Chapter VII

Molecular cytogenetics of the apomixis controlling locus in *Paspalum simplex*

Ornella Calderini, Song B. Chang, Hans de Jong, Annalisa Bietta, Sergio Arcioni, María Eugenia Caceres, Camilo L. Quarin, Diego H. Hojsgaard, Iain S. Donnison & Fulvio Pupilli

Apomictic plants give rise to seeds that are genetically identical to the mother plant. To characterize cytologically the locus controlling the apomictic reproduction of *Paspalum simplex*, FISH analysis of meiotic chromosomes was performed using an apomixis-linked BAC clone as a probe. The locus was embedded in a heterochromatin-poor region not adjacent to the centromere. Sequence analysis of the probe used confirmed the strong synteny between the apomixis locus of *P. simplex* and the telomere of the long arm of chromosome 12 of rice. Nevertheless, both large- and small-scale rearrangements of the region compared to the related area of rice were revealed. Results are discussed in relation to the use of rice genomic data for positional cloning of apomixis genes and to the possible role of gene rearrangements in the apomictic development of *P. simplex*.

KEYWORDS: apomixis, comparative sequencing, positional cloning, synteny.
INTRODUCTION

The alternation of gametophytic and sporophytic generations and the fusion of parental gametes maintain genetic variability in most living organisms. Gametophytic apomixis is an alternative plant reproductive strategy that bypasses meiosis and fertilization to give rise to seeds that are genetically identical to the mother plant, via the intermediate formation of an unreduced gametophyte (Nogler, 1984).

The potential benefits of apomixis to agriculture are enormous, especially for those crops in which the production of hybrid seed is important. Unfortunately, although apomictic reproduction is under genetic control (Savidan, 2000), the trait is present only in wild species, which makes the application of commonly used molecular techniques difficult. Furthermore, no genuine apomictic mutants have so far been recognized in model systems such as Arabidopsis, maize or Petunia. Nevertheless, several gametophytic mutants that at least resemble some aspects of apomictic reproduction have been reported in Arabidopsis and maize (Grossniklaus, 2001; Koltunow & Grossniklaus, 2003; Curtis & Grossniklaus, 2007, chapter 3). With continuing progress in large-scale sequencing, comparative mapping and reverse genetics, it is now becoming feasible to study apomixis in its natural biological background. The aim of our research is to identify the genetic determinants of apomictic reproduction in a wild apomictic species with the aim of introducing the trait into crop species.

Paspalum simplex, our chosen organism, comprises sexual diploids as well as apomictic polyploid cytotypes, with ploidy levels ranging from 3x to 6x, with the tetraploid being the most common (2n = 4x = 40) and x = 10 the base chromosome number. The apomictic cytotype is aposporous: the megaspore mother cell usually degenerates and a cell from the nucellus gives rise to an unreduced 8-nucleate embryo sac resulting from three successive mitoses (Caceres & al., 1999). Another characteristic of the P. simplex apomictic gametophyte is the presence, in a variable percentage of ovules, of multiple embryo sacs, which contrasts with sexual ovules that always carry a unique and reduced embryo sac (Caceres & al., 1999). Paspalum simplex has a number of characteristics that make it amenable to a map-based strategy to identify genes involved in apomictic reproduction. The most important of these characteristics are that sexual diploid races can be crossed with polyploid apomictic cytotypes and that the species has a relatively small genome size (0.75 pg per haploid genome [Caceres & al., 1999]). We present here data on the genetic characterization of the Apomixis Controlling Locus (ACL) in P. simplex, together with recent DNA sequence analysis and candidate gene identification.

