

A Segment of the Apospory-Specific Genomic Region Is Highly Microsyntenic Not Only between the Apomicts *Pennisetum squamulatum* and Buffelgrass, But Also with a Rice Chromosome 11 Centromeric-Proximal Genomic Region^{1[W]}

Gustavo Gualtieri², Joann A. Conner, Daryl T. Morishige, L. David Moore, John E. Mullet, and Peggy Ozias-Akins*

Department of Horticulture, University of Georgia, Tifton, Georgia 31793-0748 (G.G., J.A.C., P.O.-A.); and Department of Biochemistry and Biophysics, Texas A&M University, College Station, Texas 77843 (D.T.M., L.D.M., J.E.M.)

Bacterial artificial chromosome (BAC) clones from apomicts *Pennisetum squamulatum* and buffelgrass (*Cenchrus ciliaris*), isolated with the apospory-specific genomic region (ASGR) marker *ugt197*, were assembled into contigs that were extended by chromosome walking. Gene-like sequences from contigs were identified by shotgun sequencing and BLAST searches, and used to isolate orthologous rice contigs. Additional gene-like sequences in the apomicts' contigs were identified by bioinformatics using fully sequenced BACs from orthologous rice contigs as templates, as well as by interspecies, whole-contig cross-hybridizations. Hierarchical contig orthology was rapidly assessed by constructing detailed long-range contig molecular maps showing the distribution of gene-like sequences and markers, and searching for microsyntenic patterns of sequence identity and spatial distribution within and across species contigs. We found microsynteny between *P. squamulatum* and buffelgrass contigs. Importantly, this approach also enabled us to isolate from within the rice (*Oryza sativa*) genome contig Rice A, which shows the highest microsynteny and is most orthologous to the *ugt197*-containing C1C buffelgrass contig. Contig Rice A belongs to the rice genome database contig 77 (according to the current September 12, 2003, rice fingerprint contig build) that maps proximal to the chromosome 11 centromere, a feature that interestingly correlates with the mapping of ASGR-linked BACs proximal to the centromere or centromere-like sequences. Thus, relatedness between these two orthologous contigs is supported both by their molecular microstructure and by their centromeric-proximal location. Our discoveries promote the use of a microsynteny-based positional-cloning approach using the rice genome as a template to aid in constructing the ASGR toward the isolation of genes underlying apospory.

Apomixis refers to a process of asexual reproduction through seeds that occurs in certain angiosperm plants, where the maternal genotype is cloned into genetically identical offspring that lack the input of a male genotype in their genetic makeup (Nogler, 1984; Koltunow, 1993). Different types of apomixis can be distinguished, although all of them are characterized by the development of embryos from maternal diploid cells that divide by mitosis, instead of from male-fertilized haploid maternal cells originated by meiosis

as in sexual reproduction. Apomixis is primarily divided into sporophytic and gametophytic types. Sporophytic apomixis, also known as adventitious embryony, consists of the development of embryos from embryo initials directly originated from one or more nongenerative ovule cells. In gametophytic apomixis the embryo develops from embryo sacs that originate by mitosis of a megaspore mother cell in a process called diplospory or by mitosis of a nearby nucellar cell in a process called apospory. In both processes one of the embryo sac (female gametophyte) nuclei becomes the egg cell that continues a parthenogenetic development by mitotic divisions in the absence of fertilization. In all apomicts the production of viable seeds requires endosperm formation that in most cases needs pollination or even fertilization, although in some cases endosperm may develop autonomously from an unfertilized embryo sac central cell. Both grass family (Poaceae) apomictic species studied here, *Pennisetum squamulatum* and buffelgrass (*Cenchrus ciliaris*), display an aposporous type of gametophytic apomixis (Ozias-Akins et al., 2003).

Apomixis has significant practical implications in plant breeding and hybrid production given that agriculturally high-performing genotypes are clonally

¹ This work was supported by the National Science Foundation (award no. 0115911).

² Present address: School of Integrative Biology, Faculty of Biological and Chemical Sciences, University of Queensland, Brisbane, QLD 4072, Australia.

