Grasses, line up and form a circle

The genomes of six major grass species can be aligned by dissecting the individual chromosomes into segments and rearranging these linkage blocks into highly similar structures.

Cereals are the staple diet of most of the world's population and the study of their genetics, physiology, biochemistry and agronomy has preoccupied cereal biologists for most of the present century. We have all, however, been specialists in our own favoured crop species, and research has tended to proceed in parallel, with little academic interaction or comparison of species. In the late 1980s, when novel molecular technology allowed us, for the first time, to make dense chromosome maps, it was expected that the individual species maps would be unique. But the use of a new tool for comparing the genomic maps of different cereal crops — 'rice linkage segment' analysis — shows that this is not the case.

Comparative mapping in cereals

In 1986, when we began constructing restriction fragment length polymorphism (RFLP) maps for each of the three genomes in hexaploid (bread) wheat, we imagined that they would turn out to be different. It soon became clear, however, that the maps for each of the three diploid ancestors of hexaploid wheat — *Triticum urartu*, *Aegilops squarrosa* and the still unidentified donor of the B genome — are actually remarkably similar [1]. The ability of the

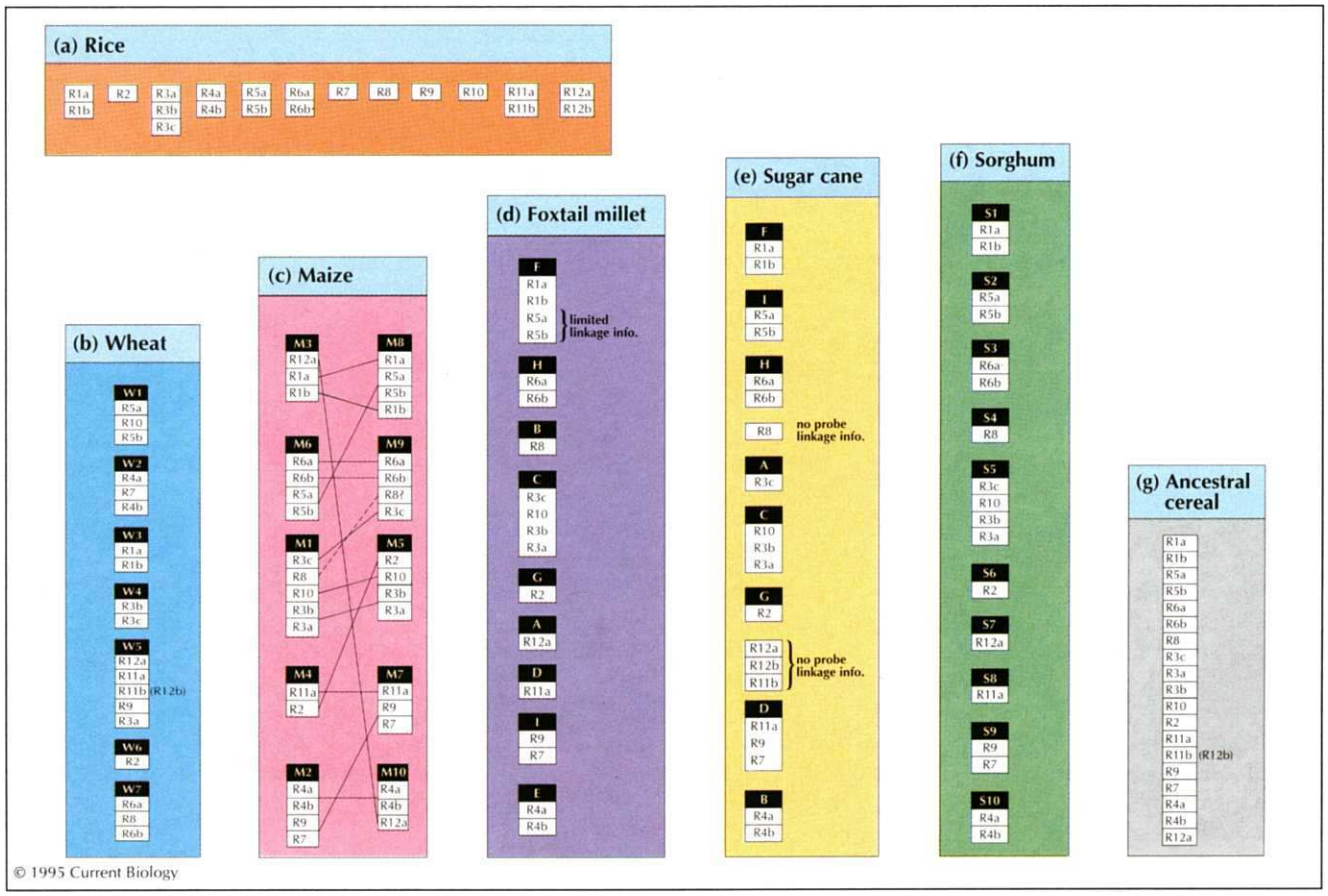

Fig. 1. Comparisons of cereal genome evolution based on rice linkage segments. **(a)** Rice chromosomes dissected into linkage blocks [3]. **(b)** Wheat, **(c)** maize, **(d)** foxtail millet, **(e)** sugar cane and **(f)** sorghum chromosomes represented as 'rice blocks' on the basis of homology and/or conservation of gene order. Connecting lines indicate duplicated segments within the maize chromosomes. **(g)** An ancestral 'single chromosome' reconstructed on the basis of these linkage blocks [3]. The designation of the different sorghum linkage groups has varied between laboratories and publications. The data in several publications [8–11] have been assembled in the form presented by Grivet [12,13]. For the purpose of this figure, the sorghum linkage groups have then been assigned (S1–S10) on the basis of their rice segments and the order of rice segments in the ancestral 'chromosome'. Open boxed linkage groups indicate that the order of segments within each linkage group needs further confirmation.