COMPARATIVE MAPPING

The genetic control of apomixis in P. simplex is consistent with a single dose dominant allele. Strongly distorted segregation ratios detected in segregating pop-
ulations were always skewed in favour of sexual genotypes. This suggests that the allele has either a lethal pleiotropic effect with incomplete penetrance, or that it is linked to a lethal factor that, in homozygous condition, causes gamete death (Pupilli & al., 2001). The ACL of \textit{P. simplex} is located on a relatively large chromosomal interval that is characterised by suppression of recombination and shows synteny with the telomeric part of the long arm of rice chromosome 12 (RC12), (Pupilli & al., 2001). This synteny has allowed the use of rice molecular markers, seven of which are linked to apomixis in \textit{P. simplex}; these markers span 14.5 cM in rice and we hypothesised that this approximates to the size of the ACL in \textit{P. simplex}. Since 1 cM corresponds roughly to 1 Mb in rice, we speculate that the physical size of the \textit{P. simplex} ACL may be in the order of 15 Mb or more. Since we are unable to finely-map this region, we are obliged to physically dissect the ACL, cloning and sequencing large DNA fragments in order to identify the possible genetic determinants of apomixis. Apomixis-specific SCARs (sequence characterized amplified regions) have already been isolated in \textit{P. simplex} and were used to screen a BAC library of the same species. Partial sequencing of positive BACs revealed a high density of non-coding regions, transposable elements while genes are rarely found. Large-scale sequencing of all BACs encompassing the ACL in \textit{P. simplex} is, therefore, likely to be a laborious, expensive, and a rather fruitless effort. The identification of ACL sequences that are conserved among different apomictic species may help to identify BACs that are more likely to contain candidate genes for apomixis. On the basis of these considerations we decided to investigate the conservation of marker order in the ACL among different \textit{Paspalum} species to identify sub-regions of the ACL in \textit{P. simplex} that are conservatively linked to apomixis to reduce the chromosome area under investigation.

Although there is remarkable synteny between the ACL region of \textit{P. simplex} and the telomeric part of the long arm of RC12, the vast majority of rice markers that mapped to this region were unlinked to apomixis. This indicates that, although no gross rearrangements occurred in this area since rice and \textit{Paspalum} diverged, a relatively high number of small-scale rearrangements may have occurred. Several kinds of rearrangements were detected at the ACL of \textit{Paspalum}; the extent of variation detected among some \textit{Paspalum} spp. was correlated with the genetic relatedness of the species analysed (Pupilli & al., 2004). The hybridising banding pattern of the rice probes R1759, C1069 and C454 and of the two AFLP-derived homologous probes (Pupilli & al., 2004) showed identical apomixis-linked fragments in the two species analysed, \textit{P. simplex} and its closely related species \textit{P. malacophyllum}. This contrasted with results obtained using the C996A probe, which showed apomixis-linked bands of different molecular weight (Pupilli & al., 2004). This indicates that small-scale rearrangements occurred, which resulted in a gain or loss of restriction sites at this locus. A comparison of the ACL structures in \textit{P. simplex} and the distantly-related species \textit{P. notatum} showed both small- and large-scale rearrangements: small-scale rearrangements were revealed by differences in the molecular weight of apomixis-related bands, whereas the large-scale rearrangements were due to a loss of synteny between the two distal markers.
R1759 and C901, and the other apomixis-linked markers derived from RC12, together with the gain of two other markers, which mapped to RC2. Linkage analysis of the homologous sexual counterpart of the ACL in *P. notatum* showed a strict colinearity of markers with the top of the long arm of RC12, including markers (C1069 C454 and C996A), which, in apomictic *P. notatum*, were linked to the two markers derived from RC2 (Fig. 1). This indicates that the large-scale rearrangements that are related to apomictic reproduction in *P. notatum* took place after the divergence of *P. notatum* and *P. simplex*. The marker 1069, together with C454, C996A and the AFLP-derived marker EM180 delimit a more restricted region of the *P. simplex* ACL and showed linkage with the apomictic trait in three *Paspalum* species. This block spanned approximately six cM and is characterised by instability in *P. notatum*, where it is located in another chromosome position. This relocation has two main effects: (1) the suppression of recombination in an otherwise recombination-competent region in the vicinity of the insertion site, and (2) the induction of apomictic reproduction. The most straightforward hypothesis is that the genetic determinants contained in this block are necessary, but not sufficient, to induce apomictic reproduction in the *Paspalum* spp. analysed (Pupilli & al., 2004).

