



ELSEVIER

## Updating the 'Crop Circle'

Katrien M Devos

Comparative analyses unravel the relationships between genomes of related species. The most comprehensive comparative dataset obtained to date is from the grass family, which contains all of the major cereals. Early studies aimed to identify chromosomal regions that have remained conserved over long evolutionary time periods, but in recent years, researchers have focused more on the extent of colinearity at the DNA-sequence level. The latter studies have uncovered many small rearrangements that disturb colinearity in orthologous chromosome regions. In part, genomes derive their plasticity from genome- and gene-amplification processes. Duplicated gene copies are more likely to escape selective constraints and thus move to other regions of the genome, where they might acquire new functions or become deleted. These rearrangements will affect map applications. The most popular applications, especially since the complete rice genomic sequence has been available, are the use of comparative data in the generation of new markers to tag traits in other species and to identify candidate genes for these traits. The isolation of genes underlying orthologous traits is the first step in conducting comparative functional studies.

### Addresses

Department of Crop and Soil Sciences, and Department of Plant Biology, University of Georgia, Athens, Georgia 30602, USA

Corresponding author: Devos, Katrien M (kdevos@uga.edu)

**Current Opinion in Plant Biology** 2005, 8:155–162

This review comes from a themed issue on  
Genome studies and molecular genetics  
Edited by Susan Wessler and James C Carrington

Available online 29th January 2005

1369-5266/\$ – see front matter  
© 2005 Elsevier Ltd. All rights reserved.

DOI 10.1016/j.pbi.2005.01.005

### Introduction

The first comparative genetic maps were produced in the late 1980s [1,2]. As the value of comparative genomics became clear and relationships between an increasing number of species were unraveled, it became necessary to devise new visualization methods that could display multiple comparisons in an adequate way. In 1995, Moore and colleagues [3] published the first 'Crop Circle' diagram. The display has since been updated to more precisely delineate syntenic relationships between species of the grass family [4,5] and has also been adopted for legumes [6].

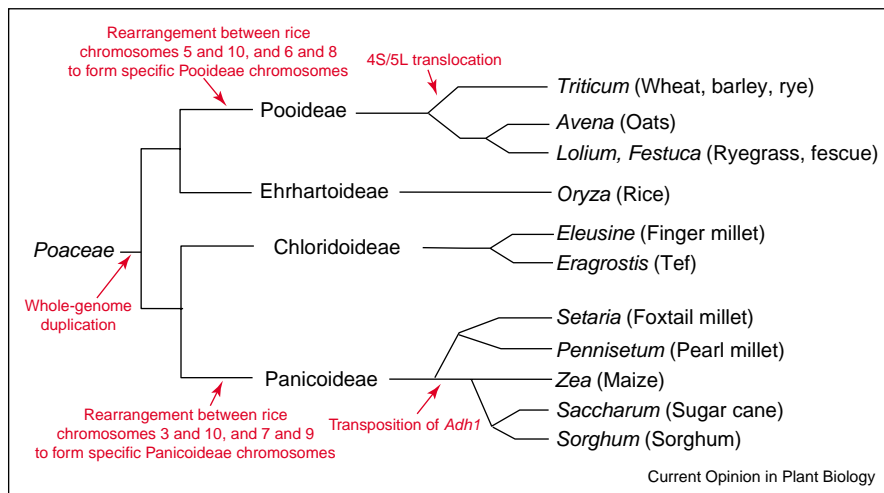
In this review, we discuss recent developments in the area of grass comparative genomics. These include the addition of new species to the comparative portfolio, and a refinement of the comparative maps through exploitation of mapped expressed sequenced tags (ESTs) and the near-completed rice genomic sequence. In particular, the sequence produced by the International Rice Genome Sequencing Project, which is ordered and anchored to the genetic map, has proven to be an important tool in the quest for understanding of the structural relationships between grass genomes. In recent years, important insights have also been gained into the extent of gene-order conservation at the DNA-sequence level. In contrast to map-based efforts, which provide an overview of chromosomal rearrangements that differentiate related species, sequence-based comparisons assess whether gene orders have remained conserved within orthologous chromosome segments. The level of gene conservation is an important criterion in determining the extent to which comparative knowledge can be applied across species. For example, high levels of colinearity are a prerequisite for the identification of candidate genes for traits of interest using cross-species information, but might be less crucial when simply using sequence data generated in one species as a source of new markers for related species.

Interpreting comparative data is not always straightforward. Partial or complete genome duplications followed by a gradual return to the diploid state might confound the identification of orthologous genes. This problem has been highlighted in *Arabidopsis*, maize and rice, but is likely to be more widespread. Indeed, on the basis of the extent of sequence divergence between duplicated rice genes, the genome-wide duplication that characterizes the rice genome has been dated to 70 million years ago, before the divergence of the major grass lineages. This review addresses some of the impacts of duplications on comparative studies.

