicting marker orders due to low LODs arising from tight linkage were sorted to the inclusion of comparatively mapped loci on the framework. ISU loci mapped in CO159/Tx303 RI (Matz *et al.* 1994) and markers retrieved from the maize genome database mapped onto the same population as above and/or Tx303/CO159 IF_2 (Davis *et al.* 1996) were cross-referenced by common markers and placed to inferred positions. In total, the resulting maize map included 422 comparative markers placed directly on, or positioned relative to, the T232/CM37 map (Figure 2).

Rationale for inferring the macroevolution of maize genome structure: Intragenomic homeology between paired segments of maize chromosomes supports the proposal that the maize genome has fixed a polyploidization event and switched from tetrasomy to disomy (Helentjaris *et al.* 1988; Ahn and Tanksley 1993). In addition, 20 chromosomal rearrangements (Tables 1 and 2) were inferred by map comparison and were classified as telomeric fusions between rice linkage groups, nested insertion of rice linkage groups, intrachromosomal inversions, and a nonreciprocal translocation. All were inferred to have occurred in maize relative to rice.

Marker order within duplicated segments, conserved on composite linkages within the maize genome, and homeology between rice and maize provided a basis for inferring the structure of ancestral chromosome arrangements in maize based on the following assumptions:

1. Rice is representative of an ancestral genome. Each of the rice linkage groups was viewed as a *component* that occurs in differing arrangements within each of the grass genomes. Many maize chromosomes were observed to be composed of more than one ancestral component and will be referred to as *compound* (two) or *composite* (three or more) linkage groups. Events leading to compound/composite linkage groups included nested insertion of, and end-to-end fusion between, ancestral components. Intra- and interchromosomal rearrangements have separated some component linkages and it is convenient to refer to these portions as *segments*.
2. A polyploidization event occurred in the lineage of modern maize, and that event could have resulted from either the doubling of a hybrid between two closely related species (allopolyploid) or doubling of a single ancestral genome (autopolyploid). The event

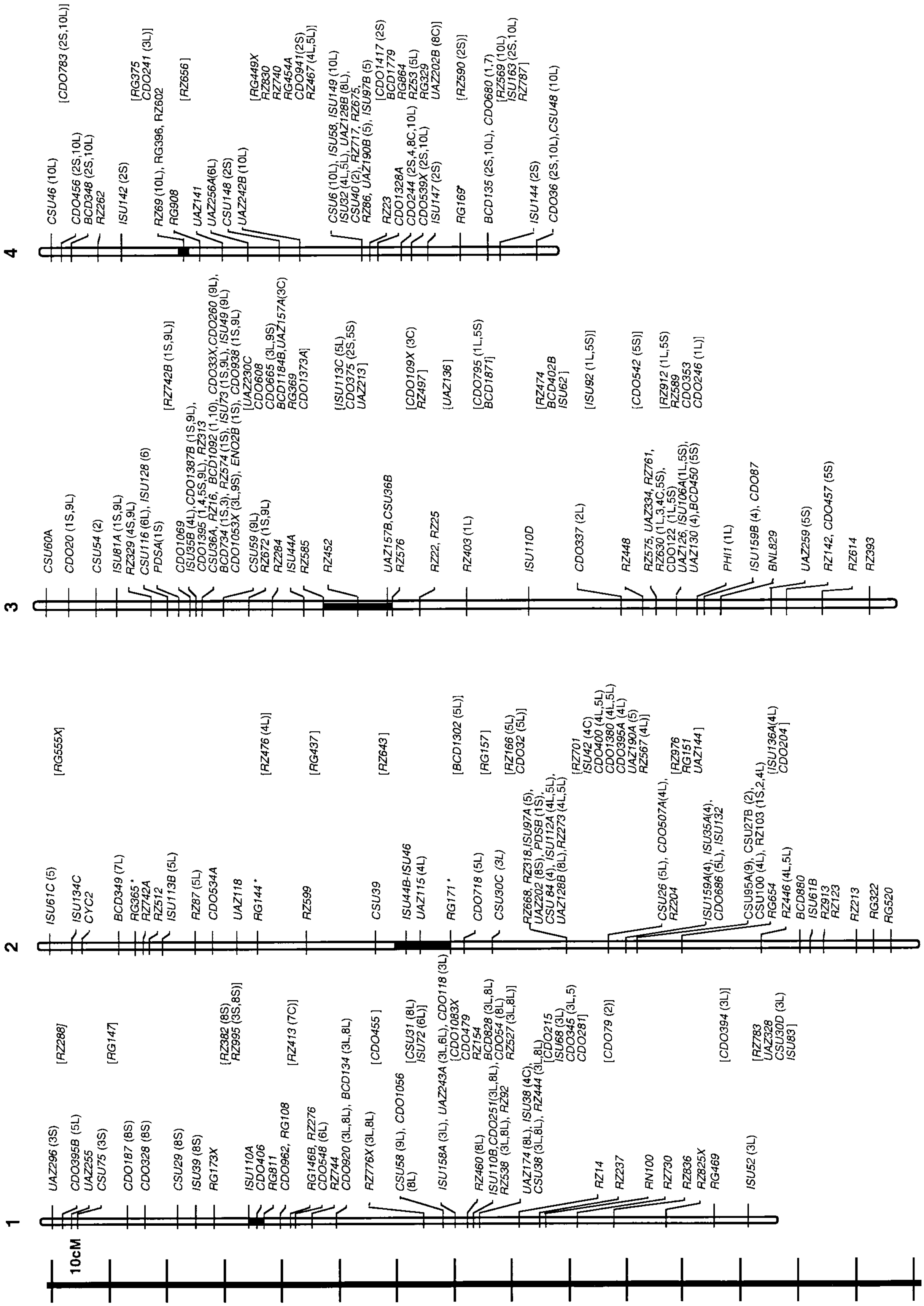

Figure 1.

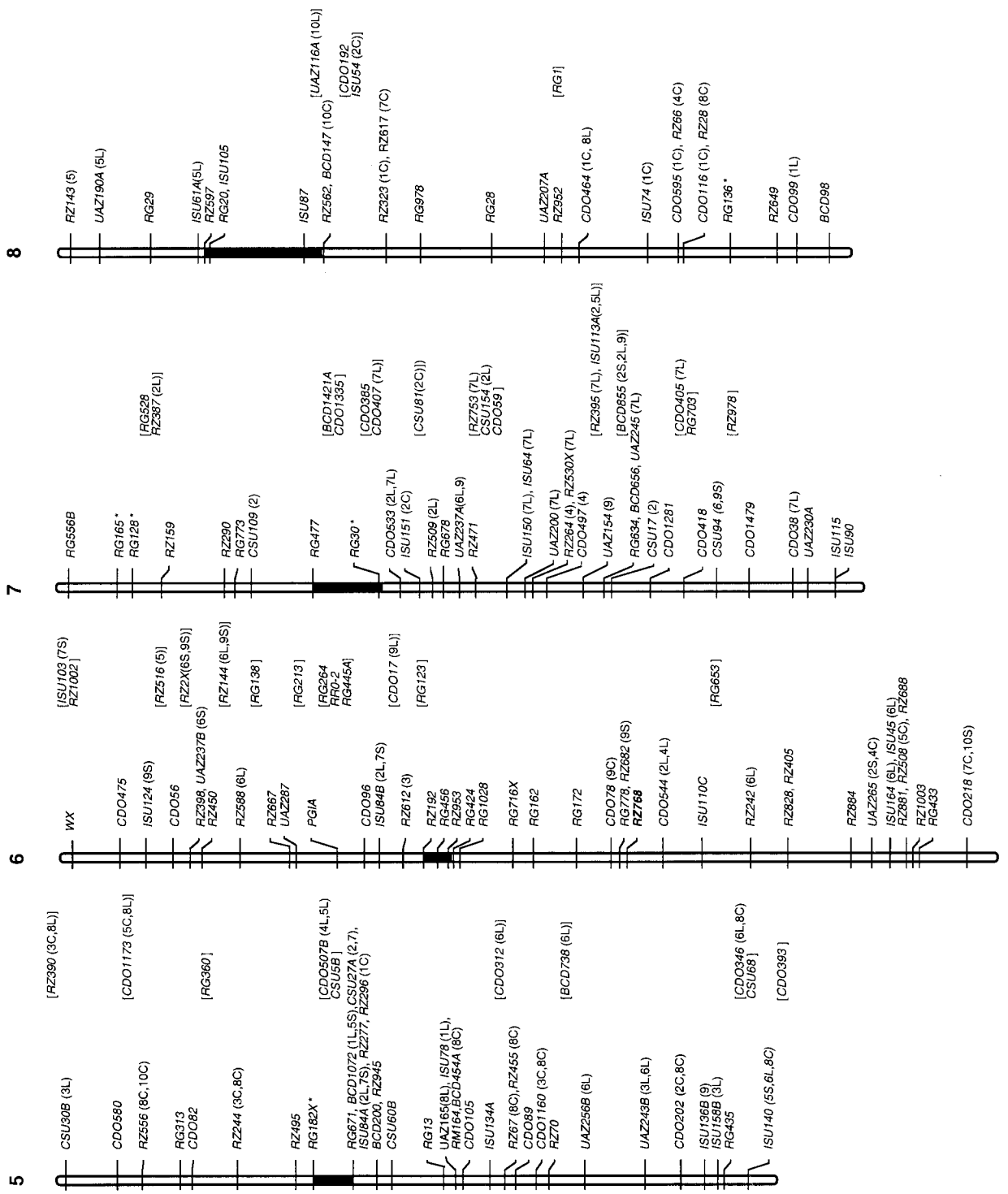

Figure 1.—Rice RFLP linkage map based on the interspecific (*O. sativa* / *O. longistaminata* // *O. sativa*) population. A comparative framework map ordered at a minimum LOD 2 was based on high confidence markers from Causse *et al.* (1994) and Ahn and Tanksley (1993) and newly placed markers detected by maize cDNA probes in this study. Markers were designated by abbreviations designating the cDNA or genomic library from which the probe was isolated as follows: RZ (rice cDNA); RG (rice genomic); BCD (barley cDNA); CDO (oat cDNA); and CSU, ISU, UAZ (maize cDNA). Centromere positioning was based on the intervals presented by Singh *et al.* (1996). Brackets denote markers placed at <LOD 2. Parentheses indicate chromosomal location of markers comparatively mapped in maize; suffixes following chromosome numbers are L, long arm; S, short arm; and C, centromere.