**PHYSICAL MAPPING**

The power of the map-based cloning is markedly reduced if the genetic trait of interest maps to chromosomal regions that lack recombination. To physically dissect such regions, genomic libraries containing large-inserts must be generated and
clones containing DNA sequences that span these regions must be identified. It has been postulated that the ACL is located within heterochromatic pericentromeric regions (Ozias-Akins & al., 1998; Pupilli & al., 2001), which may account for the suppression of recombination observed in this region. Chromosomal regions close to the centromere are known to be dense in heterochromatin and rarely participate in crossover events. However, syntenic relationships with model organisms suggest that the position of the ACL is actually (sub)distal to the centromere in Paspalum (Pupilli & al., 2001) and Brachiaria (Pessino & al., 1997). An alternative proposition suggests that crossover suppression may be due to an inversion or other chromosomal rearrangement that physically inhibit synapse formation in heterozygotes (Ozias-Akins & al., 1998; Pupilli & al., 2001; 2004; Goel & al., 2003). In such rearranged regions, chromosomes rarely pair, and when they do, crossovers give rise to dentric anaphase bridges or unbalanced chromosome segregation leading to sterility. Indeed, a heterozygous translocation has been identified at the ACL of P. notatum supporting this second hypothesis (Pupilli & al., 2004). The basis of the observed recombinational repression may have some bearing on the functional nature of the apomixis locus: if the ACL is embedded in a heterochromatic region, then the relevant genes are likely to be inactive (Bender, 2004), while genes located on segments of rearranged chromosomes may be deregulated; a gene or group of genes related to reproduction could be re-located to a different chromosomal region where local regulatory elements cause alterations in sexual development (Goel & al., 2003). Depending on the casual mechanism responsible for the apomictic trait, appropriate strategies should be adopted, firstly to identify the critical genes and their regulatory elements, and thereafter, using this knowledge, to engineer apomixis in sexual systems. To determine the physical nature of the ACL, fluorescent in situ hybridization (FISH) using pachytene chromosomes as targets, can be used to resolve the spatial relationship of adjacent probes and to visualize the dense heterochromatic regions of chromosomes (de Jong, 2003). Furthermore, FISH on late pachytene or even meiotic pro-phase stages of the cell cycle can provide information about chromosome pairing and chiasma formation within the region corresponding to the apomixis locus. FISH analysis was therefore used on P. simplex pachytene chromosomes to determine the structural context of the region (heterochromatin or euchromatin) and its relationship with the ACL. The overall organization of the P. simplex pachytene chromosomes was shown to contain densely DAPI-stained (DAPI = 4’, 6-diamidino-2-phenylindole) heterochromatic blocks in the pericentromeric regions, with a large number of smaller heterochromatic knobs dispersed along the proximal and distal euchromatic regions (Calderini & al., 2006). When a pachytene chromosome preparation from apomictic P. simplex was probed with the BAC 346H10 and the image captured with high threshold level of fluorescence signals, a single signal was clearly detectable in a weakly DAPI-stained region. Conversely, when the image was captured at low threshold level, many minor foci were observed all along the chromosomes (Calderini & al., 2006). These minor signals were thought to be due to repetitive sequences contained in the BAC 346H10, which hybridized across all chromo-
omes. Our FISH analysis of pachytene chromosomes in the apomictic *P. simplex* provides cytological evidences that: (i) the ACL is partially hemizygous and (ii) the ACL of *P. simplex* is located on a distal region of a chromosome, rather than on the densely DAPI-stained heterochromatic pericentromeric regions. However, due to the presence of many minor heterochromatic chromomeres that are dispersed along the chromosomes in this species, it is not possible to determine whether the ACL is located in the vicinity of, or even embedded within, one such heterochromatic knob. In a Metaphase-I complement, FISH analysis using BAC 346H10 as a probe, clearly revealed a small doublet on one of the two bivalents. Since this signal is adjacent to a chiasma, but not directly involved in crossing-over (in such a case it would be in the middle), we conclude that the ACL is located on two chromatids of an unbound arm of one of two chromosomes paired at the bivalent group (Fig. 2). It is well known that chromosome rearrangements, such as translocations and/or inversions, can result in local suppression of chromosome pairing (asynapsis), or reduced chiasma formation (desynapsis). Such inversions or translocations involving large chromosome segments may cause gamete sterility and are not likely to be transmitted to the progeny. However, smaller rearrangements can cause perturbations of chromosome pairing without affecting gamete viability. FISH analysis of