### Enhancement of the Crop Circles with new species

The 1998 version of the Crop Circles represented genome relationships among eight species belonging to three grass subfamilies: rice (Ehrhartoideae); foxtail millet, sugar cane, sorghum, pearl millet and maize (Panicoideae); and the Triticeae cereals and oats (Pooideae) [5]. Since then, data have emerged on wild rice (*Zizania palustris*, Ehrhartoideae) [7], rye grass (*Lolium perenne*, Pooideae) [8] and meadow fescue (*Festuca pratensis*, Pooideae) [9]. Comparative maps of finger millet (*Eleusine coracana*) (Srinivasachary, MD Dida, KM Devos, unpublished)

Figure 1



Grass phylogeny showing the four subfamilies that include most of the cereals. The timing of some of the rearrangements that differentiate lineages has been indicated.

and tef (*Eragrostis tef*) (ME Sorrells, personal communication), two Chloridoideae species, are near completion. The phylogeny of these four subfamilies (i.e. Ehrhartoideae, Panicoideae, Pooideae and Chloridoideae) is presented in Figure 1.

The main observations from these additions are, first, that wild and cultivated rice have a highly conserved genome composition, with the higher chromosome number of wild rice ( $2n = 2x = 30$ ) compared to cultivated rice ( $2n = 2x = 24$ ) being due to complete or partial duplications of chromosomes 1, 4 and 9. Second, *Lolium* ( $2n = 2x = 14$ ) and *Festuca* ( $2n = 2x = 14$ ) have a genome structure that is largely similar to that of the Triticeae, with the exception of chromosome 4, which is largely colinear to rice chromosome 3 (Figure 2a). Markers that are located distally on the long arm of rice chromosome 3 are found on the long arm of chromosome 5 (5L) in the Triticeae, indicating that this lineage underwent a translocation between chromosome arms 4S and 5L after its divergence from the Pooideae. Preliminary analysis of the finger millet maps suggests that the *Eleusine* and rice genomes are highly colinear, and that rearrangements are specific either to finger millet or to the Chloridoideae tribe.

### Improving the resolution of comparative maps

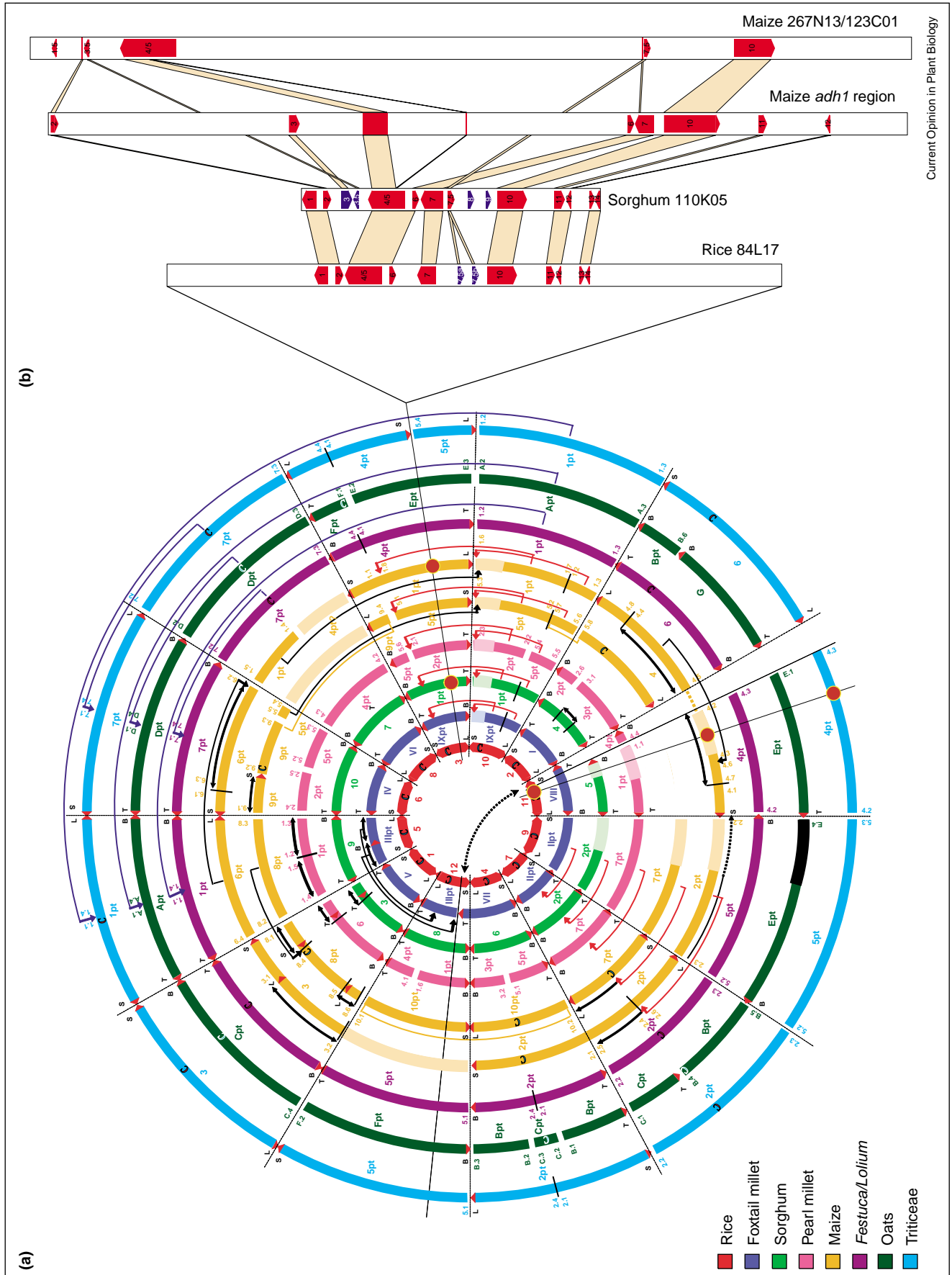
The 'traditional' way of generating comparative maps is to use common sets of probes with good cross-hybridization ability, mostly cDNAs, across different species. The recent surge in the large-scale sequencing of ESTs in several grass species has provided a more efficient way to conduct comparative mapping. *In silico* analysis of ESTs against the rice genomic sequence provides the anchor points in rice, and mapping needs to be carried out only in the species from which the ESTs

