Comparative genetics in the grasses

Katrien M. Devos and Michael D. Gale
John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, UK

Key words: comparative genetics, synteny, colinearity, grasses

Abstract

Comparative genetic studies have demonstrated that gene content and orders are highly conserved, both at the map and megabase level, between different species within the grass family. Integration of the genetic maps of rice, foxtail millet, sugar cane, sorghum, maize, the Triticeae cereals and oats into a single synthesis reveals that some chromosome arrangements characterise taxonomic groups, while others have arisen during or after speciation. A detailed analysis of the comparative maps of seven species, belonging to three subfamilies, and their applications are described below.

Introduction

History of comparative genetics

‘Comparative genetics’ is the science that exploits the results of ‘comparative mapping’ – two terms that were unknown to plant geneticists ten years ago. Although the discovery that genes in related species – and as it turns out now, quite distantly related species – tend to be ordered colinearly on chromosomes is quite new, the concept of conservation has a long history in plant genetics. In the 1920s Vavilov had already observed that ‘similar variations’ were to be found in different species. More recently, of course, DNA studies have shown that genes of similar function in different species have remarkably conserved sequences. The discovery of colinearity by comparative mapping is, however, a function of the new molecular markers – in particular RFLPs – employed in plants for the first time in the mid 1980s. Nevertheless, with the immense effort in genetic mapping of morphological mutants and, later, protein loci in both plants and animals it is surprising that gene colinearity remained hidden until the reports in 1988 of convergence of the maps of the three genomes of hexaploid bread wheat [6] and the similar convergence between the maps of tomato and potato [3].

In wheat, evidence for homoeology between the three genomes had already been provided by cytogeneticists, particularly by the late Prof. Ernie Sears, who assembled the first set of aneuploid genetic stocks in wheat. Hexaploid bread wheat (Triticum aestivum, 2n = 6x = 42) is an allopolyploid with genome constitution AABBDD, formed through hybridization of T. urartu (AA) with a B genome diploid of unknown origin, and subsequent hybridization, only about 8000 years ago, with a D genome diploid, T. tauschii. The development of aneuploid stocks in wheat led to the discovery that an extra dose of a particular chromosome could compensate for the absence of another [57]. This compensating ability of chromosomes of different ancestral origin defined their relationship, and resulted in the classification of the 21 wheat chromosomes into 7 homoeologous groups [58]. Similar compensating experiments determined the homoeologous relationships between wheat chromosomes and those of other Triticeae species such as rye, barley, and several Aegilops ssp. (for catalogue see 60). The ability of chromosomes to substitute for one another suggested that they carried similar genes. This was confirmed by the demonstration that the glutenin storage protein loci were to be found on chromosomes 1A, 1B and 1D by Shepherd [59]. Following this aneuploid analysis of other biochemical markers, and later of DNA markers, showed that most protein, isozyme and RFLP loci
are triplicated in the wheat genome. The realisation that not only gene content but also gene orders were conserved had, however, to await the construction of molecular marker-based genetic maps.

**Comparative genetics at the map level**

**Within hexaploid wheat**

The application of RFLP markers in plant genetics provided, for the first time, unlimited supplies of markers of both genic (cDNA probes) and random genomic (genomic DNA probes) origin to detect variation at the DNA level. Hybridization of cDNA clones to aneuploid lines of hexaploid bread wheat showed that most genes were triplicated on the A, B and D genomes (Figure 1a). Generally, these probes cross-hybridized strongly under high stringency conditions to wheat relatives such as rye and barley. Genomic clones, on the other hand, either showed a hybridization behaviour comparable to cDNA clones, or hybridized to one or more chromosomes apparently at random (Figure 1b). The latter ‘non-homoeologous’ probes were found to be derived from fast evolving regions of the wheat genome [30] (Figure 1d), and it was therefore not surprising that these sequences could not be detected in other Triticaceae species under stringent hybridization conditions [16]. Genetic mapping of cDNA-like or homoeologous sequences in hexaploid bread wheat clearly demonstrated that, taking into account evolutionary chromosomal rearrangements involving chromosome arms 4AS, 4AL, 5AL and 7BS [41, 15], and possibly 2BS and 6BS [16], gene orders on the A, B and D genomes were extremely highly conserved.