Fig. 2. Fluorescence in situ hybridisation of metaphase I chromosome complement of apomictic *P. simplex* probed with BAC 346H10: the double dots (arrow) show the signals of sister chromatids. A: higher magnification.
pachytene chromosomes showed that the ACL is located in the vicinity of a small asynaptic region. The suppression of recombination at the ACL of *P. simplex* is likely then to be caused by the missing capacity for local chromosome pairing, probably resulting from a nearby rearrangement (i.e., inversion and/or translocation). Asynapsis involving a single chromosome arm may also explain why molecular markers allelic to, or linked in repulsion to apomixis were not detected in the form of either heterologous or homologous co-dominant RFLPs (Pupilli & al., 2001) or AFLPs (Labombarda & al., 2002).

**SEQUENCE ANALYSIS AND COMPARISON**

Comparative mapping analysis has revealed regions of gene order conservation within the grass family (Bennetzen & Ramakrishna, 2002). Moore & al. (1995) showed that all the maps of the Gramineae can be combined to form a single map using rice as the base genome. The ongoing development of molecular tools for the analysis of model species, together with the release of the rice genome sequence, has provided the means by which to isolate a gene of interest in any crop, once the homoeologous region containing the gene has been identified using conserved rice molecular markers. In fact, one should be able to identify the locus of interest if that locus was located on a chromosome area rich in markers and genes. In general, the rice chromosomes are more gene-rich than any other homoeologous chromosomes found in other grass species. The effectiveness of this comparative map-based cloning method is confirmed if the colinearity of marker order (macrocolinearity) is conserved at the sequence level (microcolinearity). Although some types of rearrangement such as small inversions or gene duplications, are not expected to have a big effect on microcolinearity, such deletion and translocation events, even on a small scale, can greatly hamper map-based cloning using rice as a model species. For this reason, as Feuillet & Keller (2002) pointed out, “approaches based on colinearity between grass genomes must also be performed using more closely related species, e.g. within tribes or subtribes.”

On the basis of these considerations, we partially sequenced the BAC clone 346H10 to investigate the level of conservation of synteny with rice at the sequence level and to investigate the nature of the genes (if any) that were contained within this BAC. The clone 346H10, was shotgun-sequenced with 10X coverage. There were some sequence gaps that could not be closed because of repetitive DNA. However, the result was a total of 129,046 bp of sequence, representing around 99.3% of the total BAC length. The sequence data were organized in 20 contigs, whose lengths ranged from 1,463 to 23,166 bp. Using the “monocot” option of the FGENESH program, four of these contigs, comprising a total of 13,033 bp, failed to predict any peptide sequences. Thirteen peptides related to mobile elements were identified in 9 of the contigs. Among transposon-related elements, peptides belonging to the ping/pong/SNOOPY (2), En/Spm (2) and mariner (1) subclasses were identified, together with a single hAT domain that
was related to the HOBO/AC/TAM3 subclass. Among retrotransposons, peptides related to the ty3-gypsy (2), and ty1-copia (2) subclasses were identified, together with two other proteins related to unclassified retrotransposons and one general-type polyprotein. The sequences related to transposable elements may be responsible for the background signals observed during FISH analysis. Four genes, whose e-values were $e^{-34}$ or less, were identified as unrelated to transposable elements. The presence of two of these genes in the BAC under study, corresponded to previously annotated genes and were found to be conserved in the orthologous region of the rice genome. The resulting predicted protein of these two genes showed significant homology with a protein of the EXS family (PsEXS) and to a protein kinase domain (PsPKD). When the predicted translated sequences of these two genes were BLASTed against the TIGR database with the rice training set option, we identified the position of the putative orthologous rice genes that were both located on the BAC clone named OSJNBa0056D07. The genes were adjacent to the other BAC, OSJBBa0001B02, which harboured three ESTs (C454, C1069 and C60087) that were linked to apomixis in several *Paspalum* spp. (Pupilli & al., 2004). This indicates that the BAC clone under study belongs to the region linked to apomixis in several previously identified *Paspalum* spp. (Pupilli & al., 2004) and is likely to contain candidate genes for apomixis in this genus. To evaluate the level of sequence conservation between *Paspalum* and rice, the intron/exon structure of both PsEXS and PsPKD was compared between the two species. The comparison revealed major rearrangements of the *Paspalum* genes with respect to their rice counterparts: such rearrangements consisted mainly of transposable element insertions. All of these insertions interrupted the local colinearity between the *Paspalum* and rice homologues rendering them non-functional. As a consequence of these insertions, the structure of the *Paspalum* genes was fragmented, with each fragment exhibiting a high level of homology with the rice ortholog, but having apparently evolved as independent genes, as indicated by the presence of putative sites for initiation of transcription and polyadenylation. The overall sequence comparison between the two orthologs showed portions of strong homology between coding regions. These were interspersed with sequences in which the coding regions of the rice gene were significantly homologous with non-coding *Paspalum* regions, and other sequences in which no homology was detected. This may indicate that the two *Paspalum* genes evolved in such a way as to buffer the dramatic rearrangement induced by the insertion of the transposable elements. The loss of coding capacity in *Paspalum* sequences with respect to their homologs in rice was due mainly to small deletions (1–5 nucleotides) and point mutations that introduced new stop codons (Calderini & al., 2006).