are derived. This can be done relatively easily either by hybridization or PCR. The EST approach has been used to enhance the resolution of the wheat–rice [10,11\*] and maize–rice comparative maps [12]. Figure 2a has been adapted from that published by Gale and Devos [5] to reflect new information resulting from these studies.

For genomic regions covered by bacterial artificial chromosome (BAC) contigs, a further level of refinement can be obtained by aligning low-pass BAC sequence data with the rice genomic sequence. Klein and colleagues [13] demonstrated that, in sorghum, a sequence scan of 16–32 subclones per BAC (approximately 0.5X BAC coverage) was sufficient to obtain one or more genes that could be aligned with rice. Their study revealed the presence of a previously undetected inversion in the sorghum map relative to the rice map (Figure 2a). Of course, a greater BAC coverage will be needed to map the larger plant genomes, but even in the large wheat genome, 2X coverage is likely to provide sufficient sequence information to identify any genes or gene fragments that are present on the BAC.

Another level of accuracy could be provided by statisticians. They aim to remove human biases and interpretations from the process of identifying orthologous chromosome regions. The program LineUp was designed with plants in mind as it tests the statistical significance of single colinear runs while simultaneously considering rearrangements and physical distances between markers [14]. When tested for its ability to detect duplicated segments in the genetic maps of maize, an ancient tetraploid, the results varied with the mapping set. It should be noted that all but one of the colinear runs that were consistently detected had also been identified by visual

Figure 2



interpretation of the data [5]. The only exception is a duplication between maize chromosomes 1 and 3 which, on the basis of the composition of the two maize genomes (Figure 2a), would have been considered an intragenomic duplication and hence been discounted in most comparative studies. Maybe humans aren't doing such a bad job after all.

### Conservation of colinearity at the DNA-sequence level

Comparative analyses at the map level identify genome segments or 'linkage blocks' consisting of tens of megabases that are largely colinear across species. These low- and medium-resolution mapping studies do not provide information on the precise organization of genes within these blocks. When researchers first began to conduct chromosome walks in rice to isolate the ortholog of a gene underlying a trait of interest in a large-genome species, it was with mixed success [15,16]. To gain a better insight into the level of gene conservation at the microlevel, several laboratories have since undertaken sequencing of BACs that were selected to contain a particular gene across a range of crop species [17–20,21<sup>••</sup>,22–28]. The pervading message is that, by and large, colinearity is maintained at the sequence level, albeit to different levels depending on the region and the species. Disruption of colinearity is manifested by small inversions, tandem duplications, single or multiple gene insertions and/or deletions, and gene translocations (Figure 2b). By comparing orthologous regions across multiple species, it is possible to infer the relative timing of the different rearrangements, and also to assess the relative stability of the genomes analyzed. The rice genome appears to be more similar in structure to the ancestral genome organization than are the genomes of sorghum or maize [20,21<sup>••</sup>]. This is also true at the map level, at least for some of the rice chromosomes [29]. There are currently too few data, however, to allow us to conclude whether the degree of rearrangement at the map and DNA-sequence levels are correlated.

Regions that contain tandem gene clusters are particularly unstable. Testimony to this is the differential presence of clusters of disease resistance genes [19,28], small nucleolar RNA (snoRNA) genes [20] and storage protein genes [28,30] in cereals. Following amplification, these genes

might have been retained in specific lineages in response to environmental selection pressures. The mechanisms that drive rearrangements are generally unknown, but some information can be gleaned from flanking sequences. The presence of inverted repeats flanking a gene that has undergone an inversion suggests that unequal homologous recombination was the driving mechanism [20]. Short direct repeats flanking deletions and duplications might be indicative of illegitimate recombination [31,32]. Gene movement might be associated with gene amplification [22,30,33], although the evidence for this is not unequivocal. Song and colleagues [30] postulated that duplicated gene copies could be inserted elsewhere in the genome by illegitimate recombination, in a process similar to the introgression of free DNA molecules delivered by transformation procedures into the plant genome. In an independent study, Ma and Bennetzen (personal communication), however, found no evidence that transposed gene copies are significantly associated with duplication events.

Once a gene has been amplified, the fate of the paralogs can be multifold. Rearrangements between genes within a cluster can lead to new gene variants, as has been demonstrated for disease resistance genes [34,35]. The insertion of a gene into a new position might lead to a change in its transcriptional regulation that is caused by the changed chromosomal environment [36]. The accumulation of degenerative mutations in regulatory elements can cause partitioning of ancestral functions [37]. In the majority of cases, however, one of the paralogous gene copies will be deleted [21<sup>••</sup>,22,33,38]. Sequence analyses of duplicated regions and the effect of gene duplication, movement or deletion on the interpretation of comparative data are discussed in more detail in the next section.