In recent years molecular marker technology has moved to exploit the polymerase chain reaction (PCR). Microsatellites are now, in fact, the markers of choice for marker aided applications in breeding programmes. In contrast to RFLPs these systems, which rely on complete sequence identity between primer and target DNA, tend to fail to detect homoeoloci across the three wheat genomes (Figure 1c). This lack of transferability of PCR-based markers across different genomes excludes their use for comparative purposes. The evidence for the extensive synteny across the grasses discussed below stems almost entirely from cross-mapping of RFLPs.

**Within the Poaceae family**

The high levels of conservation of gene orders within polyploid wheat begged the question of how far synteny would extend. In the first instance, analyses were focused on genomes of species belonging to the same tribe with major efforts within the Triticaceae (Figure 2) and Andropogonodae tribes. Later, the leading laboratories in grass genome research exploited RFLPs as a generic tool to broaden their purview of crop species beyond tribe boundaries (Figure 3). Cornell graduated from rice to wheat, barley, maize, sugar cane and oats, while the John Innes Centre (JIC) expanded from wheat to barley, rye, rice, maize, pearl millet and foxtail millet. Where experience was not immediately available collaborations provided the necessary materials and information. Links between the Brookhaven National Laboratory maize group and Cornell and the Japanese Rice Genome Programme and JIC were key collaborations which led to the present-day synthesis.

The first publications that demonstrated the breadth of synteny among the cereals were by Ahn and Tanksley [1] showing the relationship between rice and maize, Kurata et al. [39] showing that the wheat genome could be aligned with rice, and Moore et al. [45] who showed that all the maps could be combined in a single synthesis. A recent evaluation of the extent of synteny is shown in Figure 4 which draws on many, some as yet conflicting, sources of information.

The alignment in Figure 4 is based on rice, with the other genomes arranged relative to rice in the most parsimonious manner. Rice is used as the base simply because it is the smallest cereal genome analysed in detail, and that for which we have the densest maps and most genomic tools available. It is important to note that the organisation of Figure 4 has no bearing on the ancestry of the grass genomes. Inferences can, of course, be made concerning which chromosomal arrangements are the more primitive.

The integrated grass genome map, which now includes species belonging to 6 different tribes and 3 different subfamilies, reveals three distinct genome patterns, provided by rice, a representative of the Bambusoideae, oats and the Triticaceae crops of the Pooideae, and several members of the Panicoideae (Figure 5).

**Rice.** The rice chromosomes are arranged in a circle in the order revealed by Moore et al. [46] following a comparison of the rice and maize genomes. Key information concerning the rice genome map came from the Rockefeller Rice Biotechnology Programme map
Figure 1. The chromosomal location in wheat of loci detected with a ‘homoeologous’ RFLP probe (a), a non-homoeologous RFLP probe (b), a microsatellite, all using nullisomic-tetrasomic aneuploid lines (c), and their map position on the homoeologous group 3 chromosomes (d).
Figure 2. Hybridization of a wheat cDNA probe, PSR129, to the wheat variety Chinese spring (CS), wheat nullisomic-tetrasomic lines (N7A, N7B, N7D), barley (cv. Betzes), a wheat/rye amphiploid (CS/Imperial), a CS/Ae. umbellulata amphiploid, and wheat/barley, wheat/rye and wheat/Ae. umbellulata disomic single chromosome addition lines reveals loci on homoeologous chromosomes in the four Triticeae species.

Figure 3. Differential hybridization behaviour of probes in a range of Triticeae crop species, rye grass (Lolium), maize and pearl millet (Pennisetum).