To summarize, both large- and small-scale rearrangements were identified in the structure of two genes located in the ACL of *P. simplex* when compared with their rice homologs, which were taken as a reference for the sexual homolog of apomictic *Paspalum*: large-scale rearrangements included insertions of components of transposable elements. When these were inserted in the opposite orienta-
tion with respect to *P. simplex* gene, they induced the formation of multiple “sub-genes” that appeared to be transcribed independently from each other. Conversely when an element was inserted in the same direction, its coding region appeared to be transcribed as a continuum with the other exons of the *P. simplex* genes. The small-scale rearrangements included a 110-bp duplication, frequent small deletions and single point mutations that induce premature stop codons, caused a pseudogene-nization of many exons.

**CONCLUSIONS**

The significant synteny between the ACL of *P. simplex* and the telomere of RC12 was confirmed at the sequence level for only two genes within the BAC analysed. In our results, a relatively small number of genes appear to be conserved between rice and the ACL of *P. simplex*, while interruption of colinearity through transposition of genes from other locations appears to be the rule rather than the exception. The transposed genes were mainly related to transposable elements, which, in the case of the two genes analysed in more detail, interrupted local colinearity with rice and induced the gene to evolve a strategy to counterbalance such invasion, thereby potentially retaining at least some coding capacity. This consideration raises the question of whether the extensive gene movements and the frequent disruption of colinearity may limit the use of the rice genome sequence for positional cloning of apomixis genes. In addition, as Ozias-Akins & al. (2003) pointed out, there is no common region on the grass genome responsible for aposporous apomixis. This suggests a polyphyletic evolution of the locus within the grass family or a localised disruption of synteny caused by rearrangements. On the other hand, several independent studies identified loci responsible for spikelet sterility in the same area of RC12 that was in a position homologous to the ACL of *P. simplex* (Wu & al., 1995; Wan & al., 1996; Liu & al., 2001). More recently, Kubo & Yoshimura (2004) mapped the locus hsa1, which is responsible for hybrid breakdown in *indica/japonica* crosses, to the same region. In all of these cases, spikelet sterility was related to defective development of the embryo sac. These results, together with those reported here, seem to indicate the existence of a gene complex responsible for female gametophyte development in this area of RC12. It is conceivable that genes related to apomictic reproduction are contained within such a complex. However, due to the frequent transposition of genes from other locations, as highlighted by our BAC sequence analysis, the involvement of elements (other than those present in RC12) in the expression of the apomictic phenotype should not be ruled out. In any case, large scale sequencing of apomixis-linked BACs, while cumbersome, is perhaps the most promising strategy to identify key genes for apomictic reproduction in wild apomictic species.
ACKNOWLEDGEMENTS

This study was supported by the European Union as part of the project "Natural apomixis as a novel tool in plant breeding (ApoTool),” contract number QLG2-2000-00603 of the Quality of Life and Management of Living Resources section and by a contribution of Italian Ministry of Foreign Affairs, legge 401/90 Art. 20.

LITERATURE CITED