### Effects of segmental duplications and polyploidization

Polyploidy is widespread in eukaryotes and is probably one of the main mechanisms for evolutionary divergence. The small-genome model species *Arabidopsis thaliana* was considered to be a diploid until whole-genome sequence information revealed the occurrence of three ancestral rounds of duplication [39,40]. A similar analysis of the rice genomic sequence has recently demonstrated that at least

**(Figure 2 Legend)** Crop Circle diagram showing the currently known relationships between the genomes of eight species belonging to three different subfamilies. **(a)** Each circle represents a different genome. 'Red triangles' indicate telomeres. Where known, centromeres are indicated as 'C'. Arrows indicate rearrangements relative to rice. Red and blue arrows are rearrangements common to the Panicoideae and Pooideae subfamilies, respectively. In some cases, arrows have been supplemented or replaced by numbered genome fragments. 'Red dots' represent the location, where known, of the *Adh1* and *Adh2* genes. In rice and barley, both genes are linked on the orthologous chromosomes rice 11 and barley 4H, respectively. In maize, only the *Adh2* gene is found on the orthologous chromosome 4, whereas the *Adh1* gene has transposed to maize chromosome 1. An *Adh* gene is found in an orthologous position on sorghum chromosome 1, indicating that the transposition occurred before the divergence of the sorghum and maize lineages. **(b)** Microcolinearity of *Adh*-orthologous regions of rice, sorghum and the two maize homoeologs (redrawn from [21<sup>••</sup>]). Genes are indicated by red and blue arrows. Red blocks represent gene fragments. Gene 7.5 has undergone a duplication in rice, and sorghum genes 3, 3.5, 8 and 9 (in blue) are not present at this location in rice. The adjacent genes 3 and 3.5 moved to this location in a common ancestor of sorghum and maize, whereas genes 8 and 9 independently moved to this location in a sorghum ancestor after its divergence from the maize lineage. Common genes are connected by pink bars.

53–62% of the sequence, and possibly the entire rice genome, is duplicated [41,42\*\*]. Phylogenetic dating of duplicated rice genes suggests that the duplication predates the divergence of the major grass lineages [42\*\*,43]. This could explain the high level of duplicated sequences identified in the diploid sorghum genome and implies that all cereal crops, even those currently considered to be diploids, are of polyploid origin.

The most likely fate of genes following duplication is the removal of one of the gene copies. In rice, only 21.4% of the genes within the duplicated segments retain their paralog [42\*\*]. A more precise picture of the diploidization process following genome duplication has been provided by analyses comparing duplicated segments of the maize genome with orthologous regions in related grass species [21\*\*,33,38]. These studies reported that genes had been deleted in an apparently non-preferential manner from one or the other of the homoeologous maize genomes. Depending on the region, retention of both homoeologs varied from less than 10% to 50% of the genes. It should be noted that gene functionality had not been assessed in all cases, and therefore structural retention does not imply that both homoeologs are functional. Some of the genes had been removed entirely, whereas remnants of others remained. All this has happened since the tetraploidization of maize, which is estimated to have taken place between 11.9 and 4.8 million years ago [44\*]. In synthetic polyploids, genomic changes can be observed as early as the F<sub>1</sub> generation [45,46]. Diploidization might thus be initiated as soon as different genomes are combined in one nucleus. Fu and Dooner [47] observed that gene content also varies across maize inbreds. As these authors sequenced only one of the two maize homoeologous regions, however, it remains possible that the missing genes were present on the other homoeolog. If confirmed, this would suggest that diploidization is an ongoing process that can take tens of million years to complete.

It is clear that duplication events, followed by gene movement and deletion of one of the copies, will impact colinearity. Differential gene movement after speciation will, *a priori*, cause disruption of colinearity. Another possible scenario is that duplicated genes are translocated in a common ancestor but different gene copies are subsequently deleted in diverged lineages. Unless phylogenetic studies are conducted, it is likely that paralogs will be misidentified as orthologs that occupy non-colinear positions. An analysis of the distribution of non-colinear loci in a rice–sorghum comparative map has indeed shown that a significant number of these loci were located in putative homoeologous sorghum regions [42\*\*]. This suggests that, at one time, an orthologous copy might have been present in a colinear position on the other homoeolog, and affirms the hypothesis of differential gene deletion across lineages.

### From structure to function: trait mapping across species

Orthologous genes that have conserved functions are expected to produce similar phenotypes across species. For example, the *GAI*, *Rht-1* and *D8* orthologous dwarfing genes reduce plant height in *Arabidopsis*, wheat and maize, respectively [48]. In recent years, comparative studies have been conducted to assess to what extent multigenic traits are controlled by the same genes. I do not aim to give an exhaustive list of all publications that show conserved colinearity of traits across species. Instead, a few papers with interesting and somewhat surprising conclusions have been selected for discussion.