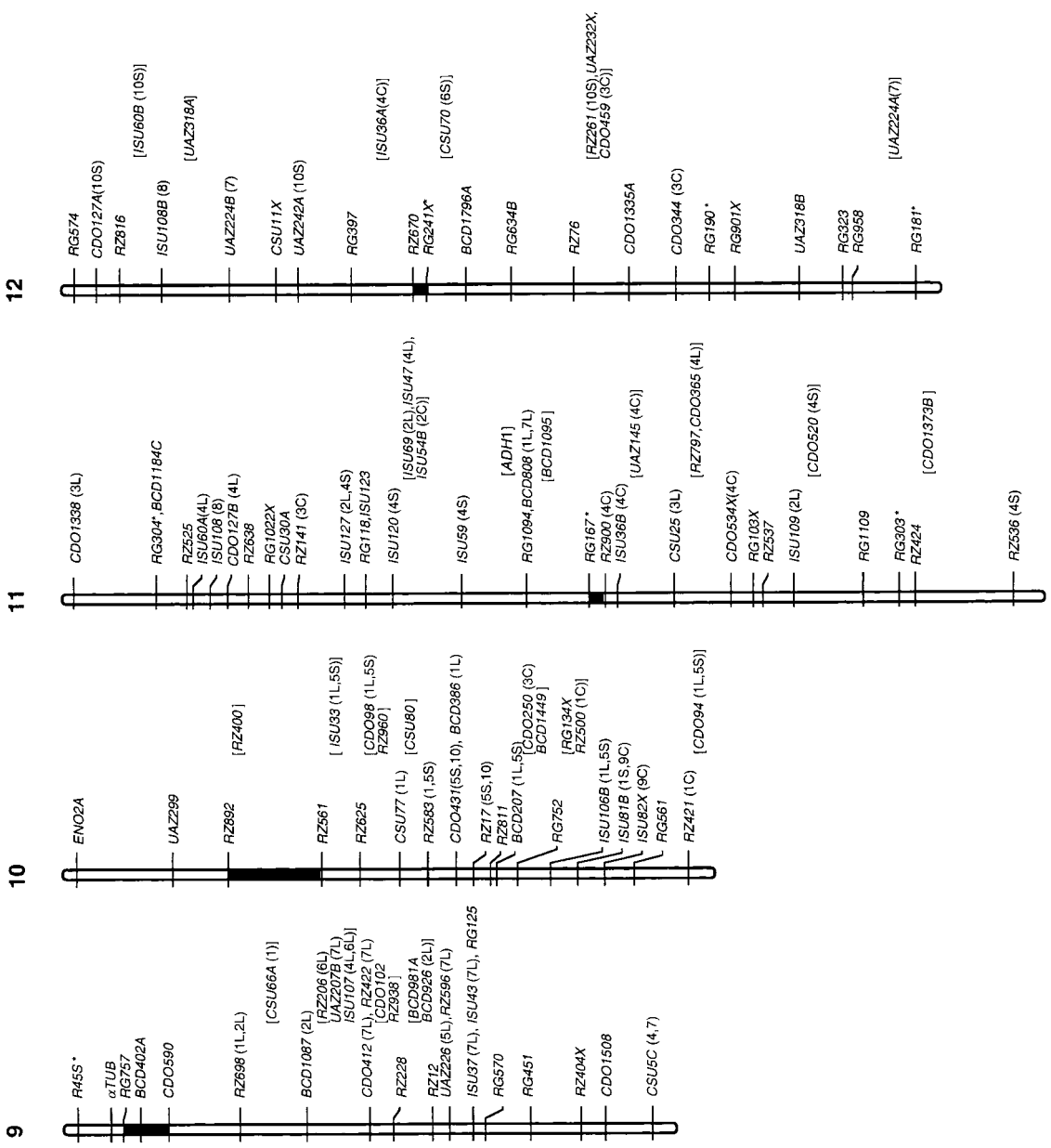

Figure 1.—*Continued.*

will be referred to as the maize duplication event (MDE).

3. Within the modern maize genome, duplicated composite linkages and chromosomal rearrangements were considered fixed prior to the MDE and those not duplicated in the maize genome occurred subsequent to the MDE. Inferences of progenitor genome structure followed Occhams razor, the simplest model based on available data from rice, maize, and sorghum.

Retracing the divergence of the maize genome from an ancient grass similar to rice requires utilization of major divergence points in the evolution of the maize lineage. Comparison of rice, maize, Sorghum, and sugarcane using systematic classification (Celarier 1957; Clark *et al.* 1995; Soreng and Davis 1998) suggests the definition of six evolutionary intervals delimited by the following divergence events (Figure 3):

1. Ancestral, prior to the divergence of Oryzoideae and Panicoideae subfamilies.
2. Progenitor Panicoideae, prior to the divergence of Paniceae (millets) from Andropogoneae-Maydeae.
3. Progenitor Andropogoneae/Maydeae prior to the divergence of Andropogoneae from Maydeae (sugarcane, sorghum, and maize).
4. Progenitor Maydeae, prior to divergence of the Maydeae, specifically Tripsacum and progenitor maize.
5. Progenitor Maize, following divergence from Tripsacum and prior to genome-wide duplication event.
6. Tetraploid Maize, duplicated genome, from tetrasomic to disomic inheritance.

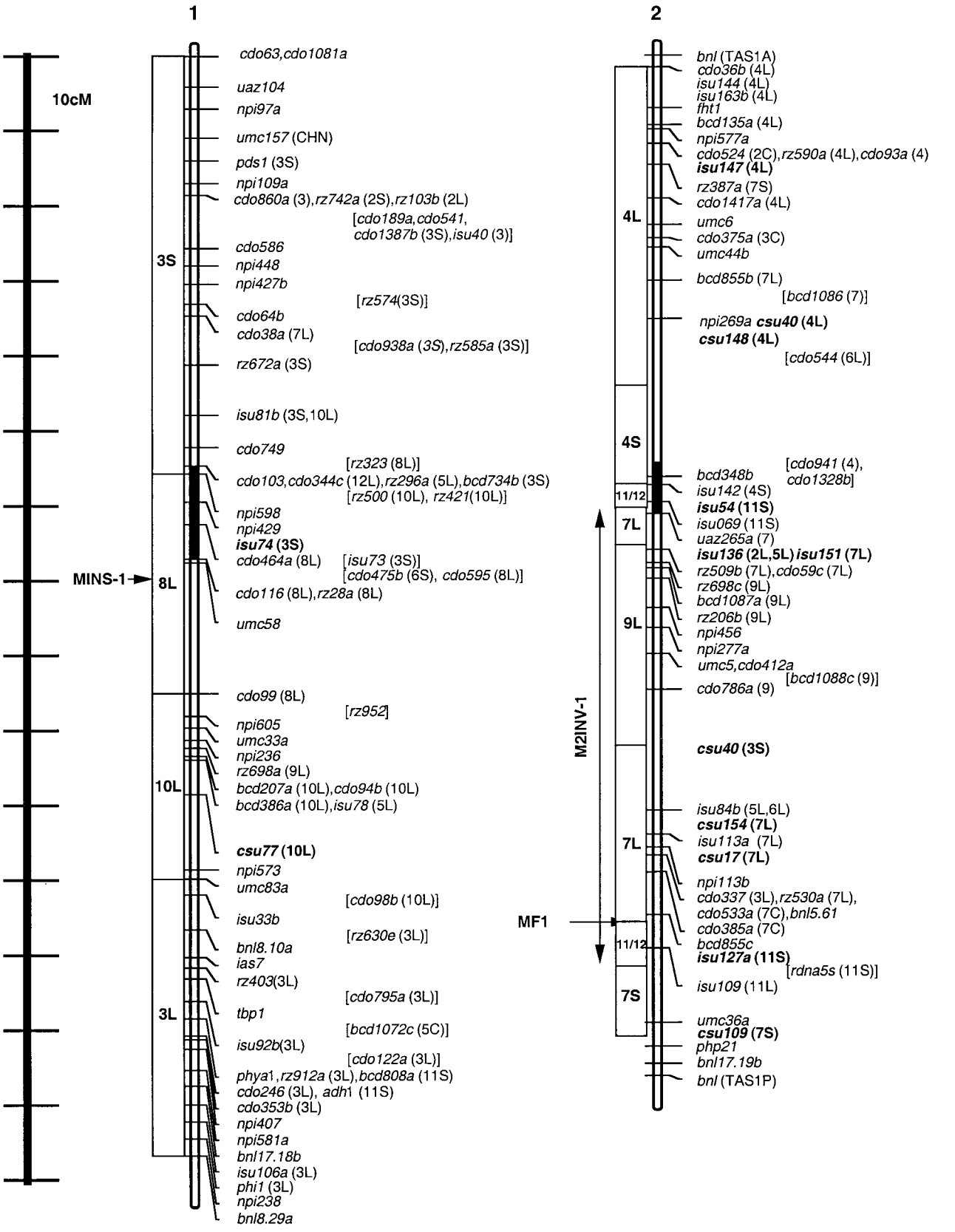
1
2
10cM
3S
8L
10L
3L
MINS-1
cdo63,cdo1081a
uaz104
npi97a
umc157 (CHN)
pds1 (3S)
npi109a
cdo860a (3),rz742a (2S),rz103b (2L)
[cdo189a,cdo541,
cdo1387b (3S),isu40 (3)]
cdo586
npi448
npi427b
[rz574(3S)]
cdo64b
cdo38a (7L)
[cdo938a (3S),rz585a (3S)]
rz672a (3S)
isu81b (3S,10L)
cdo749
[rz323 (8L)]
cdo103,cdo344c (12L),rz296a (5L),bcd734b (3S)
[rz500 (10L), rz421(10L)]
npi598
npi429
isu74 (3S)
cdo464a (8L) [isu73 (3S)]
[cdo475b (6S), cdo595 (8L)]
cdo116 (8L),rz28a (8L)
umc58
cdo99 (8L)
[rz952]
npi605
umc33a
npi236
rz698a (9L)
bcd207a (10L),cdo94b (10L)
bcd386a (10L),isu78 (5L)
csu77 (10L)
npi573
umc83a
[cdo98b (10L)]
isu33b
[rz630e (3L)]
bnl8.10a
ias7
rz403(3L)
[cdo795a (3L)]
tbp1
[bcd1072c (5C)]
isu92b(3L)
[cdo122a (3L)]
phya1,rz912a (3L),bcd808a (11S)
cdo246 (3L), adh1 (11S)
cdo353b (3L)
npi407
npi581a
bnl17.18b
isu106a (3L)
phi1 (3L)
npi238
bnl8.29a
4L
4S
11/12
7L
9L
7L
11/12
7S
M2INV-1
MF1
bnl (TAS1A)
cdo36b (4L)
isu144 (4L)
isu163b (4L)
fht1
bcd135a (4L)
npi577a
cdo524 (2C),rz590a (4L),cdo93a (4)
isu147 (4L)
rz387a (7S)
cdo1417a (4L)
umc6
cdo375a (3C)
umc44b
bcd855b (7L)
[bcd1086 (7)]
npi269a csu40 (4L)
csu148 (4L)
[cdo544 (6L)]
[cdo941 (4),
cdo1328b]
bcd348b
isu142 (4S)
isu54 (11S)
isu069 (11S)
uaz265a (7)
isu136 (2L,5L) isu151 (7L)
rz509b (7L),cdo59c (7L)
rz698c (9L)
bcd1087a (9L)
rz206b (9L)
npi456
npi277a
umc5,cdo412a
[bcd1088c (9)]
cdo786a (9)
csu40 (3S)
isu84b (5L,6L)
csu154 (7L)
isu113a (7L)
csu17 (7L)
npi113b
cdo337 (3L),rz530a (7L),
cdo533a (7C),bnl5.61
cdo385a (7C)
bcd855c
isu127a (11S)
[rdna5s (11S)]
isu109 (11L)
umc36a
csu109 (7S)
php21
bnl17.19b
bnl (TAS1P)

Figure 2.