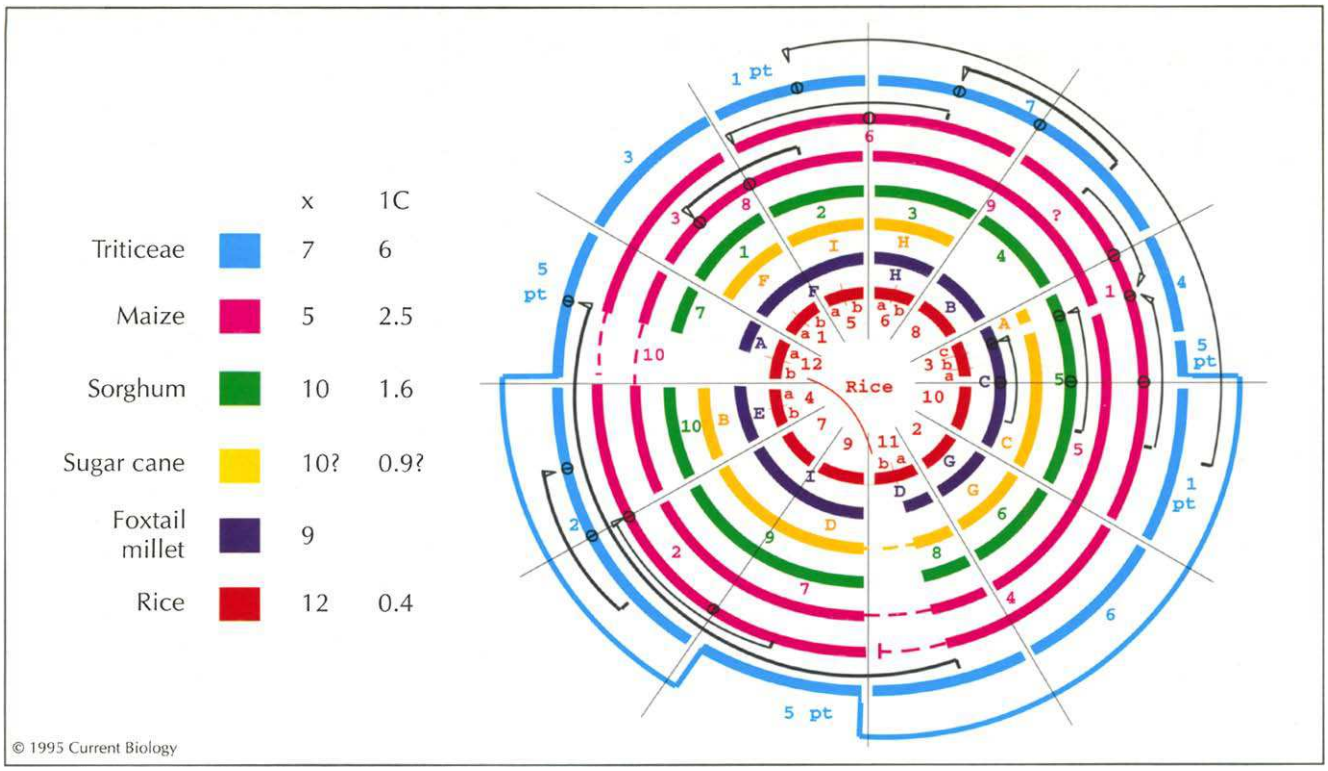

Fig. 2. Alignment of the genomes of six major grass crop species with 19 rice linkage segments, whose order reflects the circularized ancestral grass genome. The data in Fig. 1 have been redrawn as a series of rice linkage segments (defined by radiating lines) formed into chromosomes (broken colour-coded and numbered lines). The thin dashed lines correspond to the duplicated segments. Inversions of sets of sequences within a linkage segment (such as the inversion of segments 3a and 3b in maize chromosome 5) are not shown. Linkage segments forming parts (pt) of Triticeae chromosome 5 are shown as a series of segments connected by coloured lines. The alignment is based on the genetic map of the D genome of wheat. The red line indicates the duplicated segments shown as blocks 11b and 12b. Chromosomes formed by the insertion of one segment into another are shown by black lines with arrows indicating the direction and point of insertion. The points of chromosome breakage involved with insertion events are indicated by black bisected circles. The 'haploid' chromosome number of each species is shown in the column marked 'x'. The haploid DNA content of each species, shown in the column of 1C values, is per 10^9 bases.

wheat RFLP probes to cross-hybridize with homologous sequences in other cereals allowed us to extend these comparisons to the genomes of other major Triticeae species — rye and barley. Although the rye genome was found to be rearranged relative to that of wheat by a few major translocations [2], most probes revealed that the same loci in wheat, barley and rye were arranged in exactly the same order along large stretches of their chromosomes.

In 1992 we were able to start asking the same questions about the relationships between the genomes of wheat and maize, and between those of wheat and rice. We found that the gene order was conserved at the genetic level, despite these species having been isolated for up to 60 million years — a period equivalent to that during which eutherian mammals have evolved [3].

Rice linkage segment analysis

The rice genome is one of the smallest among the grasses and, following recent intense research activity, has become pivotal to grass genetic studies. When the relationship between the genetic maps of the three major cereals was examined, it became clear that the conservation of gene order was restricted to a number of rice linkage segments. It was then possible to describe the genes on the maps of individual maize and wheat chromosomes in terms of these segments, rather like building with blocks of 'Lego'. These blocks could be assembled into a single stack for a Triticeae genome and two other stacks to form a maize genome. Furthermore, all of the 19 blocks could be assembled in a specific order, forming a single stack which may represent the ancestral cereal genome. Then, for each species, the stack can be cleaved at different points to generate single blocks, or groups of blocks, which describe the gene framework of each chromosome in its genome [3].