Domestication is a relatively recent event that occurred independently and sometimes several times in the different cereals. Traits that have been selected during domestication include non-shattering of the seed, larger seed size, reduced tillering and annual growth habit. The presence of quantitative trait loci (QTL) for these traits in orthologous positions across species suggests that convergent evolution of independent mutations in orthologous genes might have played an important role in the domestication of the grasses [49,50]. It also indicates that relatively few genes underlie the traits selected for by early farmers or, alternatively, that only a few genes can be modified without detrimental effect on the agricultural performance of the crop.

Doust and colleagues [51\*\*] conducted a QTL study of the genetic control of branching in foxtail millet and evaluated the effect of *teosinte branched* (*tb1*), the gene that controls the transition in plant architecture from the wild *Teosinte* form to that of cultivated maize. Although the *tb1* region played a role in branching, it was neither the only nor the most important locus controlling vegetative branching in foxtail millet. Thus, even when orthologous QTL underlie the same trait across species, the contribution of each of the orthologous genes to the phenotype might vary.

Disease resistance genes are known for their ability to reorganize rapidly, and are often found in non-colinear positions in different species [52–54]. A recent comparative study of QTL that confer resistance to the blast fungus, *Pyricularia grisea*, in rice and barley, however, identified loci that were conserved both in their genome location and their pathogen-race specificity [55]. On first thought, conservation of race specificity over a time span of 60 million years seems unlikely, particularly considering the continuing contest between pathogen and plant to, respectively, overcome and develop new resistances. The fact that barley is normally grown in environmental conditions that are unsuitable for infection by *Pyricularia grisea*, and thus is not under selective constraint to modify the ancestral blast resistance genes, might have been a contributing factor. However, this does not explain why

the resistance specificity has been conserved in rice, a major host for the blast fungus. Apart from fortuity, a possible explanation might lie in that, in the rice population analyzed by Chen and colleagues [55], the putative orthologous rice QTL were of only minor effect. This might place them under reduced selection compared to the major resistance gene(s) that are present in the population. If this holds true, the corollary would be that one would be unlikely to find an orthologous QTL that has a major effect in conferring resistance to the same pathogen isolate.

### Current views on map applications

The increase in comparative knowledge and the development of genomic tools for most of the cereal species has resulted in a paradigm shift in how comparative data can best be exploited. Early comparative geneticists focused mainly on using the small rice genome as an intermediate for conducting chromosome walks in large genome species, such as maize and wheat. The observations that small rearrangements, including gene deletions and transpositions, disturb micro-colinearity at relatively high frequencies helped to stimulate resource development in all major crops. Recent years have seen an explosion in EST data and the availability of large-insert libraries. Furthermore, chromosome walks in rice are no longer necessary, as the entire sequence of this cereal is available. Most comparative applications now use the rice genomic sequence as a source of markers to tag genes of interest in other grass crops [24,56,57]. Starting from the conserved linkage blocks that have been identified in the comparative maps, the rice genomic sequence is used as template to gradually 'zoom in' on the target gene. Rice sequences can be used directly as hybridization probes or, more efficiently, to identify either ESTs from the species of interest or ESTs from other grasses that can be aligned to identify conserved regions for primer design. Alternatively, the genomic information from rice and other species can be scanned for putative candidate genes that are located within the region of interest [51<sup>••</sup>,56]. A perhaps more-controversial application is the use of the rice genomic sequence as a template for physical maps in other species. This approach will work well in regions of conserved colinearity [13] but might lead to the misassignment of contigs for regions that have undergone rearrangements, in particular in large genome species in which large stretches of repetitive DNA could decrease the resolving power of BAC fingerprinting and contigging.

### Conclusions

Recent advances in comparative mapping have seen a refinement of previously established chromosomal relationships. A few species have been or are being added to comparative maps; these include two species belonging to the Chloridoideae, a subfamily that was previously unrepresented in the Crop Circles. In an evolutionary context,

it would be interesting to include monocots that belong to families other than, but closely related to, the grasses in comparative analyses. We know that colinearity between monocots and *Arabidopsis* is highly fragmentary [58–60], but we have no information on species that diverged between 60 million years ago (the radiation of the grasses) and 150–200 million years ago (the monocot/eudicot split).

The main focus of grass comparative studies in the past few years has been on comparisons at the DNA-sequence level and on the effect of whole-genome or segmental duplications. Colinearity in orthologous regions is frequently disturbed by small rearrangements that appear to occur at different rates in different lineages. Rearrangements occur at a greater rate in ancient polyploids, such as maize, because diploidization, the most likely fate of duplicated segments, occurs in an apparently non-selective manner across homoeologous regions in these species. Although the rice genome has also been duplicated, this duplication is estimated to have occurred before the divergence of the grass species and should thus be manifest in all cereal crops. Furthermore, the diploidization process might have been initiated immediately following the whole-genome duplication and before the radiation of the grasses. Genes that were removed from a common ancestor do not affect grass genome colinearity and could explain why the rice genome has been identified as particularly stable.

This review has focused only on the conserved portion of the genome, the genes. However, the repetitive fraction of the genome provides useful information on the way genomes evolve [31,32]. Future comparative analyses are therefore more likely to give attention to those components of the genome that differ between species.

In addition, there will be a shift from analyses at the structural level to analyses at the functional level. Genes of common decent might, in some cases, have acquired different functions or specificities in different species. A first indication of this has been revealed in the study by Doust and colleagues [51<sup>••</sup>] who reported that *tb1* is not the major player in controlling tillering in foxtail millet. Mutant collections are becoming available in the major grass species, and will be a valuable resource for comparative functional analyses. These studies may shed some light on what determines a species' unique identity.