*Corresponding author; e-mail pozias@uga.edu; fax 229-386-7371.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Peggy Ozias-Akins (pozias@uga.edu).

^[W] The online version of this article contains Web-only data.

Article, publication date, and citation information can be found at www.plantphysiol.org/cgi/doi/10.1104/pp.105.073809.

transferred into genetically identical offspring without the detrimental disruption of desired genetic combinations that sexual reproduction generates. Unfortunately, few crops are capable of apomictic reproduction, and, therefore, from a biotechnological perspective, the molecular dissection of the pathway(s) underlying apomixis is essential toward the future possibility of genetically engineering this trait into important crops. One possibility to aid the positional cloning of the gene(s) underlying apospory in the large and complex genome polyploid species *P. squamulatum* and buffelgrass is through comparative genomics and microsynteny analysis not only between the two apomicts but also with model cereals such as rice (*Oryza sativa*).

The terms microsynteny and microcolinearity generally are used to refer to the conserved content and order of genes and markers between compared genomic regions ranging up to the megabase-size level. Many resources suitable for comparative genomics have been generated, and several reports discuss microsynteny between species belonging to the same plant family, e.g. Leguminosae (Gualtieri and Bisseling, 2002; Gualtieri et al., 2002; VandenBosch and Stacey, 2003; Young et al., 2003; Choi et al., 2004; Zhu et al., 2005), Brassicaceae, Poaceae, and Solanaceae (Bennetzen, 2000; Devos and Gale, 2000; Paterson et al., 2000; Schmidt, 2000, 2002; Devos, 2005). Moreover, microsynteny between species belonging to different plant families was observed between tomato (*Lycopersicon esculentum*; Solanaceae) and Arabidopsis (*Arabidopsis thaliana*; Brassicaceae; Ku et al., 2001; Rossberg et al., 2001); between *Lotus japonicus* and pea (*Pisum sativum*; Leguminosae), and Arabidopsis (Stracke et al., 2004); and between *Glycine max* and *Medicago truncatula* (Leguminosae), and Arabidopsis (Yan et al., 2004). In animal systems, microsynteny and its application are also well documented (O'Brien et al., 1999). Several of these plant and animal studies report local disruptions of colinearity within microsyntenic regions due to small rearrangements, such as duplications, deletions, insertions, inversions, and translocations.

The microsynteny-based positional cloning of genes is particularly useful when target genes present in species with large and complex genomes are amenable to cloning by comparative genomics with model species for which a large number of research resources are available. In plants, microsynteny analysis with model species has led to positional cloning of genes governing important traits in only a few cases, despite the increasing number of comparative genomics studies. However, the number of genes cloned using this approach is likely to increase given the extensive conservation of genome structure shown in many recent studies. One example of the value of comparative genomics was the use of the model legume *M. truncatula* as a vehicle for the positional cloning of the pea *SYM2* locus (Gualtieri and Bisseling, 2002; Gualtieri et al., 2002), a putative Nod factor receptor. Microsynteny analysis resulted in the isolation and delimitation of the *M. truncatula* *SYM2*-orthologous genomic region (Gualtieri

et al., 2002) where candidate gene families were discovered (Gualtieri and Bisseling, 2002). Subsequently, functional analysis by the RNA interference technology of individual gene family members resulted in the identification of a Nod factor receptor kinase in *M. truncatula* (Limpens et al., 2003) that maps in a position syntenic to the pea *SYM2* locus. Even across plant families, microsynteny between tomato and Arabidopsis was useful to place additional markers around the tomato fruit-shaped *ovate* locus and to facilitate positional cloning (Ku et al., 2001). Irrespective of whether the target locus is itself microsyntenic or whether the model species used as a reference genome displays the studied phenotype, analysis of overall synteny enhances the molecular characterization and isolation of markers on the target region, a result that is particularly useful in species with large and complex genomes that complicate genetic analysis.