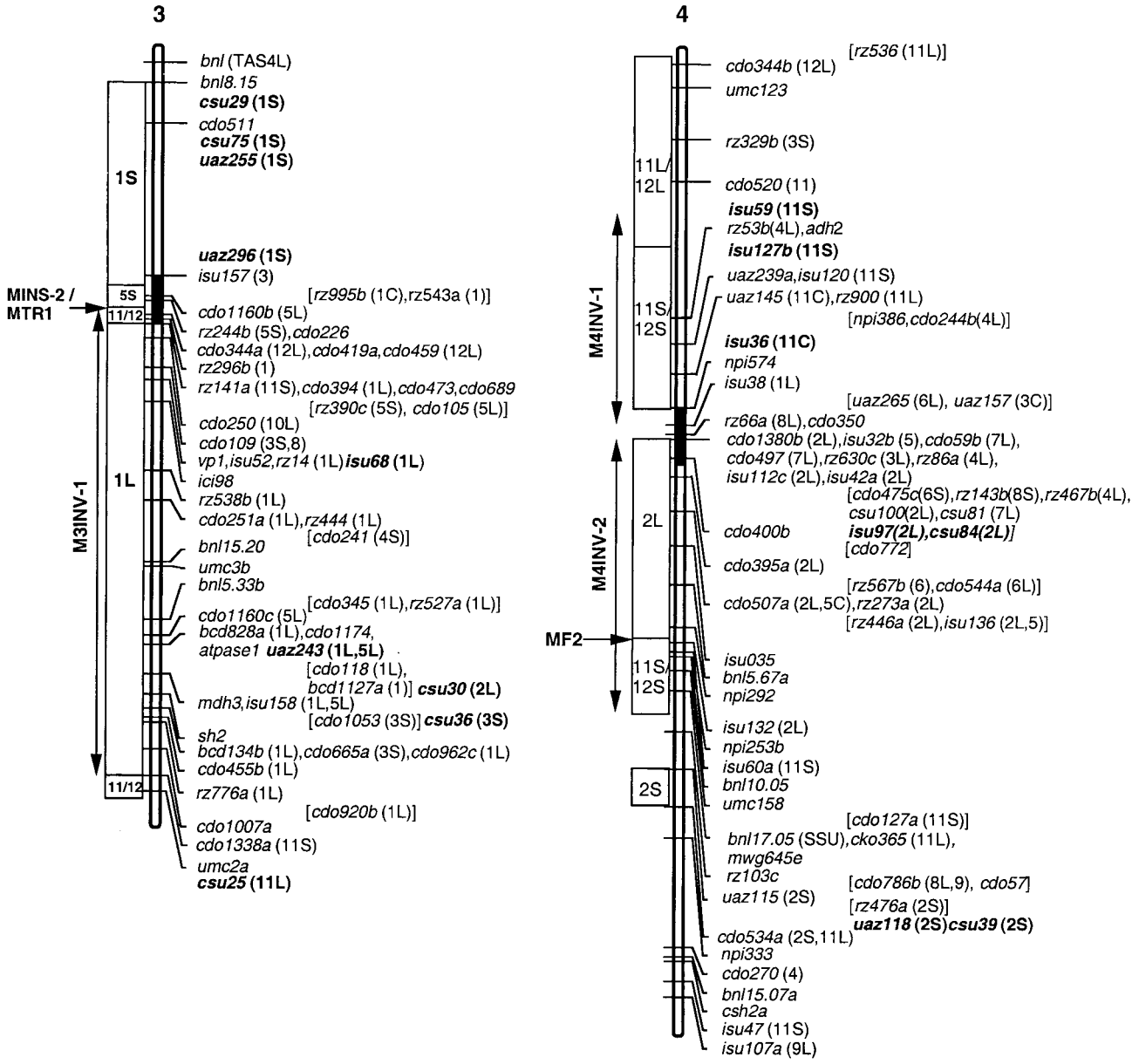

Figure 2.—Maize RFLP T232 / CM37 RI comparative linkage map. A comparative LOD 2 map was constructed using the BNL96 framework markers as a foundation for placement of loci detected by ISU cDNA probes in this study and reintegration of previously mapped CDO, BCD, and RZ markers of Ahn and Tanksley (1993). Genetic distance was estimated on the basis of recombination frequency using the Kosambi function. Parentheses and brackets in probe designations are the same as in Figure 1, with the exception that chromosomal locations in parentheses represent corresponding rice chromosomal positions. Map positions of markers retrieved from public databases were added for CSU and UAZ markers to provide reference to loci mapped in rice. Markers mapped in the CO159 / Tx303 population were positioned relative to consensus markers in the BNL maps and noted in bold. Boxes to the left of chromosomes define homeologous segments and indicate their corresponding position in the rice genome. Arrows indicate breakpoints and rearrangements as summarized in Tables 1 and 2.

Divergence of repetitive DNA fractions, subcentimorgan order, and nonrepetitive sequences were assumed to increase with each meiotic event to further separate each of the grasses. Although homeology exists among the grass linkage groups and may be conveniently compared to rice segments, it is important to note that all modern genomes have diverged from the ancestral grass genome (Figure 3, node 1). Sorting chromosomal rearrangements into intervals of maize evolution was necessary to elucidate events that differentiate the maize and rice genomes from each other and from the rest of the domesticated Poaceae.

Homeology between maize and rice: a comparison of the grass progenitor: Genomic comparison between rice and maize revealed that the genomes are segmentally homeologous (Ahn and Tanksley 1993) and that considerable rearrangement of the maize genome has occurred over the millions of meiotic events that have

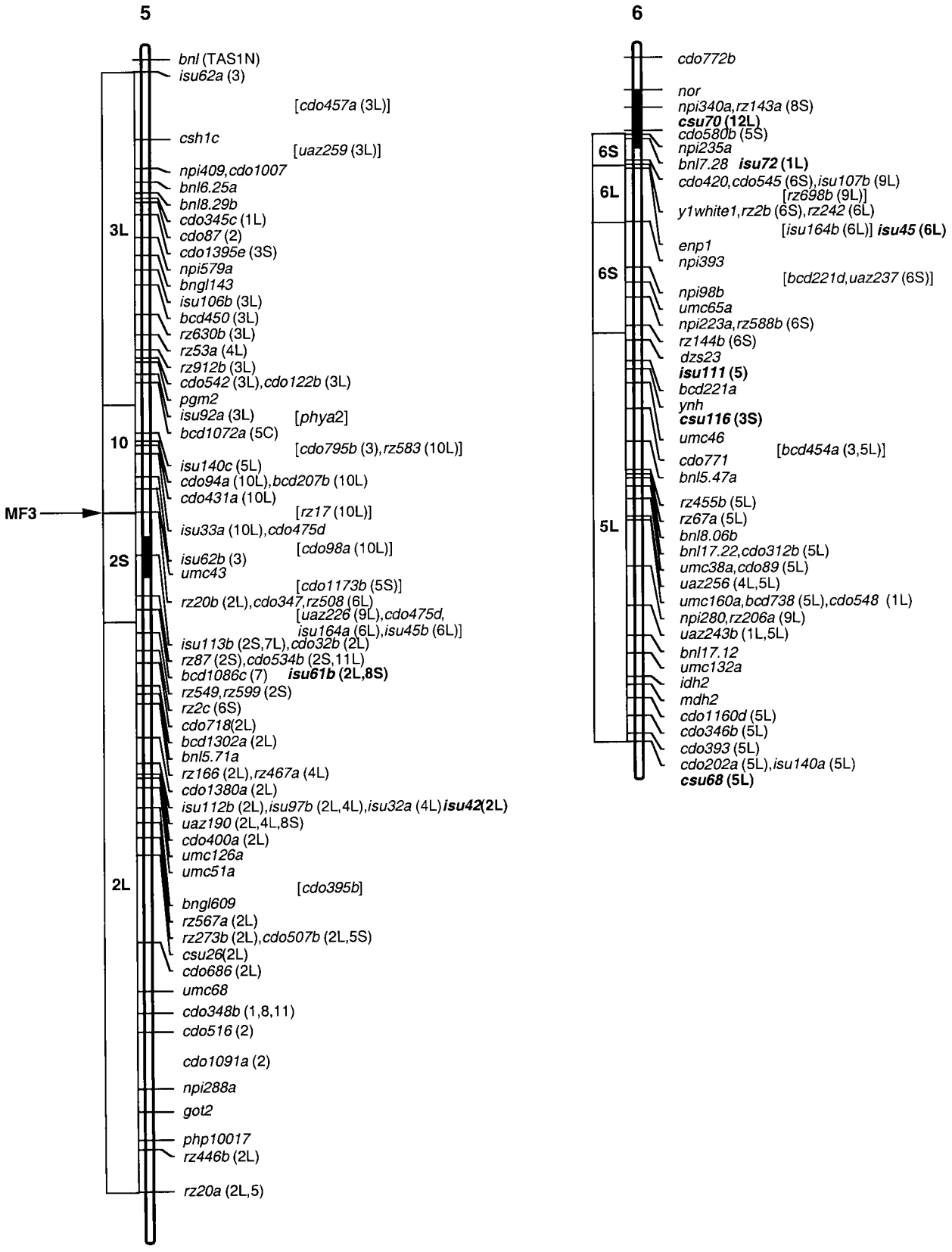

Figure 2.—*Continued.*

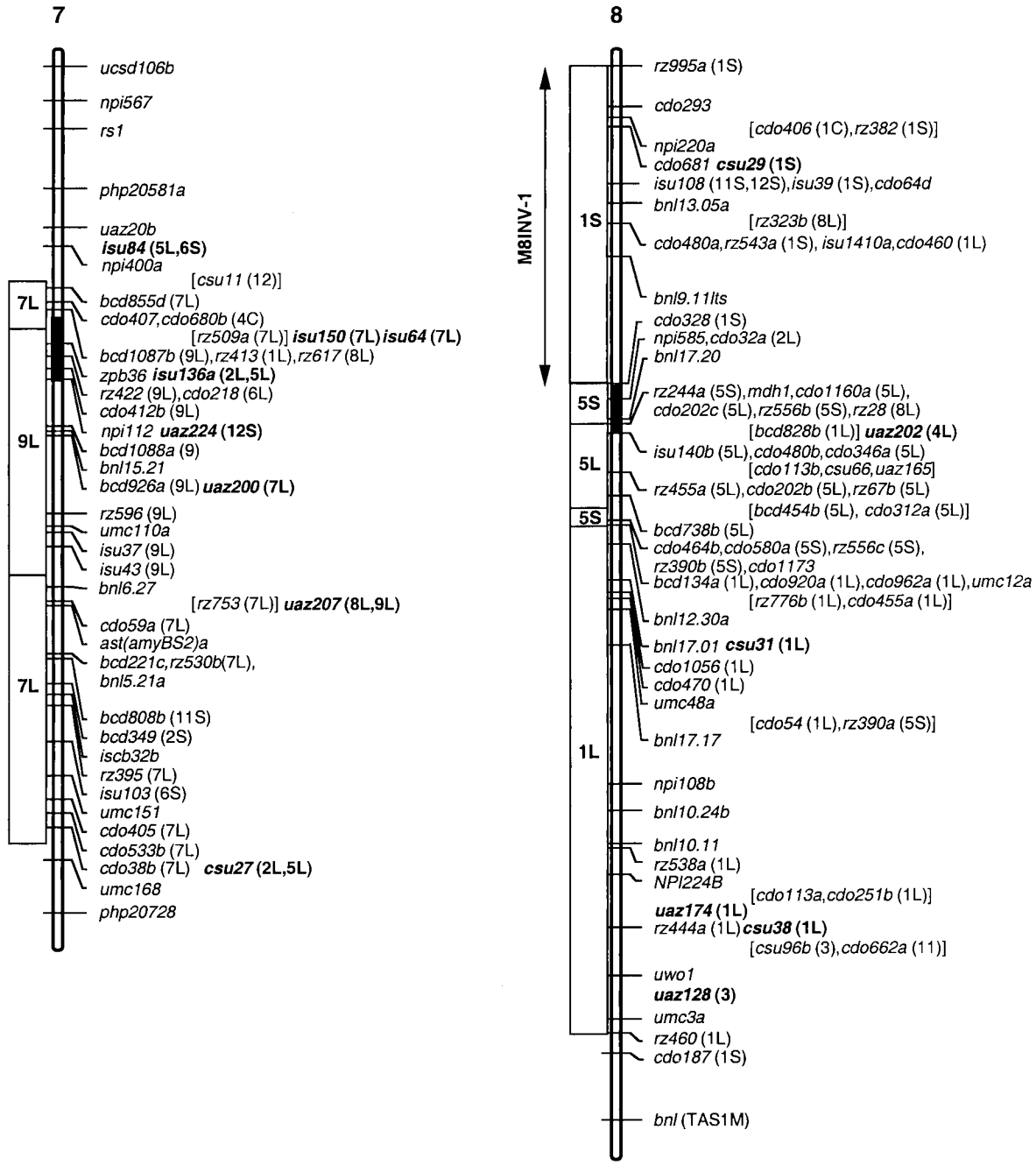

Figure 2.—*Continued.*

separated the two taxa. Anchoring an additional 182 loci in rice for maize-rice map alignments has supported and refined the definition of homeology between pairs of maize chromosomes (1-5, 1-9, 2-7, 2-10, 3-8, 4-5, 6-8, and 6-9) and between maize segments and rice linkage groups (Figure 2). The presence of segmental homeology is a remnant of the genomic constitution of the grass progenitor and indicates that a common ancestor existed and remains intact within both the rice and maize genomes.