made at Cornell by Susan McCouch and colleagues in an interspecific backcross population derived from O. sativa and O. longistaminata, two AA genomes species (5 and updates in RiceGenes, the rice genome database) and Takuji Sasaki’s group at the Japanese Rice Genome Program in Tsukuba, using an indica ×
Figure 4. Aligned maps of rice, foxtail millet, sugar cane, sorghum, maize, the Triticeae crops and oats. Chromosome nomenclature: Chromosomes are numbered conventionally – rice (short and long) after Singh et al. [61]; foxtail millet (top and bottom) after Wang et al. [66]; sugar cane linkage groups (top and bottom) after Dufour [20]; sorghum (top and bottom) after Dufour [20]; oats (top and bottom) after O’Donoughue et al. [53]. All maps are aligned relative to rice, ends of maps of chromosomes (in some cases telomeres) are shown with coloured arrows. Where known, the positions of centromeres are marked. Double ended arrows show inversions (relative to rice), open arrows show segments that require transposition to return to linear chromosome maps. Hatched areas indicate regions of uncertainty where the cross-mapping data does not allow definite alignment with rice. Arrows in solid red indicate ‘evolutionary’ transpositions which define the Panicoideae and Pooideae subfamilies. The intra-genomic duplication of R11b and R12c is considered as part of R12 only in this analysis. All references are in the text.
japonica O. sativa cross (40 and updates in the JRGP database). These two maps have been merged [43] to provide in excess of 2000 loci. Discrepancies in marker orders between the maps are few. An earlier comparison between a map of the cultivated AA-genome species O. sativa and the wild CC-genome O. officinalis showed an 8.6% reduction in map length but marker orders that were mostly conserved. Two large inversions were observed in rice chromosome 1, and several markers appeared to be translocated between rice 11 and 12 [34].

Locations of the centromeres on the rice maps have been provided by analysis of tertiary trisomic stocks by Singh et al. [61]. The construction of detailed maps has also revealed a number of segmental duplications within the largely diploid rice genome. More and more of these are becoming recognised, but only one major duplication involving the distal short arm segments of R11 and R12 [47] is taken into account in Figure 4. Duplications within the rice and other grass genomes may confound an unequivocal alignment of the species chromosomes.

_Triticeae_. Within the Triticeae tribe, comparative research has determined precise relationships among the genomes of _T. tauschii_ [26], barley [65, 48], rye [56, 12], _T. monococcum_ [19], and most recently, _Aegilops umbellulata_ [69], all of which share a basic chromosome number of 7 and a monoploid DNA content of about 5.5 pg. The genomes of wheat, _T. monococcum_, _T. tauschii_ and barley were shown to be highly colinear [62, 16, 48, 19], although the presence of a few minor inversions is still being debated. The genomes of rye and _Ae. umbellulata_, on the other hand, are distinguished from the wheat genomes by a minimum of 7 and 11 rearrangements respectively [12, 69]. Colinearity, however, was conserved within the translocated segments. The presence of extensive rearrangements in the rye and _Ae. umbellulata_ genomes was surprising, and there is currently no answer as to why certain species accumulate and fix rearrangements more readily than others. It is clear that this differential rate of accumulation of structural changes excludes sheer numbers of rearrangements as a measure of evolutionary age.

The Triticeae maps in Figure 4 are a consensus of wheat, barley and rye and the arrangement shown is that of the D genome of the wheat variety Chinese Spring. The main wheat maps providing RFLP clones for comparative mapping were developed at JIC in a Chinese Spring × Synthetic population [13, 16, 68, 35, 15, summarised in 25] and by the Cornell and Clermont Ferrand groups in the International Triticeae Mapping Initiative population, derived from the cross Opata × Synthetic [62, 50, 51, 49, 42]. As yet there are too few points of correspondence to allow these maps to be closely aligned. The barley maps of the North America mapping group [38] and the group at Munich

Figure 5. Taxonomic relationships within the grass family.
Ae. umbellulata characterise rye and iceae tribe. Other rearrangements, such as those that rearrangements characterise the species within the Triticeae tribe. Other rearrangements, such as those that characterise the group, each species has further rearrangements which distinguish it from the others.

**Oats.** The oat map shown in Figure 4 is from a cross between two diploid species Avena atlantica and A. hirtula by O’Donoughue et al. [53] which has since been populated with comparative wheat, rice, barley and maize markers by Van Deynze et al. [63] and comparisons drawn with rice, maize and wheat. Hexaploid oats (AACCDD) is characterised by multiple intergenomic translocations, at least some of which have occurred after the formation of the AACC tetraploid [33, 52]. A number of duplications are present in both the diploid and hexaploid oat maps, indicating that they preceded polyploidization.

The analysis by Van Deynze et al. [63] shows that the diploid oat genome can be aligned with that of rice in only 20 syntenic regions, and with wheat and maize in 18. Many of the rearrangements, relative to the other grasses, are however novel. Exceptions are the insertions of R10 into R5 and the R8/R6 rearrangement that characterise both oats and the Triticeae and demonstrates their relationship as members of the Pooidae.