Previous molecular genetic research done on *P. squamulatum* and buffelgrass generated crucial information and molecular resources that were valuable in our comparative genomics studies at the apospory-specific genomic region (ASGR). In these studies a number of markers linked to the apospory phenotype in *P. squamulatum* (Ozias-Akins et al., 1998) and buffelgrass (Roche et al., 1999) were isolated that defined the ASGR. Subsequently, bacterial artificial chromosome (BAC) libraries from the two species were generated, and several BAC clones were isolated by screening with the ASGR-linked molecular markers (Roche et al., 2002). In this article we focus on a segment of the ASGR from the two apomictic species that contains the *ugt197* marker (Ozias-Akins et al., 1998). We assembled BACs containing the *ugt197* marker into contigs and extended them by chromosome walking. Gene-like sequences were identified from these contigs and used to isolate putative orthologous rice contigs. Microsynteny analysis was carried out by constructing detailed contig molecular maps showing the spatial distribution of markers and gene-like sequences. We show the existence of considerable microsynteny at the *ugt197* genomic region not only between *P. squamulatum* and buffelgrass, but also with the respective orthologous region of the rice genome. In addition, we describe an interestingly conserved localization of highly microsyntenic contigs proximal to centromeres or centromere-like sequences in the three compared species. The results presented here support the use of rice as a reference genome to assist with the construction and characterization of the ASGR toward the cloning of apospory-governing genes.

RESULTS AND DISCUSSION

Contig Assembly and Chromosome Walking

Marker *ugt197* was previously shown to be linked to the apospory phenotype both in *P. squamulatum* and in buffelgrass (Ozias-Akins et al., 1998; Roche et al.,

1999). Eight *P. squamulatum* and seven buffelgrass BAC clones were isolated after screening approximately $3.5 \times$ and $5 \times$ library coverage of the two genomes, respectively, by hybridization with *ugt197* (Roche et al., 2002). *P. squamulatum* and buffelgrass BAC clones were assembled into contigs by restriction and hybridization fingerprinting in a similar way as described previously (Gualtieri and Bisseling, 2002; Gualtieri et al., 2002). Contig assemblies were generated via a comparative analysis of BAC clone fingerprint band patterns that enabled the ordering of restriction fragments along contigs and the definition of contig regions representing BAC clone overlapping arrays within contigs (Fig. 1). Marker and gene-like sequences were mapped onto the assembled contigs by further restriction and hybridization analysis. Results were integrated and schematically represented by linear contig molecular maps showing the distribution of marker and gene-like sequences (Fig. 2). These contig maps provided the basis for chromosome walk-

ing and microsynteny comparisons between the two apomicts and also with the model cereal rice.

The *ugt197*-derived *P. squamulatum* and buffelgrass BACs were assembled into a single contig named P2, and two contigs named C1A and C1C, respectively (Fig. 1; Supplemental Table I). Contig C1C was extended from one end by chromosome walking, thereby gaining 40 kb and resulting in a final size of 195 kb (Fig. 1; Supplemental Table I). A third contig named C1B was assembled from three identical buffelgrass BAC clones that had been isolated using the same probe that extended the C1C contig (Fig. 1; Supplemental Table I). Contig C1B lacks marker *ugt197* and represents a duplicate chromosome segment (see below; Fig. 2).

Three and 15 BAC clones were isolated in chromosome walks attempted from the two ends of the *P. squamulatum* P2 contig, respectively (Supplemental Table II); however, although all of the BACs isolated clearly hybridized with the end probes, none had the