Assignment of homeology between rice chromosomes 11 and 12 to segments within the maize genome has been tenuous due to a complex arrangement of duplicate loci on these linkage groups (Panaud *et al.* 1996; Nagamura *et al.* 1997). Of 15 single- or low-copy maize probes detecting loci on the short arms of R11 or R12, 3 detected loci on both chromosomes. Intervals delimited by *isu60a*/*isu60b* and *isu36b*/*isu36a* on R11 and R12, respectively, indicate that duplicate loci may be expected on the short arms of R11 and R12. It is also apparent from duplicate loci detected by *isu36*, *isu60*, and *isu85* on sorghum linkage groups I and J (Pereira *et al.* 1994) that this duplication may have existed in the ancestral genome and predates the divergence of

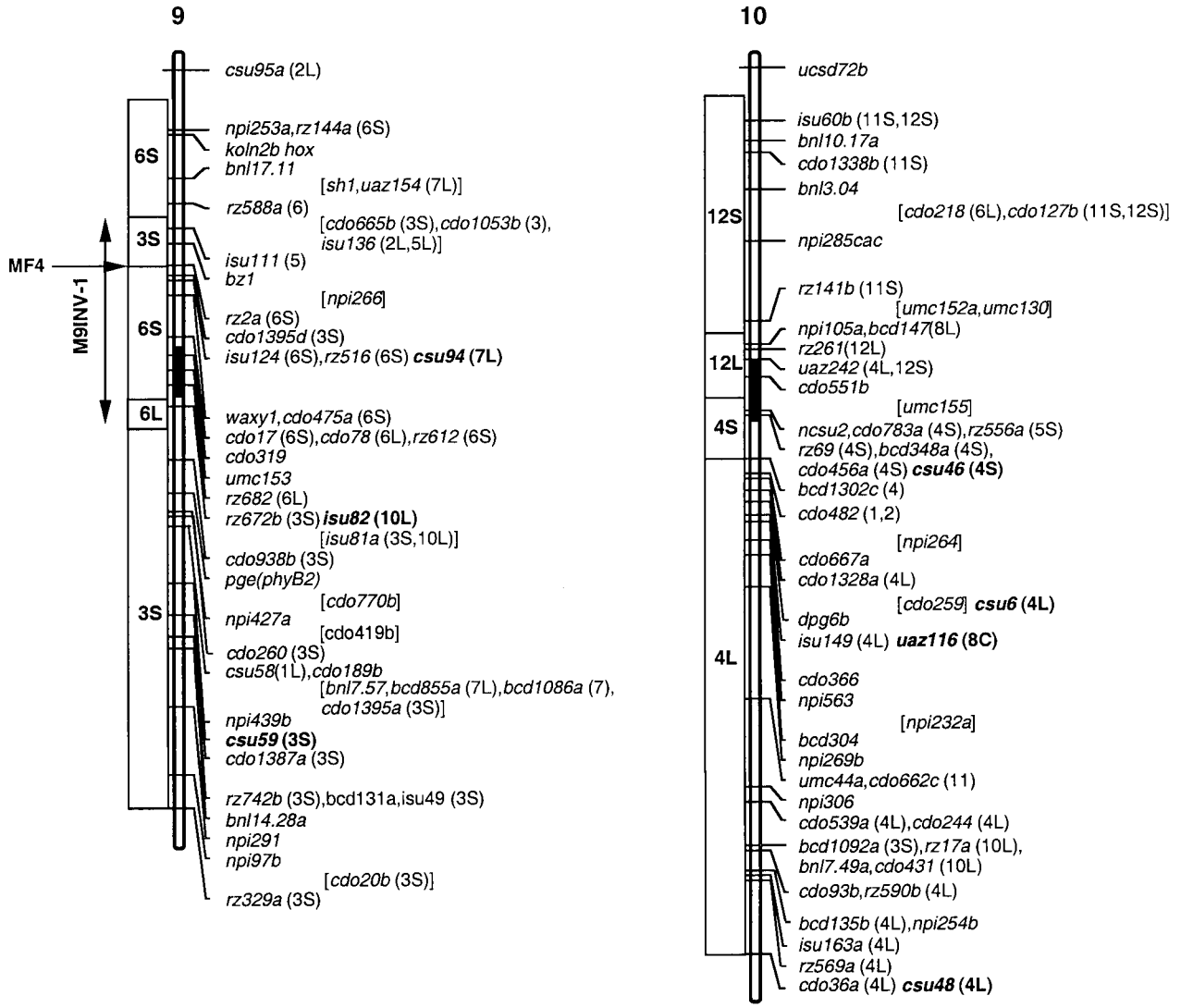

Figure 2.—*Continued.*

Panicoideae and Oryzoideae. If the duplication of these loci predates the divergence of maize, sorghum, sugarcane, and rice, quadruplication of these loci and their chromosome segments would be expected within the maize genome. Genetic quadruplication is difficult to define due to selection against the use of multiple-copy probes in mapping studies and the reduced likelihood of detecting four polymorphic fragments in a single cross. As a result of these complications, the detection of probes mapping to one, but not both, segments is common and gives the appearance of incongruency between a locus and the map location.

Probes detecting loci on R11 detected loci on M2, M3, M4, M8, and M10 (Figures 1 and 2). Probes detecting loci on R12 detected loci on M2, M4, M6, M7, M8, and M10. Segments showing homeology (more than four probes) to R11 or R12 (designated R11/12) are located on M2, M3, M4, and M10. The most extensive conservation of gene content and order was detected on segments of M4 that showed considerable homeology to R11/12, after accounting for a rearrangement within the long and short arms of M4. The expected duplicate counterpart of M4 (R11/12) was difficult to identify in the maize genome. Five probes (*isu54*, *isu69*, *isu109*, *isu123*, and *bcd808*) on M2 detected loci on R11. These loci define a segment, although separated into two clusters by >50 cM in maize, that was assigned homeology to R11/12 after accounting for an inversion event on the long arm of M2.

While selected markers can be used to assign paired homeology between these segments in the maize genome, compound linkages that predate the duplication of the maize genome may be a more reliable method. Alignment of maize segments into a consensus segment representative of paired duplicate segments and evaluation of compound linkages in maize relative to rice indicated that paired homeology of R11/12 segments consisted of M3-M4 and M2-M10. This is in contrast to

TABLE 1

Genome rearrangements inferred to have occurred in evolutionary intervals prior to the tetraploidization event in maize

Event designation[a]	Evolutionary internode[b]	Chromosomal rearrangement	Homeologous rice components	Breakpoint(s)/fusion point(s) (chromosome location in rice)	Resulting genome feature	Modern maize chromosome affected
P INS-1	2	Centric insertion	R10 → R3	*rz585-isu113*/*rz585-rz421c*(10L), *isu33*(101)-*isu113c*(3C)	P3 compound	M1, M5, M9
P INS-2	2	Insertion	R9 → R7L	*cdo59*(7L)-*uaz200*(7L)/*cdo59-bcd1087*, *uaz200*(7L)-*isu37*	P7 compound	M2, M7
PM INS-1	4 or 5	Centric insertion	R5 → R1	*rz995-cdo920*/*rz995-rz390*, *isu140-cdo920*	PM1 compound	M3, M8
PM INV-1	4 or 5	Pericentric inversion	R5	*rz556*(5S)-*rz244*(5S)/*rz556*(5S)-TELOMERE(5L)	PM1 inversion	M3, M8
PM INV-2	4 or 5	Pericentric inversion	R6	*cdo475*(6S)-*uaz237*(6S)/*cdo475*-TELOMERE(6L)	PM6 inversion	M6, M9
PM F-1	5	End-to-end fusion	R4S-R11/12L	NBP/*csu46*(4S)-*rz261*(11/12L)	PM4 compound	M2, M10
MDE	5	Polyploidization	R1-R12	Genome duplication to tetraploid level	Tetraploid genome	All

[a] P, Panicoideae genome; PM, progenitor Maydeae; M, maize; MDE, maize duplication event; INS, insertion; INV, inversion; F, fusion; NBP, no breakpoint.
[b] Refer to Figure 4 for numerical designation of evolutionary internodes.

the homeology assigned on the basis of sampled markers and to the previous assignment of homeology between M3-M10 proposed by Moore *et al.* (1995).

Previous distinction of interspersed single-copy loci and linkage segments throughout the maize genome (Ahn and Tanksley 1993) has been reduced by the mapping of duplicated loci (Figure 2). Of all grass components, only the duplicate counterpart of the segment homeologous to R8 has not been identified in the maize genome at the resolution of these maps. Fifteen loci mapped on R8 have been anchored to loci within the maize genome on M1, M2, M4, M5, M7, M8, and M10. Segmental homeology was clearly defined between R8L and a centrically located segment on M1 by six loci. These loci accounted for the entire long arm of R8 and one of the two copies of the linkage group expected in the maize genome. Of the remaining nine markers anchored on R8 and in maize, all were scattered across the maize genome and were located in the centromeric regions of seven maize linkage groups.

The progenitor maize genome: On the basis of analysis of composite linkages exhibiting intragenomic homeology within modern maize, we deduced that the progenitor maize genome had eight independent linkage groups (Figure 4). Four linkage groups of progenitor maize (PM), PM1, PM3, PM4, and PM7 were inferred to represent compound linkage groups in the modern maize genome. The compound and composite arrangements of PM1, PM3, PM4, and PM7 can be identified within the intragenomic paired homeology of M3-M8, M1-M5 and M1-M9, M2-M10, and M2-M7, respectively. Relative to gene order and segment orientation in rice, inversions were noted on PM1 and PM6 (Table 1). Two inversions were located within composite linkage groups (PM1 and PM7) and all involved regions homeologous to a rice centromeric interval (Figures 1 and 4).

Comparisons between sorghum, sugarcane, and progenitor maize served to identify structural rearrangements that predate divergence of the Panicoideae. Detailed comparisons between sorghum and maize based on many of the loci detected by ISU probes utilized in the maize-rice comparison will be presented to further define the Panicoideae interval (W. Woodman and M. Lee, unpublished results). PM3 and PM7 are common to sorghum and indicate that two compound linkage groups were present in the progenitor Panicoideae and account for the reduction in base chromosome number from $x = 12$ (ancestral genome) to $x = 10$ (Panicoideae). The base number of common chromosomes in Panicoideae genomes is inferred to have been further reduced to eight in the progenitor maize genome by the formation of two novel compound linkages, PM1 and PM4. Based on these two compound linkages, PM1 and PM4, and an inversion on PM6, the maize progenitor genome ($x = 8$) was structurally differentiated from sorghum.

Conservation of compound progenitor linkages and

TABLE 2

Genomic rearrangements in the maize genome that occurred subsequent to the tetraploidization event

Event designation[a]	Last event in tetraploid genome structure	Type of chromosomal rearrangement	Homeologous group/segment	Breakpoint(s)/fusion point(s)	Resulting chromosome arrangement
M INS-1	MDE	Insertion	PM8 → PM3	*bcd207-rz500*/*rz500-cdo464*, *bcd207-cdo99*	M1 composite
M FS-1	MDE	Fission at centromere(?)	PM3	*isu62-cdo665*/NFP	PM3a-PM3b
M F-1	MDE	End-to-end fusion	PM4-PM7	NBP/*isu109*-TELO	M2 composite
M F-2	MDE	End-to-end fusion	PM2-PM5	NBP/TELO(*rz446*)-TELO(*isu60*)	M4 compound
M F-3	MDE	End-to-end fusion	PM3b-PM2	NBP/*isu62b*-TELO	M5 composite
M F-4	MDE	End-to-end fusion	PM3a-PM6	NBP/*cdo665*-TELO	M9 compound
M INS-2	MDE	Insertion (or fusion)	PM5 → PM1	TELO(*cdo1338*-CENT1S-*rz995b*), TELO(*csu25-cdo394*)	M3 composite
M TR-1	MDE	Translocation	R5L → PM6	DISTAL TO *cdo105*(*isu111*)/*isu111-rz144b*	M6 compound
M2 INV-1	Composite M2	Inversion	R11/12	*isu127-isu54*, *cdo385-csu109*/*csu109-isu127*, *isu151-isu69*	M2
M3 INV-1	Composite M3	Inversion	R11/12	*cdo344*(12L)-*csu25*(11L)	M3
M4 INV-1	MDE	Inversion	R11/12	*cdo127b*(11S)-*isu127*(11S)/CENTROMERE	M4
M4 INV-2	Compound M4	Inversion	R11/12	*uaz145*(11C)-*cdo127b*(11S)	M4
M8 INV-1	MDE	Inversion	R1S	*rz995a*-CENTROMERE	M8
M9 INV-1	Compound M9	Inversion	R3S, R6	*cdo665b-cdo672b*/*cdo78-rz588*	M9

[a] M, maize; R, rice; INS, insertion; INV, inversion; FS, fission; MDE, maize duplication event; TR, translocation; NBP, no breakpoint; NFP, no fusion point.