**The Panicoideae group.** The foxtail millet, sugar cane, sorghum and the two maize genomes shown in Figure 4 are characterised by two common rearrangements relative to rice. These involve, once again, R10 which is inserted into R3, and R9 which is found inserted into R7. While these evolutionary translocations characterise the group, each species has further rearrangements which distinguish it from the others.

**Maize.** The main sources of mapping information are from the USDA group led by Ed Coe in Columbia, Missouri (10 and updates in MaizeDB), Ben Burr’s programme at the Brookhaven National Laboratory (4 and updates in MaizeDB) and Tim Helentjaris [31]. The comparative information has been provided by Ahn and Tanksley [1] and Devos et al. [14] for maize chromosome 9. Other information used is from Sorrells et al. [63, 64] and the CIRAD group at Montpellier [20, 22].

The key observation is that maize is clearly a tetraploid, probably of relatively recent origin since most DNA probes hybridise to duplicated loci. Doebley dates the polyploidisation of maize at 16 million years ago. However, the two ‘genomes’, each with five chromosomes are, in contrast to wheat, extremely differentiated in that no completely ‘homoeologous’ chromosomes remain. The fact that most of the larger maize chromosomes (1, 2, 3, 4 and 6) make up one genome and the smaller chromosomes (5, 7, 8, 9 and 10) make up the other, may relate to differences retained from the time of allotetraploidization. However, the chromosome size differences could also be fortuitous. A comparison at the sequence level of duplicated genes indicated that some duplicated loci appear to have diverged less (16 million years) than others (27 million years) (Doebley, pers. comm.). This could have been achieved if some chromosome pairing and recombination took place before ‘homoeologous’ pairing was restricted, resulting in segments of the duplicated genome having been retained from the two ancestral five-chromosome diploid parents and other parts representing duplicated segments from one or other of genomes which have diverged only in the intervening 16 million years. This theory that maize is a ‘segmental’ allotetraploid will be supported if the two types of genes are found to lie in linkage blocks, but more evidence is yet required.

Several rearrangements and inversions, relative to both wheat and rice, are apparent in addition to the common Panicoideae group translocations and these would appear to have occurred after the divergence of maize from other members of the subfamily.

**Sorghum.** Within the Andropogonodae, comparative research was undertaken to increase the understanding of sorghum genetics. Several groups established the relationship between the sorghum and maize genomes following the lead of Hulbert et al. [32, 67, 44, 55, 29, 22, 20, 54]. The most complete information to date is provided by Dufour et al. [20] and their synthesis and chromosome numbering and orientations are used in Figure 4. Gene orders appear to be largely conserved between sorghum and maize and only a limited number of rearrangements have been identified. With the exception of major evolutionary translocations which characterise the Panicoideae, extreme colinearity also appears to have been maintained with rice, with seven
linkage groups reflecting precisely, within the limits of the analysis to date, seven rice chromosome linkage blocks.

The close taxonomic relationship between maize and sorghum, both with a 10 chromosome haploid complement, and maize having been shown clearly to be an allotetraploid as of 16 million years ago, begs the question as to whether sorghum is also a tetraploid. Although there are clearly some regions of duplication [55, 9, 20, 21], the evidence for tetraploidy, at least of recent origin, is not strong. In both Chittenden et al. [9] and Dufour et al.’s [20] analysis only 8% and 11% respectively of duplicated loci were present on the map, although the number of probes that detected more than one hybridizing fragment amounted to 23% [20]. In maize the equivalent figure is about 70% [67, 18]. Certainly comparative maps of maize and sorghum show single sorghum linkage groups relating to, usually, two regions of the maize genome, as indicated in Figure 4. For these reasons sorghum is maintained as a 2x = 2n = 20 diploid in this review.

Sugar cane. Sugar cane, another important crop belonging to the Andropogonodae, has a complex polyploid nature and mapping studies were initiated at CIRAD to help define its intricate genome structure [29]. Modern sugar cane varieties have originated through introgression of part of the genome of the wild species Saccharum spontaneum (x = 8, 2n = 40 to 128) into S. officinarum (x = 10, 2n = 80). Introgressions are mostly complete chromosomes but some intergenicomic translocations have been observed [11]. Due to its genome complexity, the numbers in Figure 4 refer to composite linkage groups rather than chromosomes. Extensive colinearity was maintained between sugar cane and sorghum.