This analysis is based on maps derived from recombination data, expressed in centimorgans. Whether the gene orders are conserved at the megabase level is, at present, still an assumption. Similarly, it is clear that this conservation applies only to genes; the nature and amount of intergenic DNA varies greatly between species, depending on the extent of accumulation of sequences, mostly repeated, following isolation by speciation.

An extended grass genome framework

We have now been able to extend the analysis to include foxtail millet (*Setaria italica*), sorghum (*Sorghum bicolor*) and sugar cane (*Saccharum* spp.). The results, based on the cross-mapping of common DNA probes, are shown in

Figures 1 and 2. It is apparent that only 19 linkage blocks, comprised of large chromosome segments or whole rice chromosomes (ignoring the duplicated segments [4] shown as blocks 11b and 12b in Figs 1 and 2), are needed to reconstruct all the cereal chromosomes compared so far. Moreover, with the rice linkage segments arranged as shown, they can be fitted together in the same order to compose most of 56 chromosomes in the diagrams. The recognition that a segment of maize chromosome 3 is duplicated on the long arm of chromosome 10 (Fig. 1c), effectively circularizes the consensus grass genome. This single 'chromosome' provides the basis of a consensus map that may reflect the genome of the ancestral grass from which these cereals evolved.

The conservation of genes and gene orders is irrespective of 'haploid' chromosome numbers (x) of the six species. These vary from 5 (maize), 7 (wheat, barley and rye), 9 (foxtail millet), and ~ 10 (sugar cane, sorghum) to 12 (rice). It is also irrespective of the haploid DNA contents (1C value) which vary 40-fold — rice being the smallest (400 megabases) and the Triticeae cereals the largest. Moreover, it is clear from information presented at the recent Plant Genome III meeting in San Diego (15–19 January 1995) that hexaploid oats (*Avena sativa*; x = 7) [5] and rye-grass (*Lolium perenne*; x = 7) [6] share similar levels of conservation, and could soon be incorporated into the rice-linkage analysis.

Cereal evolution

The analysis confirms that maize is a complete tetraploid, although the two genomes are rearranged relative to one another. Interestingly, most of the larger chromosomes (1, 2, 3, 4 and 6) form one genome, and a shorter set (5, 7, 8, 9 and 10) form the other. Doubts as to the ploidy of sorghum can also be resolved, as it is clear that the ten chromosomes form a complete cereal genome. Furthermore, the break in the ancestral 'chromosome' between blocks 3c and 3b in the millet, sorghum, maize and sugar cane genomes, together with the insertion of block 10 into the same break in millet, sorghum and one of the maize genomes, clearly places these species in a group (see Fig. 1). The Triticeae species and Oryzeae (rice) species form two further groups.