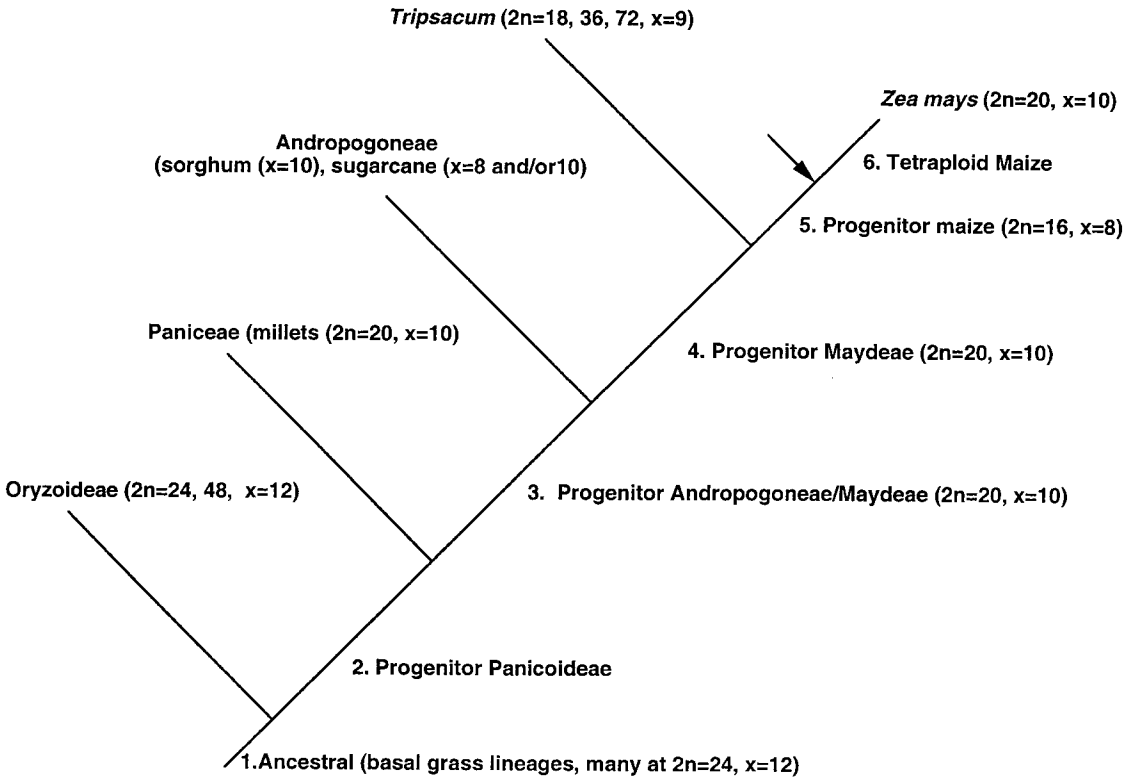

Figure 3.—Evolutionary intervals based on divergence points along the maize lineage. Divergent taxa are noted at each node. Inferred genome structures within the internodes represent progenitor taxa to subsequent lineages.

inversions in modern maize linkage groups suggested that the tetrasomic state arose either from a cross between species sharing these composite linkages and inversions or via autopolyploidy. Of the 8 progenitor linkage groups duplicated (16 groups after polyploidization), only M7 (PM7), M8 (PM1), and M10 (PM4) have not been further combined subsequent to polyploidization. Both copies of the duplicated PM2, PM3, PM5, PM6, PM8 formed novel compound or composite linkages in modern maize. Relative to ancestral components, all modern maize linkage groups are compound or composite and these linkage groups were fixed over many divergence points in the maize lineage.

The modern maize genome: Due to genome-wide duplication and numerous chromosomal rearrangements and fusions, direct comparison between rice and maize is facilitated by comparing maize linkage groups to rice on the basis of segmental homeology rather than component linkage groups. Points of chromosome breakage and fusion are estimated on rice linkage groups on the basis of informative markers in one or both of the corresponding duplicated segments in maize. Often only one of the two segments is noted in any comparison as presented in the following three examples for M2, M3, and M4.

Maize chromosome 2: Homeology was assigned between M2 and R4, R7, R9, and R11/12 by the anchoring of 11, 12, 4, and 5 loci on these linkage groups, respectively. Ten incongruent loci corresponding to loci on R2, R3, R5, R6, and R7 were scattered along the linkage group in no discernible pattern. The short arm of M2 from *cdo36b* through *isu142* contained 11 loci anchoring M2-R4 homeology and accounted for >95% of R4 within the interval.

A portion of the long arm of M2 is composed of segments homeologous to R7 and R9 and was considered a duplicate counterpart of M7. The association of R7 and R9 components in each linkage group is representative of a progenitor maize linkage group. In progenitor maize, an inversion was deduced on the basis of the inverted marker order on both M2 and M7 relative to R7, including the rearrangement of *cdo533* and *cdo385* (both in the centric region of R7) to a distal location on M2 and M7. After accounting for this inversion in both maize linkage groups, the position of the insertion was estimated to lie between *cdo59* and *uaz200*, two anchored loci within the long arm of R7.

The long arm of M2 is complicated by several homeology transitions resulting from composite linkage formation and a subsequent intrachromosomal rearrangement that occurred after the MDE. A segment homeologous to R11/12 was detected on M2L by the positioning of five loci detected by five maize probes. Two R11/12 loci (*isu69* and *isu54*) were located near the centromere of M2, interstitially located between segments homeologous to R4S and R7L, and the other three R11/12 loci

(*isu109* and *isu127*, R5S) were interstitial to segments homeologous to R7L and R7S at the distal end of M2L. A likely inversion was noted by the separation of *isu109*, *isu127*, and the R5S locus from *isu69* and *isu54* that would result from breakpoints within the segments homeologous to R11/12 and R7. Both breakpoints occurred in centromeric intervals: one breakpoint with the segment homeologous to R7 occurred in an interval homeologous to R7C and the other breakpoint was located at or near the M2C. Such an inversion would have involved the majority of the long arm of M2 and may have resulted in the transfer of the R11C/12C within this segment to a more distal position and the transfer of the region orthologous to R7C to a centromeric position. Ordering of the corresponding segment duplicated on M7 and the corresponding progenitor maize consensus indicates that the inversion is unique to M2 and happened subsequent to the MDE, perhaps as a result of composite linkage group formation resulting from an end-to-end fusion between PM4 and PM7.

Maize chromosome 3: Maize chromosome 3 was observed to be the most complex chromosome in maize due to the presence of segments homeologous to R1, a portion of R5, and R11/12. Direct comparison of M3 and M8, based on observed homeology, reveals that each has an arrangement reflecting the insertion of R5 into R1 and presumably both arose from a PM1 linkage group.

The short arm of M3 is homeologous to R1S within the region delimited by *csu75* and *rz995*. The order of comparatively mapped markers on M3S is uncertain due to difficulties associated with positioning markers to intervals and to a deficiency of comparatively mapped markers. Tight linkage and small population size precludes precise ordering of markers within the centromeric region, but a similar structural arrangement consisting of R5 inserted into R1 in the centromeric region was observed in the centromeric region of M8. The M8 arrangement was considered the duplicate counterpart of M3 and served as a better template to infer the ancestral arrangement.

Close inspection of marker order on M8 revealed markers anchoring M8 to R5S were separated by an interval corresponding to R5L. Separation of portions of segments homeologous to R5S indicated an inversion had occurred within the M8 lineage relative to rice. On M3, markers corresponding to one set of R5S markers separated by the M8 inversion event were detected in the centromeric region. Therefore, we deduce that the segments homeologous to R5 on M3 and M8 were inverted prior to duplication of PM1. It was not possible to determine if the inversion was associated with compound linkage group formation resulting in PM1. If M3 is presumed to have arisen from PM1, the segment corresponding to the long arm of R5L that appears on M6L would have been translocated subsequent to the MDE.

A third homeology transition on M3 involves segments homeologous to R11 or R12 and R1L and is demarcated by *cdo344a-cdo459* and *rz296b*. The position of *csu25* and *cdo1338a* on the end of M3L and a reversal of order of markers of R1L indicate an inversion has occurred within M3 and is likely to have occurred subsequent to the insertion of R11/12 in a position at or adjacent to the centromeric region on M3. This is further supported by the observation that the markers along the large segment corresponding to R1L appear in reverse order to those from R1S, which comprises the short arm of M3 and PM1.

Maize chromosome 4: Maize chromosome 4 can be aligned to rice by 55 loci and is composed of segments homeologous to R2 and R11/12. Of the 55 loci, 22 were located on M4 and R2, 10 were on M4, and 12 were on R11/12. An R11/12-R2 homeology transition occurs at or near the centromere in an interval between *isu127b* and *cdo1380*. Within this centromeric interval, 15 incongruent loci on M4 indicated point relationships with R1, R3, R4, R5, R6, R7, and R8. Of these incongruent loci, all but one were located near the centromere and the first homeology transition region with no apparent positional relationships to other portions of the maize or rice genomes.

The order of loci in segments homeologous to R2 and R11 indicated that end-to-end fusions (PM2 and PM5) and two inversion events were necessary to account for the modern structural arrangement of M4. The breakpoints of the most recent inversion were located to the intervals delimited by *uaz115-cdo1380b* (corresponding to the aligned rice centromeric region in the segment homeologous to R2) and *uaz145-cdo127a* (corresponding to the centromeric region of R11 and located in the centromeric region of M4). Reversion of this interval leads to reconstitution of the linear order corresponding to R2 and juxtaposition of a small segment on the M4L that is homeologous R11/12 to the remainder of the segment composing the short arm of M4. The first inversion on M4 occurred within the short arm of M4 (homeologous to R11) with breakpoints at or near the M4 centromere (telomere R11S) and between *cdo520* and *cdo534*. Neither of these inversions is present on the duplicated M2, M3, and M10, or M5 segments that are homeologous to R2 and R11/12, respectively, and therefore they were presumed to have occurred subsequent to the MDE.

Hybridization of maize probes in grasses other than rice and maize: Comparative mapping of the grasses is based upon the use of common probes, usually derived from conserved but anonymous genes, to map related loci and to anchor linkage relationships in different species, tribes, genera, or clades (Van Deynze *et al.* 1998). Maize probes utilized in this study were evaluated for copy number and strength of hybridization signal across the grasses and two close relatives of the Poaceae. More than 90% of maize probes surveyed on sugarcane,

sorghum, and rice produced strong clear signal. Hybridization efficiency dropped dramatically within the Pooideae to 60% (wheat) or below (barley and oat). Hybridization efficiency of maize probes on bamboo (79%), Joinvillea (80%), and Flagelleria (68%) demonstrated the efficacy of maize cDNAs to detect and map loci beyond the grass family.

DISCUSSION

Genome rearrangements in the grasses: An important consideration in comparing genome structure across the grasses is the role of chromosomal rearrangement in speciation because, although chromosome breakage and rearrangements are common, retention of rearrangements requires selection and/or the ability to buffer genetic deficiencies. Of the domesticated grasses mapped to date, the structure of the maize genome is the most complex in the Gramineae due to fixation of 10 composite linkage groups, at least 10 inversions, one nonreciprocal translocation, a polyploidization event, and aneuploidy consisting of one duplicate linkage group relative to rice. Early fixation of some of these events has served to separate the Panicoids from a progenitor grass similar to rice, and subsequent events separated maize from millet, sugarcane, sorghum, and Tripsacum. In addition, rearrangements below the level of resolution of the maps presented in this study have occurred, as suggested by the numerous markers whose map position does not coincide with currently defined regions of synteny (Figure 2), and are expected to be resolved as genomic resolution increases (Chen *et al.* 1997).

The compound nature of linkage groups in all the domesticated grasses except rice (Ahn *et al.* 1993; Kurata *et al.* 1994; Van Deynze *et al.* 1995a,b) has led to the proposition that rice has the most basic genome. The incompletely dysploidic arrangement of the 12 grass components into species with a basic chromosome number of 7 (Triticeae, Aveneae), 8 (Progenitor maize and *Saccharum spontaneum*), and 10 (Sorghum, *S. officinarum*) accounts for the reduction of the base chromosome number in each lineage and serves to clarify phylogenetic relationships. The unique events involving components homeologous to R8 and R10 in each lineage (Ahn and Tanksley 1993; Van Deynze *et al.* 1995a,b) add further support for rice as the most basic genome because a link to chromosome fission has not been developed that fits the chromosome complements of the Pooids and the Panicoids. Indeed, a reductional trend in chromosome number is apparently irreversible (Stebbins 1950) as no evidence has been presented within the domesticated grasses of fixation of two linkage groups that have arisen from a compound linkage group that fits with evolutionary direction. While the mechanisms enforcing the reductional trends in grass genomes are unknown, gross structural rearrangement altering the number, function, and positioning of centromeres could be very important.