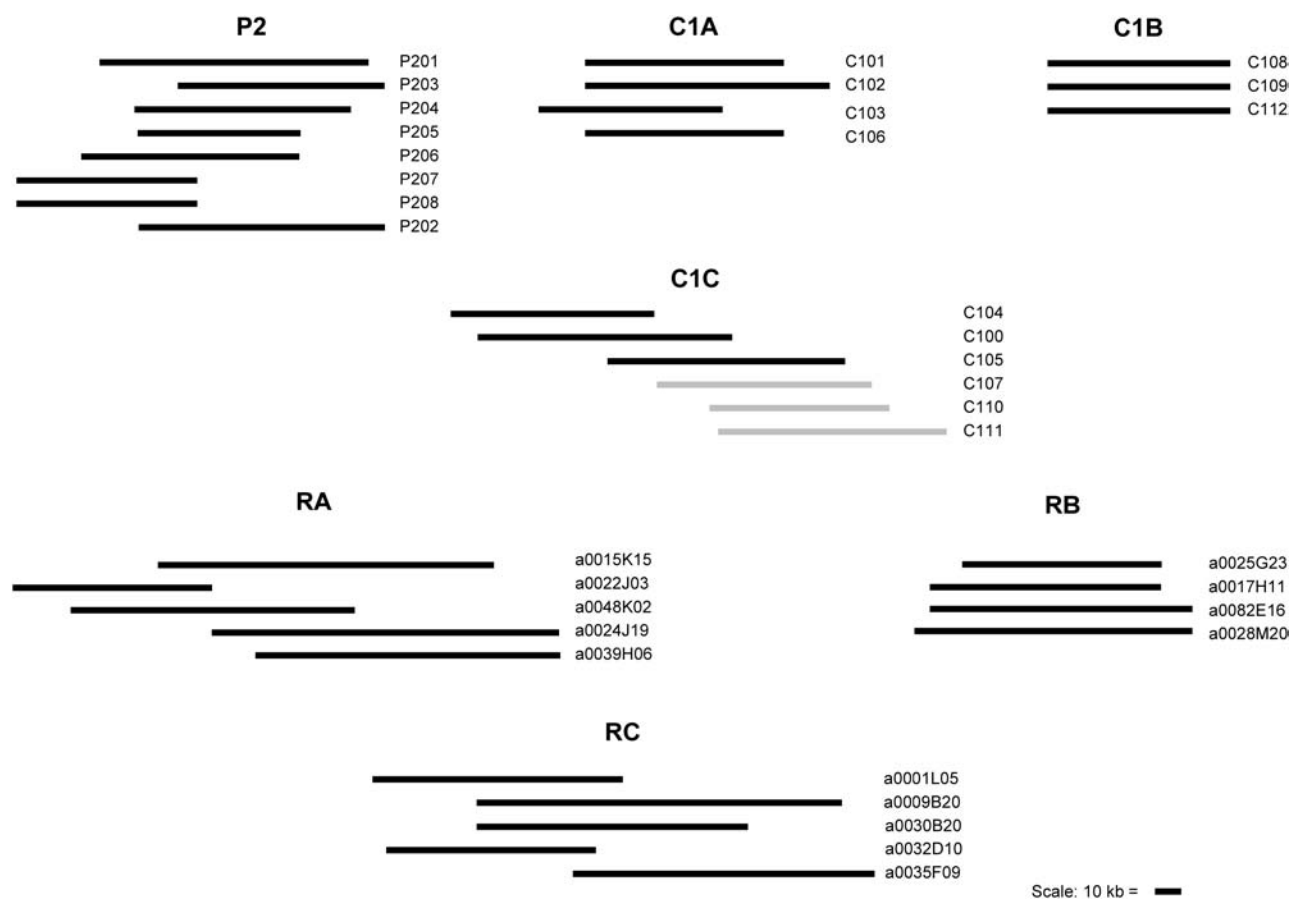


Figure 1. BAC composition and structure of contigs. The overlapping pattern of BACs in each contig assembly is shown for each species: P2, *P. squamulatum* contig; C1A, C1B, and C1C, buffelgrass contigs; RA, RB, and RC, rice contigs. Each BAC is represented by a black or gray line. The identity of BACs is given at the right of each contig. In the apomicts, black lines represent BACs initially isolated by screening the BAC libraries with the *ugt197* marker (Roche et al., 2002) and assembly into contigs. BACs depicted as gray lines in the buffelgrass C1C contig were isolated by chromosome walking and represent a contig extension. All rice BACs were isolated by screening the rice BAC library with the AP2 and XYL sequences. A 10-kb scale bar is provided.