### References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
  - of outstanding interest
1. Bonierbale MW, Plaisted RL, Tanksley SD: **RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato.** *Genetics* 1988, **120**:1095-1103.

2. Chao S, Sharp PJ, Worland AJ, Warham EJ, Koebner RMD, Gale MD: **RFLP-based genetic maps of wheat homoeologous group 7 chromosomes.** *Theor Appl Genet* 1989, **78**:495-504.
  3. Moore G, Devos KM, Wang Z, Gale MD: **Grasses, line up and form a circle.** *Curr Biol* 1995, **5**:737-739.
  4. Devos KM, Gale MD: **Comparative genetics in the grasses.** *Plant Mol Biol* 1997, **35**:3-15.
  5. Gale MD, Devos KM: **Plant comparative genetics after 10 years.** *Science* 1998, **282**:656-659.
  6. Choi H-K, Mun J-H, Kim D-J, Zhu H, Baek J-M, Mudge J, Roe B, Ellis N, Doyle J, Kiss GB *et al.*: **Estimating genome conservation between crop and model legume species.** *Proc Natl Acad Sci USA* 2004, **101**:15289-15294.
  7. Kennard WC, Phillips RL, Porter RA, Grombacher AW: **A comparative map of wild rice (*Zizania palustris* L.  $2n=2x=30$ ).** *Theor Appl Genet* 2000, **101**:677-684.
  8. Jones ES, Mahoney NL, Hayward MD, Armstead IP, Jones JG, Humphreys MO, King IP, Kishida T, Yamada T, Balfourier F *et al.*: **An enhanced molecular marker based genetic map of perennial ryegrass (*Lolium perenne*) reveals comparative relationships with other *Poaceae* genomes.** *Genome* 2002, **45**:282-295.
  9. Alm V, Fang C, Busso CS, Devos KM, Vollan K, Grieg Z, Rognli OA: **A linkage map of meadow fescue (*Festuca pratensis* Huds.) and comparative mapping with other *Poaceae* species.** *Theor Appl Genet* 2003, **108**:25-40.
  10. Sorrells ME, La Rota M, Bermudez-Kandianis CE, Greene RA, Kantety R, Munkvold JD, Miftahudin MA, Ma XF, Gustafson PJ, Qi LL *et al.*: **Comparative DNA sequence analysis of wheat and rice genomes.** *Genome Res* 2003, **13**:1818-1827.
  11. La Rota M, Sorrells ME: **Comparative DNA sequence analysis of mapped wheat ESTs reveals the complexity of genome relationships between rice and wheat.** *Funct Integr Genomics* 2004, **4**:34-46.
- The wheat-rice comparative map represents the most comprehensive whole-genome comparison carried out to date and involves some 3000 anchor points.
12. Salse J, Piégou B, Cooke R, Delseny M: **New *in silico* insight into the synteny between rice (*Oryza sativa* L.) and maize (*Zea mays* L.) highlights reshuffling and identifies new duplications in the rice genome.** *Plant J* 2004, **38**:396-409.
  13. Klein PE, Klein RR, Vrebalov J, Mullet JE: **Sequence-based alignments of sorghum chromosome 3 and rice chromosome 1 reveals extensive conservation of gene order and one major chromosomal rearrangement.** *Plant J* 2003, **34**:605-621.
  14. Hampson S, McLysaght A, Gaut BS, Baldi P: **LineUp: statistical detection of chromosomal homology with application to plant comparative genomics.** *Genome Res* 2003, **13**:999-1010.
  15. Kilian A, Chen J, Han F, Steffenson B, Kleinhofs A: **Towards map-based cloning of the barley stem rust resistance genes *Rpg1* and *Rpg4* using rice as an intergenomic cloning vehicle.** *Plant Mol Biol* 1997, **35**:187-195.
  16. Foote T, Roberts M, Kurata N, Sasaki T, Moore G: **Detailed comparative mapping of cereal chromosome regions corresponding to the *Ph1* locus in wheat.** *Genetics* 1997, **147**:801-807.
  17. Tikhonov AP, SanMiguel PJ, Nakajima Y, Gorenstein NM, Bennetzen JL, Avramova Z: **Colinearity and its exceptions in orthologous *adh* regions of maize and sorghum.** *Proc Natl Acad Sci USA* 1999, **96**:7409-7414.
  18. Dubcovsky J, Ramakrishna W, SanMiguel PJ, Busso CS, Yan LL, Shiloff BA, Bennetzen JL: **Comparative sequence analysis of colinear barley and rice bacterial artificial chromosomes.** *Plant Physiol* 2001, **125**:1342-1353.
  19. Ramakrishna W, Emberton J, SanMiguel P, Ogden M, Llaca V, Messing J, Bennetzen JL: **Comparative sequence analysis of the sorghum *Rph* region and the maize *Rp1* resistance gene complex.** *Plant Physiol* 2002, **130**:1728-1738.