An obvious result of compound linkage group formation is the juxtaposition of two or more centromeres on a chromosome. Of the 10 inversions defined between maize and rice, all had breakpoints at or near a centromere in maize or an interval homeologous to a rice centromere, and at least 8 occurred within compound or composite linkage groups. Inversions on PM1, PM7, M2, M3, M4, and M5 were pericentric from the perspective of the inclusion of regions that were orthologous to rice centromeres. Pericentric inversions detected in this study would affect the physical position of centromeres and they may represent a mechanism of aggregating and/or dispersing centromeres or centromeric activity in compound and composite linkage groups.

The concept of the grass family as an incomplete dysploidic series of genomes provides an opportunity to evaluate the fixation rate for gross structural rearrangements. These rearrangements can be estimated on the basis of the data presented in this study as summarized in Figure 5. Divergence of Panicoids from Oryzoids has been estimated to be 60 mya (Figure 3, node 1). Based on RFLP linkage maps of sorghum (Chittenden *et al.* 1994; Pereira *et al.* 1994), rice, and maize, two compound linkage groups and two inversions formed in the progenitor Panicodeae genome and account for reduction in basic chromosome number from 12 to 10. A basic chromosome of 10 is common throughout the Panicoideae, although exceptions have been well documented in previous studies (Celarier 1956, 1957). Andropogoneae-Maydeae divergence is estimated at 25 mya (Gaut and Doebley 1997). Subsequent to this divergence (Figure 3, node 3), two additional compound linkages occurred accounting for reduction from $x = 10$ to $x = 9$ (Tripsacum) to $x = 8$ (progenitor maize). The polyploidization event in the maize lineage has been estimated to have occurred 16–25 mya (Gaut and Doebley 1997). Subsequent to this doubling, seven new compound or composite linkage groups, seven inversions, one translocation, and perhaps one deletion (R8) were fixed to account for the modern maize genome. Given the rate of chromosomal rearrangement prior to duplication, it is apparent that relatively rapid genome rearrangement occurred subsequent to the polyploidization event.

Tetraploid maize: Cytogenetic (reviewed in Molina and Naranjo 1987) and molecular evidence (Gottlieb 1982; Helentjaris *et al.* 1988; Ahn and Tanksley 1993) have supported the early contentions (Rhoades 1951) that maize is an ancient tetraploid that exhibits disomic inheritance. While cytogenetic studies have proposed 5 as the basic chromosome number (reviewed in Molina and Naranjo 1987), we see no evidence supporting this contention based upon comparative mapping. Our progenitor genome model presents evidence that the genome(s) that gave rise to a duplicated maize genome

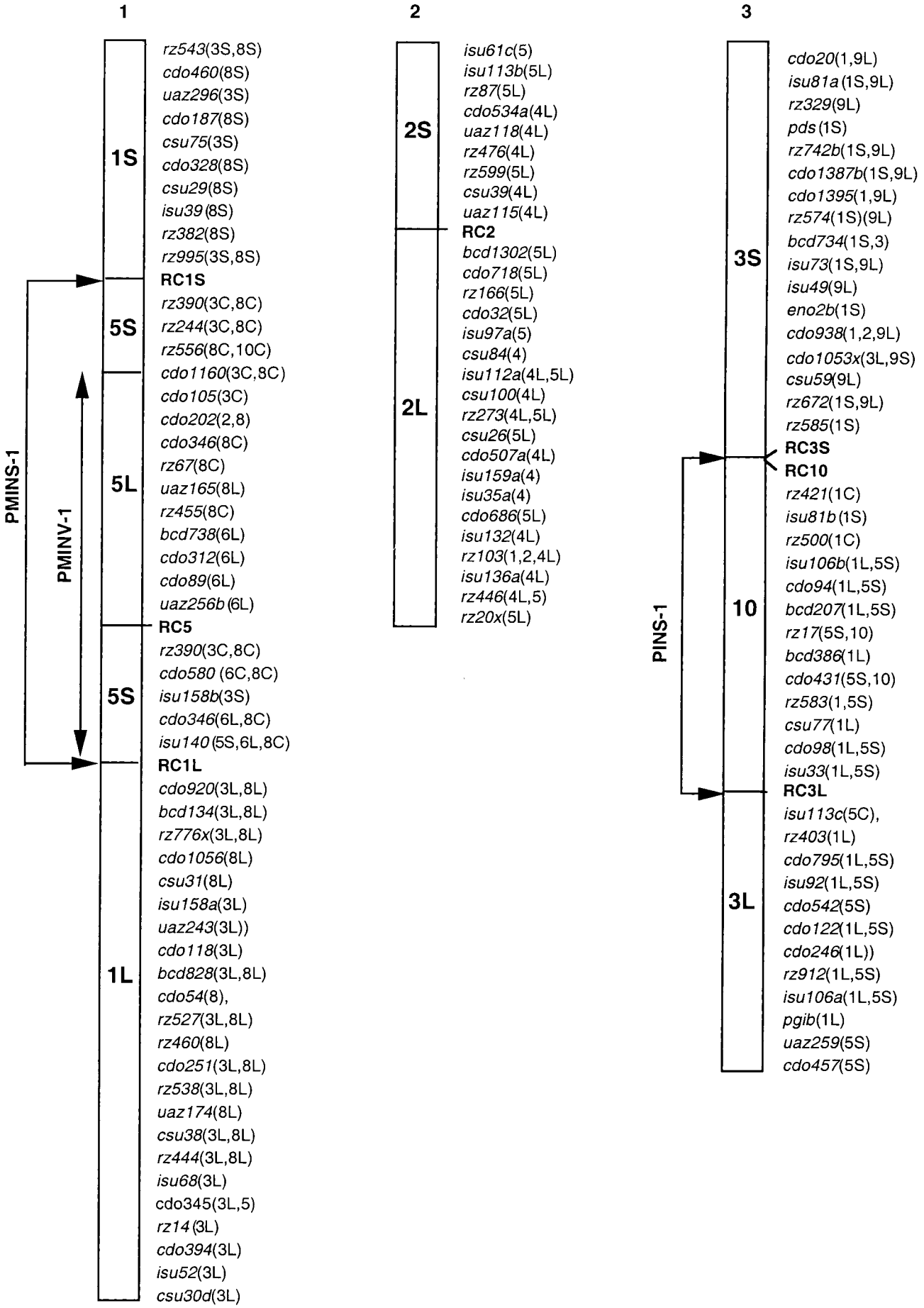
1
2
3
1S
5S
5L
5S
1L
2S
2L
3S
10
3L
PMINS-1
PMINV-1
PINS-1
rz543(3S,8S)
cdo460(8S)
uaz296(3S)
cdo187(8S)
csu75(3S)
cdo328(8S)
csu29(8S)
isu39(8S)
rz382(8S)
rz995(3S,8S)
RC1S
rz390(3C,8C)
rz244(3C,8C)
rz556(8C,10C)
cdo1160(3C,8C)
cdo105(3C)
cdo202(2,8)
cdo346(8C)
rz67(8C)
uaz165(8L)
rz455(8C)
bcd738(6L)
cdo312(6L)
cdo89(6L)
uaz256b(6L)
RC5
rz390(3C,8C)
cdo580(6C,8C)
isu158b(3S)
cdo346(6L,8C)
isu140(5S,6L,8C)
RC1L
cdo920(3L,8L)
bcd134(3L,8L)
rz776x(3L,8L)
cdo1056(8L)
csu31(8L)
isu158a(3L)
uaz243(3L))
cdo118(3L)
bcd828(3L,8L)
cdo54(8),
rz527(3L,8L)
rz460(8L)
cdo251(3L,8L)
rz538(3L,8L)
uaz174(8L)
csu38(3L,8L)
rz444(3L,8L)
isu68(3L)
cdo345(3L,5)
rz14(3L)
cdo394(3L)
isu52(3L)
csu30d(3L)
isu61c(5)
isu113b(5L)
rz87(5L)
cdo534a(4L)
uaz118(4L)
rz476(4L)
rz599(5L)
csu39(4L)
uaz115(4L)
RC2
bcd1302(5L)
cdo718(5L)
rz166(5L)
cdo32(5L)
isu97a(5)
csu84(4)
isu112a(4L,5L)
csu100(4L)
rz273(4L,5L)
csu26(5L)
cdo507a(4L)
isu159a(4)
isu35a(4)
cdo686(5L)
isu132(4L)
rz103(1,2,4L)
isu136a(4L)
rz446(4L,5)
rz20x(5L)
cdo20(1,9L)
isu81a(1S,9L)
rz329(9L)
pds(1S)
rz742b(1S,9L)
cdo1387b(1S,9L)
cdo1395(1,9L)
rz574(1S)(9L)
bcd734(1S,3)
isu73(1S,9L)
isu49(9L)
eno2b(1S)
cdo938(1,2,9L)
cdo1053x(3L,9S)
csu59(9L)
rz672(1S,9L)
rz585(1S)
RC3S
RC10
rz421(1C)
isu81b(1S)
rz500(1C)
isu106b(1L,5S)
cdo94(1L,5S)
bcd207(1L,5S)
rz17(5S,10)
bcd386(1L)
cdo431(5S,10)
rz583(1,5S)
csu77(1L)
cdo98(1L,5S)
isu33(1L,5S)
RC3L
isu113c(5C),
rz403(1L)
cdo795(1L,5S)
isu92(1L,5S)
cdo542(5S)
cdo122(1L,5S)
cdo246(1L))
rz912(1L,5S)
isu106a(1L,5S)
pgib(1L)
uaz259(5S)
cdo457(5S)

Figure 4.