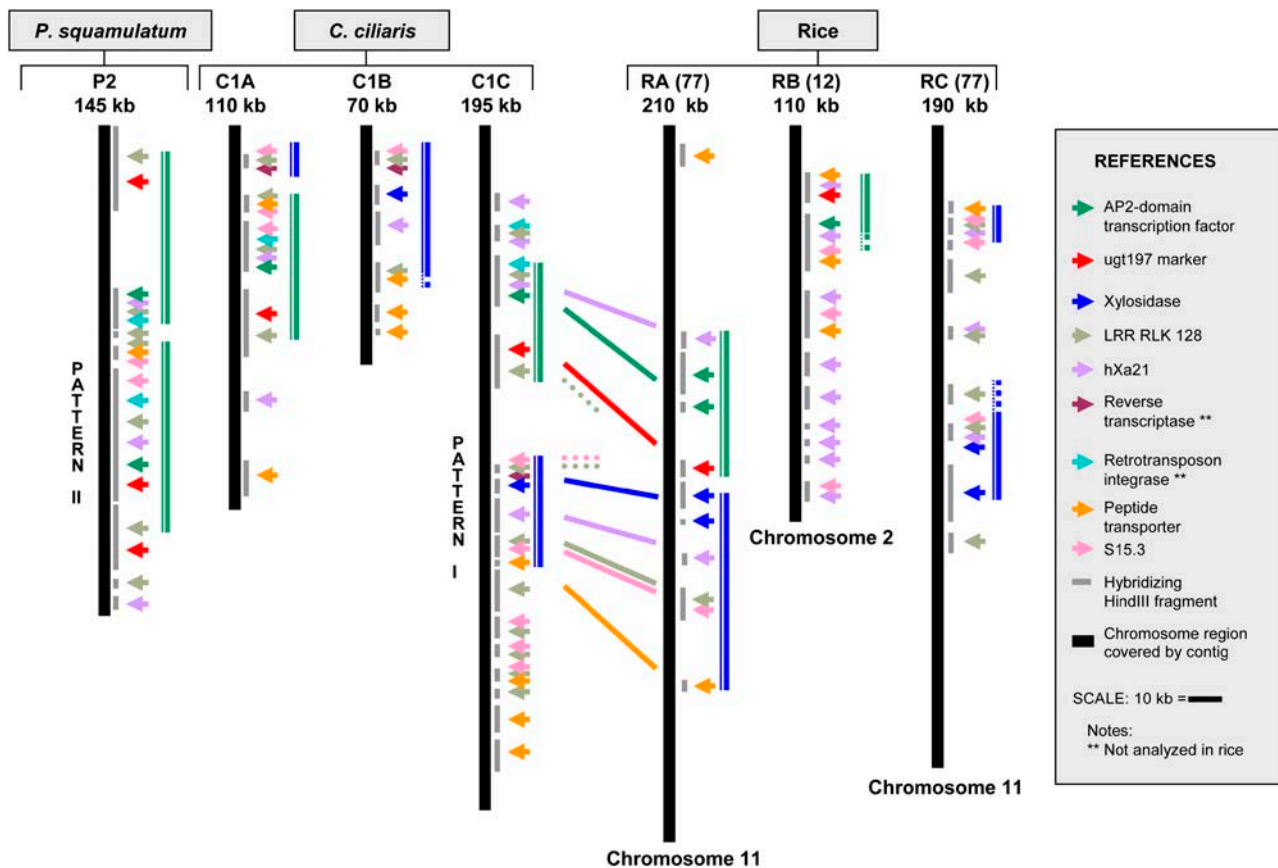


Figure 2. Microsynteny between the ASGRs of *P. squamulatum* and buffelgrass at the *ugt197*-containing contigs and their rice orthologous contigs. The chromosomal region spanned by contigs is represented by black thick vertical lines. Gray lines at the right side of contigs represent *Hind*III restriction fragments to which the analyzed gene-like sequences and markers represented by colored arrows hybridize. A kilobase scale, color codes, and notes are given in the picture reference frame. Microsynteny patterns I and II referred to in the text are quoted and depicted by double-blue and double-green fine vertical lines beside contig C1C and contig P2, respectively. When these patterns show disruption in one or more sequences, the region of disruption is indicated by dashed double lines. For the contigs C1C and RA that show the highest microsynteny between the apomicts and rice, oblique colored lines are drawn that connect each syntenic gene-like sequence and marker. When these lines are dashed, they indicate a synteny disruption. Rice contig names are followed by a number between parentheses corresponding to the rice genome database FPC contigs to which they belong, according to the September 12, 2003, build (www.genome.arizona.edu). The chromosome location of each rice contig is indicated below the drawings.