  20. Ramakrishna W, Dubcovsky J, Park Y-J, Busso C, Emberton J, SanMiguel PJ, Bennetzen JL: **Different types and rates of genome evolution detected by comparative sequence analysis of orthologous segments from four cereal genomes.** *Genetics* 2002, **162**:1389-1400.
  21. Ilic K, SanMiguel PJ, Bennetzen JL: **A complex history of rearrangement in an orthologous region of the maize, sorghum, and rice genomes.** *Proc Natl Acad Sci USA* 2003, **100**:12265-12270.
- This paper provides the first example of a comparative sequence analysis across homoeologous regions of the ancient tetraploid maize. It demonstrates that maize is well on its way to diploidization through the non-selective deletion of duplicated gene copies in one of the homoeologous maize genomes.
22. Swigonova Z, Bennetzen JL, Messing J: **Structure and evolution of the *r/b* chromosomal regions in rice, maize and sorghum.** *Genetics* 2005, in press.
  23. Wicker T, Yahiaoui N, Guyot R, Schlagenhauf E, Liu Z-D, Dubcovsky J, Keller B: **Rapid genome divergence at orthologous low molecular weight glutenin loci of the A and A<sup>m</sup> genomes of wheat.** *Plant Cell* 2003, **15**:1186-1197.
  24. Brunner S, Keller B, Feuillet C: **A large rearrangement involving genes and low-copy DNA interrupts the microcollinearity between rice and barley at the *Rph7* locus.** *Genetics* 2003, **164**:673-683.
  25. Gu YQ, Coleman-Derr D, Kong X-Y, Anderson OD: **Rapid genome evolution revealed by comparative sequence analysis of orthologous regions from four triticeae genomes.** *Plant Physiol* 2004, **135**:459-470.
  26. Kong X-Y, Gu YQ, You FM, Dubcovsky J, Anderson OD: **Dynamics of the evolution of orthologous and paralogous portions of a complex locus region in two genomes of allopolyploid wheat.** *Plant Mol Biol* 2004, **54**:55-69.
  27. Gu YQ, Anderson OD, Londeoré CF, Kong X-Y, Chibbar RN, Lazo GR: **Structural organization of the barley D-hordein locus in comparison with its orthologous regions of wheat genomes.** *Genome* 2003, **46**:1084-1097.
  28. Guyot R, Yahiaoui N, Feuillet C, Keller B: ***In silico* comparative analysis reveals a mosaic conservation of genes within a novel colinear region in wheat chromosomes 1AS and rice chromosome 5S.** *Funct Integr Genomics* 2004, **4**:47-58.
  29. Gale MD, Devos KM: **Comparative genetics in the grasses.** *Proc Natl Acad Sci USA* 1998, **95**:1971-1974.
  30. Song R, Llaca V, Messing J: **Mosaic organization of orthologous sequences in grass genomes.** *Genome Res* 2002, **12**:1549-1555.
  31. Devos KM, Brown JKM, Bennetzen JL: **Genome size reduction through illegitimate recombination counteracts genome expansion in *Arabidopsis*.** *Genome Res* 2002, **12**:1075-1079.
  32. Ma J, Devos KM, Bennetzen JL: **Analyses of LTR-retrotransposon structures reveal recent and rapid genomic DNA loss in rice.** *Genome Res* 2004, **14**:860-869.
  33. Lai J, Ma J, Swigonova Z, Ramakrishna W, Linton E, Llaca V, Tanyolac B, Park Y-J, Jeong O-Y, Bennetzen JL, Messing J: **Gene loss and movement in the maize genome.** *Genome Res* 2004, **14**:1924-1931.
  34. Mondragon-Palomino M, Meyers BC, Michelmore RW, Gaut BS: **Patterns of positive selection in the complete NBS-LRR gene family of *Arabidopsis thaliana*.** *Genome Res* 2002, **12**:1305-1315.
  35. Richter TE, Pryor TJ, Bennetzen JL, Hulbert SH: **New rust resistance specificities associated with recombination in the *Rp1* complex in maize.** *Genetics* 1995, **141**:373-381.
  36. Song R, Llaca V, Linton E, Messing J: **Sequence, regulation, and evolution of the maize 22-kD zein gene family.** *Genome Res* 2001, **11**:1817-1825.
  37. Force A, Lynch M, Pickett FB, Amores A, Yan Y-L, Postlethwait JH: **Preservation of duplicate genes by complementary, degenerative mutations.** *Genetics* 1999, **151**:1531-1545.