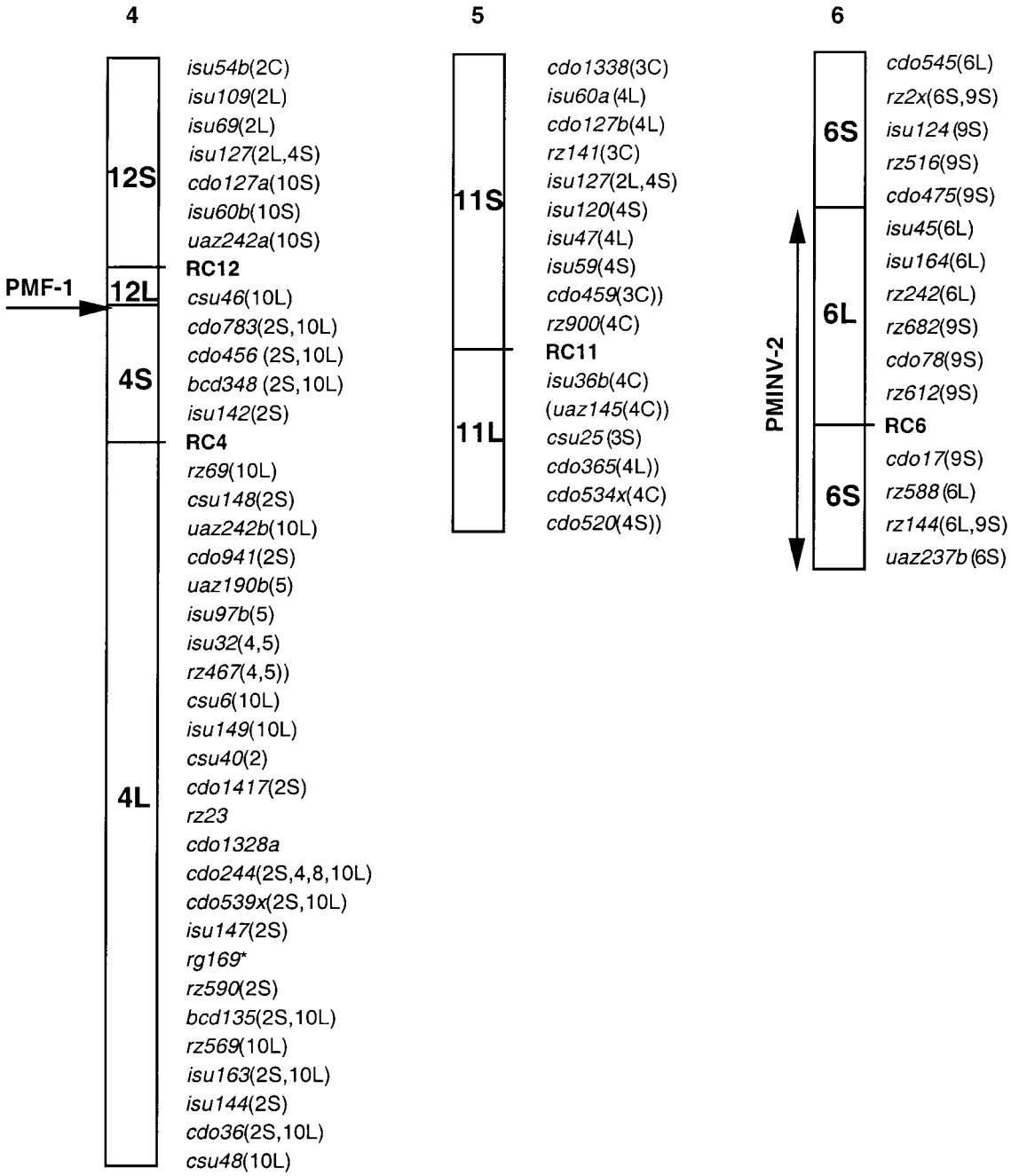

Figure 4.—Inferred consensus map of a progenitor maize genome based on associations of conserved chromosomal segments within the rice and maize genomes. Numbered blocks represent rice linkage groups that correspond to the progenitor maize segment. All rearrangement breakpoints were estimated on the basis of order within duplicated segments within the maize genome and rearrangements noted relative to the rice linkage segments (as summarized in Tables 1 and 2). Numbers in parentheses behind each marker indicate chromosomal positions in the maize genome. Designations in boldface type indicate putative positions of rice centromeres (RC) at rearrangement breakpoints.

had already diverged from sorghum by two novel compound linkage groups and at least two inversions. We propose the polyploidization event would have involved one (auto) or two (allo) species with these linkages in common. The relatedness of these genomes was based on common compound linkages in modern maize, but the unique events associated with these linkage groups indicate that maize was derived via autopolyploidy or via a cross between two species more closely related to each other than either was to sorghum or sugarcane.

After polyploidization, the progenitor maize genome switched from tetrasomic to disomic inheritance. The switch has been proposed to have been segmental based on grouping sequence divergence of duplicated loci (Gaut and Doebley 1997). While autopolyploidy with segmental reduction from polysomy has not been clearly

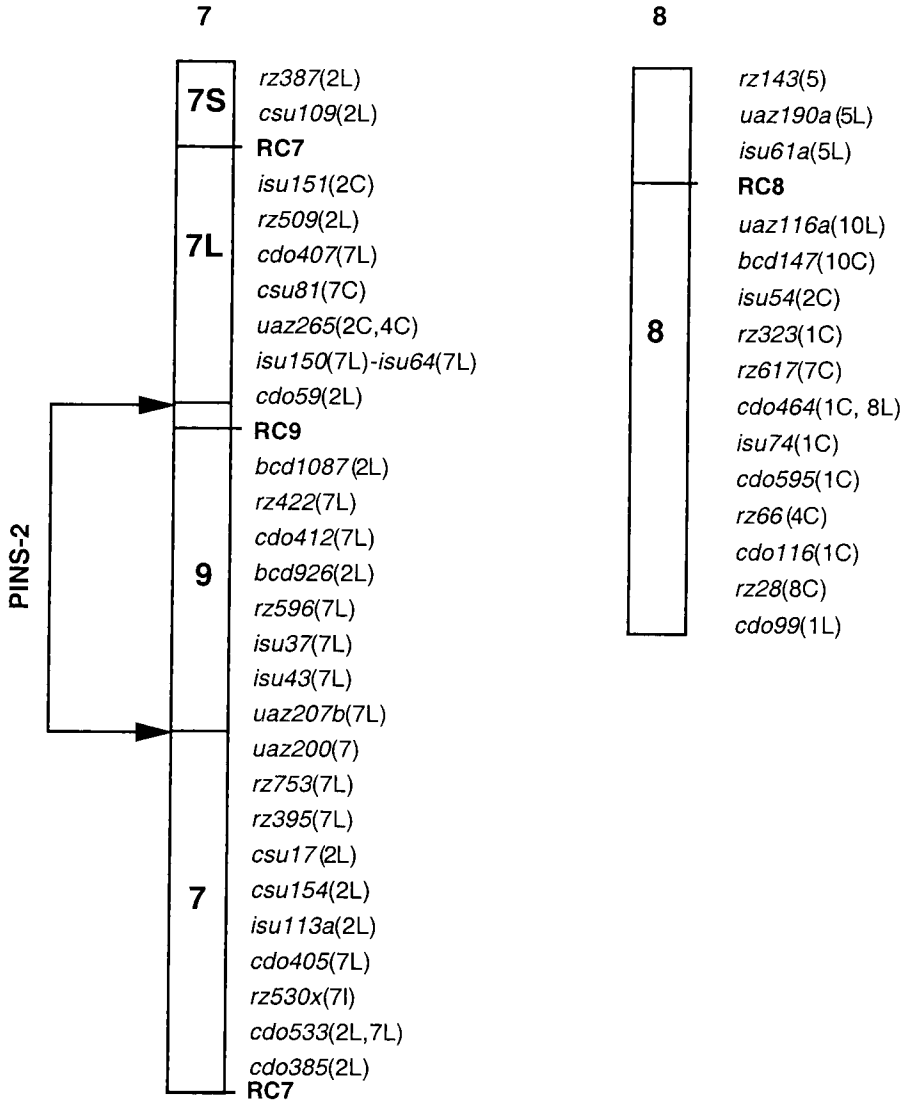

Figure 4.—*Continued.*

demonstrated in the domesticated grasses, sugarcane has been postulated to be an autooctoploid ($x = 8$; Da Silva *et al.* 1993) that exhibits some preferential pairing (Al-Janabi *et al.* 1994; Ming *et al.* 1998). The autopolyploid origin and segmental reduction to disomic inheritance proposed for maize parallels observations of the sugarcane genome. Chromosome rearrangements may have contributed to the shift from tetrasomic to disomic inheritance by introducing structural heterology among linkage groups. Formation of compound or composite linkages followed by inversion events may have reduced pairing between duplicated segments and contributed to the transition to disomic inheritance. Future sequence comparisons between paralogous loci, like those of Gaut and Doebley (1997), may lead to an understanding of the role of chromosome rearrangements in the transition to disomic inheritance.

The tribe Maydeae has seven genera: Zea, Tripsacum, Coix, Trilobachne, Polytoca, Sclerachne, and Chionachne (Celarier 1957). Most genera have a haploid chromosome number based on 10 similar to maize (*i.e.*, $2n = 20$ or 40), except for *Tripsacum* spp. ($2n = 18$, 36, or 72) and Coix ($2n = 10$, 20). Coix is a unique member of the tribe and has been proposed to represent the $x = 5$ progenitor of the Maydeae lineage. Our model of progenitor maize, coupled with an assumption that changes in basic chromosome number are only reductional, would predict that Coix is dramatically rearranged relative to maize due to novel compound and composite linkages and, rather than being a progenitor, is highly divergent from other Maydeae genomes. *Tripsacum* spp. genomes range from $2n = 18$ to $2n = 72$, which indicates 9 as the base number. A future comparison between Sorghum, Tripsacum, and the inferred progenitor maize might identify a compound linkage group common only to maize and Tripsacum and provide the intermediate genome structure between progenitor maize ($x = 8$) and sorghum ($x = 10$).

The authors gratefully acknowledge Jerrold Davis for many useful conversations and suggestions in preparing this manuscript, Matthew Blair and Kevin Livingstone for their critical reviews of the manuscript,

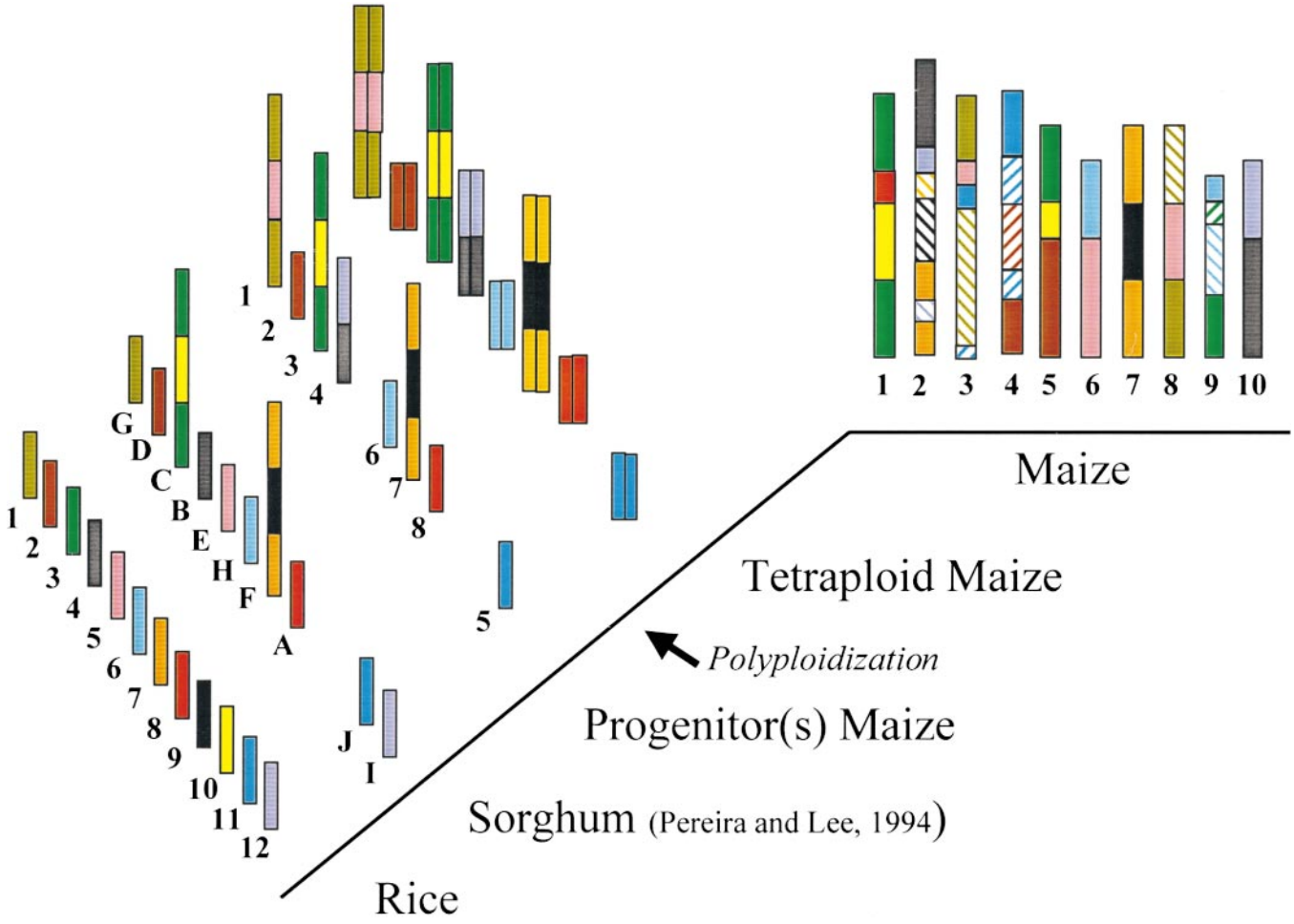

Figure 5.—Genome structures of domesticated grasses along the continuum of maize evolution. The rice genome, with 12 basic chromosome structures represented as color blocks, marks the divergence of the Orzyoideae and Panicoideae subfamilies. Sorghum, progenitor maize, duplicated progenitor maize, and modern maize are drawn relative to rice chromosome structures. Chromosome inversions inferred in modern maize are indicated as hatched shading of blocks representing the inverted linkage segments. To faciliate structural comparisions Sorghum mapping data from Pereira *et al.* (1994) was included and compared to rice on the basis of ISU markers.