Foxtail millet. All of the genome information relating to Setaria italica (2x = 2n = 18) derives from the work of Zhimin Wang from Hebei Academy, China during his PhD work at JIC. In addition to being a crop of importance in Northern China, foxtail millet is very attractive as a genetic organism since it has a genome almost as small as rice with 1C = 490 mb DNA (O. Panaud, pers.comm.). Maps were constructed in an intraspecific population between two S. italica accessions and an interspecific cross, S. italica × S. viridis. Both maps had equivalent genetic lengths and displayed highly conserved marker orders [66]. The foxtail millet genome also shows extreme colinearity with rice [17] with five entire chromosomes unchanged in marker content and order. More interesting possibly is the very close relationship between foxtail millet, which has a haploid chromosome number of 9, and sorghum with n = 10. The difference in chromosome number is accounted for by the synteny of foxtail millet chromosome III with sorghum chromosomes E and I. Elsewhere we can currently detect only one inversion, in sorghum chromosome D, and one translocation involving foxtail millet chromosomes III and VII, which differentiate the two species.

Centromeres. The comparative maps of genomes with chromosome numbers varying between 5 and 12 raises the question as to whether centromeres are syntenic. Plainly they cannot all be completely so.

Precise positioning of the centromeres on the maps is limited to wheat, maize and rice. Only in these three species have the appropriate genetic stocks been available to map RFLP markers to chromosome arms and then to delineate, often very precisely, the locations of the centromeres. A comparison of maps and centromere positions across these three species indicates that the centromeres of maize chromosomes 1, 2, 3, 5, 7, 8, 9 and 10 correspond to those of rice chromosomes 8, 4, 1, 2, 7, 1, 6 and 4 respectively. Conservation of colinearity of the centromeres of maize chromosomes 4 and 6 across rice could not, as yet, be established, but it is expected that they correspond with the centromeres of either rice chromosomes 2 or 11, and 5 or 6 respectively. Similarly, colinearity between the centromeres of wheat and rice can be established for five of the seven wheat chromosomes. Centromeres of wheat chromosomes 2, 3, 4, 6 and 7 appear to be colinear with those of rice chromosomes 7, 1, 3, 2 and 8.

Conservation of colinearity of centromeres across grass species could be exploited to predict their location in species such as the millets for which the appropriate genetic stocks are not available. Using the comparative approach, putative centromere positions were established for foxtail millet chromosomes I, II, IV, V, VI, VII, VIII and IX [17]. Particularly on chromosomes I, IV, V, VI and IX, the putative centromere regions coincide with regions of severe marker clustering. This is a further indication that these regions indeed contain centromeres, following the observations in wheat that reduced recombination results in clustering of markers on genetic maps around centromeres [7]. Similar arguments have been used to predict the locations of the oat centromeres [63], five of which are shown in Figure 4.
Recent research has identified cereal centromere-specific probes [2, 36] and microsatellites that are indicative of centromere positions [27]. It may thus soon be feasible to map both functional and redundant centromeres. This will go some way towards answering the question of ancestry since, if rice has retained the more primitive grass genome organisation, then wheat, for example, should carry five redundant centromeres in addition to the seven functional centromeres. Moreover these should all be present at predictable locations in the Triticeae genomes.

Ancestry. The question as to exactly what the genome of the ancestral grass was like is still open. Of those grasses studied to date for synteny, rice has the smallest genome which may be an indicator, if we assume that genomes accumulate repeats, as demonstrated by Bennetzen [8]. Another pointer may lie in the chromosomal rearrangements that characterise the Pooidae and Panicoideae subfamilies relative to rice, shown in red on Figure 4. Each involves the insertion of one rice linkage group into the centromeric region of another. Rice chromosome 10 is involved in two of these key rearrangements, inserted in maize into R3 and in wheat into R5, which may argue that an independent R10 block is the more primitive. Certainly the argument by Moore et al. [46], which uses the comparison of the rice and maize genomes and shows that the rice genome can be equated to a single chromosome stack of blocks is intriguing. However we must await further research before assuming there was indeed a single ancestral grass chromosome.