38. Langham RJ, Walsh J, Dunn M, Ko C, Goff SA, Freeling M: **Genomic duplication, fractionation and the origin of regulatory novelty**. *Genetics* 2004, **166**:935-945.
39. Vision TJ, Brown DG, Tanksley SD: **The origins of genomic duplications in *Arabidopsis***. *Science* 2000, **290**:2114-2117.
40. Blanc G, Hokamp K, Wolfe KH: **A recent polyploidy superimposed on older large-scale duplications in the *Arabidopsis* genome**. *Genome Res* 2003, **13**:137-144.
41. Guyot R, Keller B: **Ancestral genome duplication in rice**. *Genome* 2004, **47**:610-614.
42. Paterson AH, Bowers JE, Chapman BA: **Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics**. *Proc Natl Acad Sci USA* 2004, **101**:9903-9908.
- An analysis of the rice genomic sequence suggests that the rice genome underwent a whole-genome duplication some 70 million years ago and before the radiation of the grasses. All cereal crops, including those currently considered diploids, are thus potentially of tetraploid origin.
43. Vandepoele K, Simillion C, Van de Peer Y: **Evidence that rice and other cereals are ancient aneuploids**. *Plant Cell* 2003, **15**:2192-2202.
44. Swigonova Z, Lai J, Ma J, Ramakrishna W, Llaca V, Bennetzen JL, Messing J: **Close split of sorghum and maize genome progenitors**. *Genome Res* 2004, **14**:1916-1923.
- Comparative sequence analysis of five duplicated regions of the maize genome dated the divergence of the ancestral sorghum and maize genomes to 11.9 million years ago, and the tetraploidization event to the time-span 11.9–4.8 million years ago. These values are considerably more recent than earlier projected datings.
45. Han F, Fedak G, Ouellet T, Liu B: **Rapid genomic changes in interspecific and intergeneric hybrids and allopolyploids of Triticeae**. *Genome* 2003, **46**:716-723.
46. Kashkush K, Feldman M, Levy AA: **Gene loss, silencing and activation in a newly synthesized wheat allotetraploid**. *Genetics* 2002, **160**:1651-1659.
47. Fu H, Dooner HK: **Intraspecific violation of genetic colinearity and its implications in maize**. *Proc Natl Acad Sci USA* 2002, **99**:9573-9578.
48. Peng JR, Richards DE, Hartley NM, Murphy GP, Devos KM, Flinham JE, Beales J, Fish LJ, Worland AJ, Pelica F *et al.*: **'Green Revolution' genes encode mutant gibberellin response modulators**. *Nature* 1999, **400**:256-261.
49. Paterson AH, Lin Y-R, Li Z, Schertz KF, Doebley JF, Pinson SRM, Liu S-C, Stansel JW, Irvine JE: **Convergent domestication of cereal crops by independent mutations at corresponding genetic loci**. *Science* 1995, **269**:1714-1718.
50. Hu FY, Tao DY, Sacks E, Fu BY, Xu P, Li J, Yang Y, McNally K, Khush GS, Paterson AH, Li Z-K: **Convergent evolution of perenniality in rice and sorghum**. *Proc Natl Acad Sci USA* 2003, **100**:4050-4054.
51. Doust AN, Devos KM, Gadberry MD, Gale MD, Kellogg EA: **Genetic control of branching in foxtail millet**. *Proc Natl Acad Sci USA* 2004, **101**:9045-9050.
- In maize, *tb1* is the main player determining the transition from the wild to cultivated phenotype. Doust and colleagues showed that the *tb1* ortholog only plays a minor role in determining vegetative branching in foxtail millet, another Panicoideae species. This suggests some level of functional divergence between orthologous genes across species.
52. Leister D, Kurth J, Laurie DA, Yano M, Sasaki T, Devos KM, Graner A, Schulze-Lefert P: **Rapid reorganization of resistance gene homologues in cereal genomes**. *Proc Natl Acad Sci USA* 1998, **95**:370-375.
53. Bruggeman R, Rostoks N, Kudrna D, Kilian A, Han F, Chen J, Druka A, Steffenson B, Kleinhofs A: **The barley stem rust-resistance gene *Rpg1* is a novel disease-resistance gene with homology to receptor kinases**. *Proc Natl Acad Sci USA* 2002, **99**:9328-9333.
54. Gallego F, Feuillet C, Messmer M, Peng A, Graner A, Yano M, Sasaki T, Keller B: **Comparative mapping of the two wheat leaf rust resistance loci *Lr1* and *Lr10* in rice and barley**. *Genome* 1998, **41**:328-336.
55. Chen H, Wang S, Xing Y, Xu C, Hayes PM, Zhang Q: **Comparative analyses of genomic locations and race specificities of loci for quantitative resistance to *Pyricularia grisea* in rice and barley**. *Proc Natl Acad Sci USA* 2003, **100**:2544-2549.
56. Sutton T, Whitford R, Baumann U, Dong C, Able JA, Langridge P: **The *Ph2* pairing homoeologous locus of wheat (*Triticum aestivum*): identification of candidate meiotic genes using a comparative genetics approach**. *Plant J* 2003, **36**:443-456.
57. Gottwald S, Stein N, Börner A, Sasaki T, Graner A: **The gibberellic-acid insensitive dwarfing gene *sdw3* of barley is located on chromosome 2HS in a region that shows high colinearity with rice chromosome 7L**. *Mol Genet Genomics* 2004, **271**:426-436.
58. Devos KM, Beales J, Nagamura Y, Sasaki T: ***Arabidopsis*—rice: will colinearity allow gene prediction across the eudicot-monocot divide?** *Genome Res* 1999, **9**:825-829.
59. Mayer K, Murphy G, Tarchini R, Wambutt R, Volckaert G, Pohl T, Dusterhoff A, Stiekema W, Entian KD, Terry N *et al.*: **Conservation of microstructure between a sequenced region of the genome of rice and multiple segments of the genome of *Arabidopsis thaliana***. *Genome Res* 2001, **11**:1167-1174.
60. Vandepoele K, Saeys Y, Simillion C, Raes J, Van de Peer Y: **The automatic detection of homologous regions (ADHoRe) and its application to microcolinearity between *Arabidopsis* and rice**. *Genome Res* 2002, **12**:1792-1801.