Carole Morehouse and Lois Swales for help in formatting the manuscript and figures, and the financial support of the United States Department of Agriculture Plant Genome Initiative (USDA NRI grants 94-37300-0324 to M.S., 94-37310-0661 to S.M., and 97-35300-4939 to M.L.); International Atomic Energy Agency (research contract 8494/R1 to M.L.); and Cooperative State Research, Education and Extension Service (NYC 149-401 to S.M.).

LITERATURE CITED

Ahn, S. N., and S. D. Tanksley, 1993 Comparative linkage maps of the rice and maize genomes. Proc. Natl. Acad. Sci. USA **90:** 7980–7984.

Ahn, S., J. A. Anderson, M. E. Sorrells and S. D. Tanksley, 1993 Homeologous relationships of rice, wheat and maize chromosomes. Mol. Gen. Genet. **241:** 483–490.

Al-Janabi, S. M., R. J. Honeycutt and B. W. S. Sobral, 1994 Chromosome assortment in Saccharum. Theor. Appl. Genet. **89:** 959–963.

Arumuganathan, K., and E. D. Earle, 1991 Nuclear DNA content of some important plant species. Plant Mol. Biol. Rep. **9:** 208–218.

Bennetzen, J. L., and M. Freeling, 1993 Grasses as a single genetic system: genome composition, collinearity, and compatibility. Trends Genet. **9:** 259–261.

Burr, B., F. A. Burr, K. H. Thompson, M. C. Albertsen and C. W. Stuber, 1988 Gene mapping with recombinant inbreds in maize. Genetics **118:** 519–526.

Burr B., F. A. Burr and E. C. Matz, 1993 Maize molecular map (*Zea mays*) 2*n* = 20, pp. 190–203 in *Genetic Maps*, edited by S. J. O'Brien. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Causse, M., T. M. Fulton, Y. G. Cho, S. N. Ahn, J. Chunwongse *et al.*, 1994 Saturated molecular map of the rice genome based on an interspecific backcross population. Genetics **138:** 1251–1274.

Celarier, R. P., 1956 Cytotaxonomy of the Andropogoneae 1. Subtribes Dimeriinae and Saccharinae. Cytologia **21:** 272–291.

Celarier, R. P., 1957 Cytotaxonomy of the Andropogoneae 2. Subtribes Ischaeminae, Rottboellinae, and the Maydeae. Cytologia **22:** 160–183.

Chao, S., C. Baysdorfer, O. Heredia-Diaz, T. Musket, G. Xu *et al.*, 1994 RFLP mapping of partially sequenced leaf cDNA clones in maize. Theor. Appl. Genet. **88:** 717–721.

Chen, M., P. San Miguel, A. C. de Oliveira, S. S. Woo, H. Zhang *et al.*, 1997 Microcolinearity in *sh2*-homologous regions of the maize, rice, and sorghum genomes. Proc. Natl. Acad. Sci. USA **94:** 3431–3435.

Chittenden, L. M., K. F. Schertz, Y. R. Lin, R. A. Wing and A. H.

Paterson, 1994 A detailed RFLP map of Sorghum bicolor × S. propinquum suitable for high-density mapping suggests ancestral duplication of Sorghum chromosomes or chromosomal segments. Theor. Appl. Genet. **87:** 925–933.

Clark, L. G., W. Zhang and J. F. Wendel, 1995 A phylogeny of the grass family (Poaceae) based on ndhF sequence data. Syst. Bot. **20:** 436–460.

Da Silva, J. A. G., M. E. Sorrells, W. L. Burnquist and S. D. Tanksley, 1993 RFLP linkage map of *Saccharum sponteanum*. Genome **36:** 782–791.

Davis, G., M. D. McMullen, E. H. Coe and M. Polacco, 1996 Working maps. Maize Genet. Coop. Newslett. **70:** 118–132.

Davis, J. I., 1995 A phylogenetic structure for the monocotyledons, as inferred from chloroplast DNA restriction site variation, and a comparison of measures of clade support. Syst. Bot. **20:** 503–527.

Devos, K. M., and M. D. Gale, 1997 Comparative genetics in the grasses. Plant Mol. Biol. **35:** 3–15.

D'Hont, A., L. Grivet, P. Feldmann, P. S. Rao, N. Berding *et al.*, 1996 Characterization of the double genome structure of modern sugar cane cultivars (*Saccharum* spp.) by molecular cytogenetics. Mol. Gen. Genet. **250:** 405–413.

Doyle, J. J., J. I. Davis, R. J. Soreng, D. Garvin and M. J. Anderson, 1992 Chloroplast DNA inversions and the origin of the grass family (Poaceae). Proc. Natl. Acad. Sci. USA **89:** 7722–7726.

Duvall, M. R., M. T. Clegg, M. W. Chase, W. D. Clark, W. J. Kress *et al.*, 1993 Phylogenetic hypotheses for the monocotyledons constructed from rbcL sequence data. Ann. Mo. Bot. Gard. **80:** 607–619.

Flavell, R. B., M. D. Bennett, J. B. Smith and D. B. Smith, 1974 Genome size and the proportion of repeated nucleotide sequence DNA in plants. Biochem. Genet. **12:** 257–269.

Gaut, B. S., and J. F. Doebley, 1997 DNA sequence evidence for the segmental allotetraploid origin of maize. Proc. Natl. Acad. Sci. USA **94:** 6809–6814.

Gottlieb, L. D., 1982 Conservation and duplication of isozymes in plants. Science **216:** 373–380.

Harushima, Y., M. Yano, A. Shomura, M. Sato, T. Shimano *et al.*, 1998 A high-density rice genetic linkage map with 2275 markers using a single F2 population. Genetics **148:** 479–494.

Helentjaris, T., D. Weber and S. Wright, 1988 Identification of the genomic locations of duplicate nucleotide sequences in maize by analysis of restriction fragment length polymorphisms. Genetics **118:** 353–363.

Kurata, N., Y. Nagamura, K. Yamamoto, Y. Harushima, N. Sue *et al.*, 1994 A 300 kilobase interval genetic map of rice including 883 expressed sequences. Nature Genet. **8:** 365–372.

Kurata, N., Y. Umehara, H. Tanoue and T. Sasaki, 1997 Physical mapping of the rice genome with YAC clones. Plant Mol. Biol. **35:** 101–113.

Lander, E. S., P. Green, J. Abrahamson, A. Barlow, M. J. Daly *et al.*, 1987 MAPMAKER: an interactive computer package for maps of experimental and natural populations. Genomics **1:** 174–181.

Matz, E. C., F. A. Burr and B. Burr, 1994 Molecular map based on TxCM and CoxTx recombinant inbred families. Maize Genet. Coop. Newslett. **68:** 198–208.

McCouch, S. R., G. Kochert, Z. H. Yu, Z. Y. Wang, G. S. Khush *et al.*, 1988 Molecular mapping of rice chromosomes. Theor. Appl. Genet. **76:** 815–829.

Ming, R., S. C. Liu, Y. R. Lin, J. Da Silva, W. A. Wilson, *et al.*, 1998 Detailed alignment of saccharum and sorghum chromosomes: comparative organization of closely related diploid and polyploid genomes. Genetics **150:** 1663–1682.

Molina, M. C., and C. A. Naranjo, 1987 Cytogenetic studies in the genus *Zea*: 1. Evidence for five as the basic chromosome number. Theor. Appl. Genet. **73:** 542–550.

Moore, G., K. M. Devos, Z. Wang and M. D. Gale, 1995 Grasses, line up and form a circle. Curr. Biol. **5:** 737–739.

Nagamura, Y., B. A. Antonio and T. Sasaki, 1997 Rice molecular genetic map using RFLPs and its applications. Plant Mol. Biol. **35:** 79–87.

Panaud, O., X. Chen and S. R. McCouch, 1996 Development of microsatellite markers and characterization of simple sequence length polymorphism (SSLP) in rice (*Oryza sativa* L.). Mol. Gen. Genet. **252:** 597–607.

Paterson, A. H., Y.-R. Lin, S. Li, K. F. Schertz, J. F. Doebley *et al.*, 1995 Convergent domestication of cereal crops by independent mutations at corresponding genetic loci. Science **269:** 1714–1717.

Pereira, M. G., M. Lee, P. Bramel-Cox, W. Woodman, J. Doebley *et al.*, 1994 Construction of an RFLP map in sorghum and comparative mapping in maize. Genome **37:** 236–243.

Rhoades, M. M., 1951 Duplicate genes in maize. Am. Nat. **85:** 105–110.

San Miguel, P., A. Tikhonov, Y. K. Jin, N. Motchoulskaia, D. Zakharov *et al.*, 1996 Nested retrotransposons in the intergenic regions of the maize genome. Science **274:** 765–768.

Shen, B., N. Carneiro, I. Torres-Jerez, B. Stevenson, T. McCreery *et al.*, 1994 Partial sequencing and mapping of clones from two maize cDNA libraries. Plant Mol. Biol. **26:** 1085–1101.

Singh, K., T. Ishii, A. Parco, N. Huang, D. S. Brar *et al.*, 1996 Centromere mapping and orientation of the molecular linkage map of rice (*Oryza sativa L.*). Proc. Natl. Acad. Sci. USA **93:** 6163–6168.

Song, W. Y., G. L. Wang, L. Chen, H. S. Kim, B. Wang *et al.*, 1995 The rice resistance gene, Xa-21, encodes a receptor-like protein kinase. Science **270:** 1804–1806.

Soreng, R. J., and J. I. Davis, 1998 Phylogenetics and character evolution in the grass family (Poaceae): simultaneous analysis of morphological and chloroplast DNA restriction site character sets. Bot. Rev. **64:** 1–85.

Stebbins, G. L., 1950 Evolutionary trends I: the karyotype, pp. 442–475 in Variation and Evolution in Plants, edited by L. C. Dunn. Columbia University Press, New York.

Umehara, Y., A. Inagaki, H. Tanoue, T. Yasukochi, Y. Nagamura *et al.*, 1994 Construction and characterization of rice YAC libraries for physical mapping. Mol. Breed. **1:** 79–89.

Van Deynze, A. E., J. Dubcovsky, K. S. Gill, J. C. Nelson, M. E. Sorrells *et al.*, 1995a Molecular-genetic maps for group 1 chromosomes of Triticeae species and their relation to chromosomes in rice and oat. Genome **38:** 45–59.

Van Deynze, A. E., J. C. Nelson, S. E. Harrington, E. S. Yglesias, D. Braga *et al.*, 1995b Comparative mapping in grasses. Wheat relationships. Mol. Gen. Genet. **248:** 744–754.

Van Deynze, A. E., M. E. Sorrells, W. D. Park, N. M. Ayres, H. Fu *et al.*, 1998 Anchor probes for comparative mapping of grass genera. Theor. Appl. Genet. **97:** 356–369.

Veldboom, L. R., M. Lee and W. L. Woodman, 1994 Molecular marker-facilitated studies in an elite maize population. I. Linkage analysis and determination of QTL for morphological traits. Theor. Appl. Genet. **88:** 7–16.

Wang, Z. X., M. Yano, U. Yamanouchi, M. Iwamoto, L. Monna *et al.*, 1999 The Pi-b gene for rice blast resistance belongs to the nucleotide binding and leucine rich repeat class of plant disease resistance genes. Plant J. **19:** 55–64.

Yamamoto, K., and T. Sasaki, 1997 Large-scale EST sequencing in rice. Plant Mol. Biol. **35:** 135–144.

Yoshimura, S., U. Yamanouchi, Y. Katayose, S. Toki, Z.-X. Wang *et al.*, 1998 Expression of Xa1, a bacterial blight-resistance gene in rice, is induced by bacterial inoculation. Proc. Natl. Acad. Sci. USA **95:** 1663–1668.

Zhang, H. B., and R. A. Wing, 1997 Physical mapping of the rice genome with BACs. Plant Mol. Biol. **35:** 115–127.

Communicating editor: J. A. Birchler