Applications

For those interested in the more detailed mapping of any region of any grass genome, Figure 2 shows, at a glance, the chromosomal regions of other species that will serve as sources of potential DNA or protein markers. For those interested in traits, and the genes that control them, Figure 2 can also be used to reveal similar traits mapping to corresponding linkage segments which are likely to be controlled by homologous genes. For example, a radius in Figure 2 from block 3b will pass through *Rht1* and *Rht2*, the wheat gibberellin-insensitive semi-dwarfing genes, and *d*9 and *d*8, the maize gibberellin-insensitive dwarfing genes. Similarly, a radius from block 4a passes through the *liguless* genes on barley block 2, maize block 2 and rice block 4.

The way is now open to capitalize on the wealth of information available for the many grass species, by exploiting the extensive resources being devoted to the analysis of the small genome of rice. Genes located in any grass can be isolated by map-based cloning of the homologue in rice, and then by homology in the target species. The reconstruction of the rice genome in the form of ordered yeast artificial chromosomes will reveal the general gene order and content of all cereal genomes. This may provide clues as to how and why particular regions of some cereal genomes have undergone extensive expansion by the non-random accumulation of repetitive sequences following speciation [7].

The concept thus describes a framework for the eventual collation of most of the information derived from the genetical studies of individual cereal crops and their wild relatives since the rediscovery of Mendel's laws.

References

1. Chao S, Sharp PJ, Worland AJ, Warham EJ, Koebner RMD, Gale MD: **RFLP-based genetic maps of the wheat homoeologous group 7 chromosomes.** *Theor Appl Gen* 1989, **78**:495–504.
2. Devos KM, Atkinson MD, Chinoy CN, Francis HA, Harcourt RL, Koebner RMD, Liu CJ, Masojé P, Xie DX, Gale MD: **Chromosomal rearrangements in the rye genome relative to that of wheat.** *Theor Appl Gen* 1993, **85**:673–680.
3. Moore G, Foote T, Helentjaris T, Devos K, Kurata N, Gale M: **Was there a single ancestral cereal chromosome?** *Trends Genet* 1995, **11**:81–82.
4. Shimano T, Inove T, Antonio B, Kajiya H, Shomura A, Yang Lin S, Kuboki Y, Nagamura N, Yano M, Sasaki T: **Extensive conservation in linkage alignment of RFLP markers between rice chromosome 11 and 12.** *Plant Genome III* 1995, 212.
5. Van Deynze AE, Dubcovsky J, Gill KB, Lagudah ES, McCouch SR, Appels R: **Molecular genetic maps for group 1 chromosomes of Triticeae species and their relation to chromosomes in rice and oat.** *Plant Genome III* 1995, 45.
6. Forster J, Hayward M, Richard C, Evans C, Devos K, Gale M: **Comparative genetic mapping between forage grasses and the Triticeae.** *Plant Genome III* 1995, 61.
7. Abbo S, Dunford RP, Foote TN, Reader SM, Flavell RB, Moore G: **Organization of retroelement and stem-loop repeat families in the genomes and nuclei of cereals:** *Chromosome Res* 1995, **3**:5–15.
8. Whitkus R, Doebley J, Lee M: **Comparative genome mapping of sorghum and maize:** *Genetics* 1992, **132**:1119–1130.
9. Melake Berhan A, Hulbert SH, Butler LG, Bennetzen JL: **Structure and evolution of the genomes of *Sorghum bicolour* and *Zea mays*:** *Theor Appl Gen* 1993, **86**:598–604.
10. Periera MG, Lee M, Bramel-Cox P, Woodman N, Doebley J, Whitkus R: **Construction of an RFLP map in sorghum and comparative mapping in maize.** *Genome* 1993, **37**:236–243.
11. Chittenden LM, Schertz KF, Lin Y-R, Wing RA, Paterson AH: **A detailed RFLP map of *Sorghum bicolour* x *S. propinquum* suggests ancestral duplication of sorghum chromosomes or chromosomal segments.** *Theor Appl Gen* 1994, **87**:925–933.
12. Grivet L, D'Hont A, Dufour P, Hamon P, Rogues D, Glaszmann JC: **Comparative genome mapping of sugarcane with other species with the *Andropogoneae*.** *Heredity* 1994, **73**:500–508.
13. Grivet L: **Marquage moleculaire chez la canne a sucre (*Saccharum* spp.); decomposition d'une structure genetique complexe et application a l'amelioration varietale.** *PhD Thesis, University de Paris XI Orsay* 1995.

Graham Moore, Katrien M. Devos, Zhimin Wang and Mike D. Gale, John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, UK.