Comparative genetics at the level of gene organisation

As discussed above there is now overwhelming evidence for the existence of extensive regions of conserved colinearity among grass species at the genetic map level. This knowledge can already be exploited to advance marker studies on all grass species and to extend our knowledge of key syntenic agronomic genes as they become placed on the genetic maps. Application of comparative genome research as a tool for gene isolation, however, also requires that conservation of gene presence and organisation is maintained at the megabase level. Dunford et al. [23] demonstrated that the order of markers, contained on the same YAC and thus separated by $\leq 1$ mb in rice, was indeed maintained in barley. In a similar comparative study of a region of barley chromosome 7H with rice chromosome 6, Kilian et al. [37] also concluded that a very high degree of synteny exists at the sub-cM level. Foote et al. [24] have similarly shown extensive colinearity between a small region of wheat chromosome 5 and rice chromosome 9. The similarity in gene content and order are adequate to allow Andris Kleinhof’s group to continue their search for the barley rust resistance gene, $Rpg1$, in rice and Graham Moore’s group to continue towards the wheat gene controlling homoeologous chromosome pairing, $Ph1$, from rice.

These studies have, however, also shown disruption of simple relationships, and small deviations from strict colinearity are likely to be common. Kilian’s study has revealed a duplicated segment of their target region proximal to the site of the $Rpg1$ gene in barley. Foote et al. have shown that part of their target region is duplicated on rice chromosome 9, and that duplications and, possibly, triplications may disrupt synteny at the microlevel in wheat. The same study also revealed that close by the target region the syntenic relationship wheat 5 – rice 9 was disrupted by a short segment with homoeology to rice chromosome 11 intercalated into the larger R9 region in wheat and barley. In a detailed study of the $Sh2$ – $al$ region in maize, sorghum and rice, Chen et al. [8] demonstrated a direct duplication of $al$ at the homoeolocus in sorghum. The $Sh2$ – $al$ region, however, exhibited extensive colinearity, but spanned 140 kb in maize, and only 19 kb in rice and sorghum. Sequence homology between $Sh2$ and $al$ in maize, sorghum and rice amounted to at least 80%. Most of the intergenic regions, however, were populated with species-specific repetitive elements and failed to cross-hybridize across species [8]. This result serves to remind us that the comparative story that is emerging relates only to genes. Intergenic regions are likely to have accumulated species-specific sequences which prohibits prediction of physical distances between homoeologous genes in related species.

Based on the current comparative data, it is clear that gene content and gene orders are highly conserved between species within the grass family, and that the amount and organisation of repetitive sequences has diverged considerably. Rearrangements, both at the map and megabase level have taken place, but 60 million years of evolution has left large chromosome regions apparently untouched.

Conclusions

The stage is now set to exploit the synteny between grass genomes in many ways. Moreover, more species will be added. The comparative maps of pearl mil-
let, *Pennisetum glaucum*, finger millet, *Eleusine coracana* and rye grass, *Lolium perenne*, will soon be available, the latter from work currently being completed at the Institute of Grassland and Environmental Research, Aberystwyth and the former two projects are from work at the John Innes Centre. In the immediate future we can expect successes in cross-genome map-based cloning, but in the longer term we can expect an increased level of understanding from a coalescing of information about genes and traits as key agronomic genes in different species become linked on the comparative map.

Figure 6 shows a few such genes which have already been demonstrated to show synteny across the three major cereals. If we assume that a cereal genome carries about 25,000 genes, as has been predicted in *Arabidopsis*, then we should be able to align homoeoalleles across the grass genomes on 25,000 radii on the circles. We can expect the mapping to continue apace, but integration and applications will be hampered without advances in bioinformatics. The amount of data in individual species genome projects is overwhelming. The total data for all grass programmes is vast and requires a new approach to bioinformatics that is compatible with the species-specific databases, such as GrainGenes, MaizeDB, RiceGenes, MilletGenes, which already serve the community.

This conservation of genome organisation has since been observed in other groups of species, for example over various legume crops and the brassicas, including *Arabidopsis* which is a member of the Cruciferae. Possibly the next ten years may reveal that conservation of colinearity is maintained across the monocots and dicots, and maybe even further.
References

35. Jia J, Miller TE, Reader SM, Devos KM, Gale MD: RFLP- based maps of the homoeologous group 6 chromosomes of...


65. Xie DX, Devos KM, Moore G, Gale MD: RFLP-based genetic