end-restriction pattern of the clone used to start the walk. This suggested that none of these BACs was linked to the contig assemblies or represented contig extensions, implying that the *P. squamulatum* BAC library has sequence gaps at the corresponding contig ends, which may be possible given the approximately $3.5 \times$ genome coverage of this library (Roche et al., 2002). Alternatively, some of these BACs could be linked in repulsion to the P2 contig, where the lack of a matching end restriction pattern could reflect polymorphisms that might be expected due to the polyploid and heterozygous genotype used for BAC library construction. However, this possibility seems less likely given that pairing and chiasmata formation between the ASGR carrier chromosome and its homolog have been infrequently observed in this primarily hemizygous region (Akiyama et al., 2004). In addition,

some of these BACs share extensive sequence similarity with the ASGR-linked contigs as shown by cross-hybridizations, indicating that they represent duplications of ASGR contig segments; and they may still map onto the ASGR given the large size of this region as recently deduced (Akiyama et al., 2004) combined with the fact that the cis-chromosome mapping of duplicated genes and chromosomal segments is well documented in other species, such as *Arabidopsis*, *M. truncatula*, and rice (*Arabidopsis* Genome Initiative, 2000; Gualtieri et al., 2002b; Shiu et al., 2004; Tian et al., 2004). Indeed, we present evidence that rice contigs Rice A (RA) and Rice C (RC) represent duplicated segments, both of them cis-located on rice fingerprint contig (FPC) contig 77 (see below).

The polyploid and polymorphic nature of the apomicts' genomes considerably complicates chromosome

walking assessment as compared to diploid species. Mapping of expressed sequence tag polymorphisms from homoeologous loci in hexaploid wheat was accomplished with nullisomic/tetrasomic lines and extensive expressed sequence tag sequence data (Mochida et al., 2004), genetic resources that are not developed for our species of interest. Fluorescence in situ hybridization (FISH) mapping could then be an alternative to address complex chromosome walking situations.

Gene Identification and Isolation of Rice Orthologous Contigs

DNA sequences encoding functions are likely to be conserved between species and are the ideal tool in comparative structural genomics (i.e. macro- and microsynteny; Bennetzen and Ma, 2003). Gene-like sequences were isolated from the apomicts or rice by three different approaches. Five gene-like sequences from shotgun clones of the *ugt197*-containing contigs C1C and P2 showed similarity to known or predicted genes after BLASTX analysis (Table I). These sequences were amplified by PCR (Supplemental Table III) and used for microsynteny analysis by hybridization. Two genes identified corresponded to a retrotransposon integrase and a reverse transcriptase and were used only in our comparative genomic analysis with the apomict contigs (Table I). Amplicons of two sequences, namely, an AP2-domain transcription factor (AP2) and a xylosidase (XYL) homolog, were used to screen a rice BAC library. The isolated rice BACs a0015K15 and a0048K02 cohybridized with both AP2 and XYL, and this colocalization was a preliminary indication of microsynteny. Additional BACs were isolated, five containing only AP2 (BACs a0017H11, a0025G23, a0028M20, a0082E16, and a0022J03) and seven containing only XYL (BACs a0001L05, a0009B20, a0030B20, a0032D10, a0035F09, a0024J19, and a0039H06). These BACs were assembled into the three contigs RA (210 kb), Rice B (RB; 110 kb), and RC (190 kb; Fig 1; Supplemental Table I). An anchor point coordinate for each contig, RA and

RC on chromosome 11 pseudomolecule and RB on chromosome 2 pseudomolecule, are given in Supplemental Table I as per The Institute for Genomic Research (www.tigr.org) rice database. Throughout this article we correlated our rice contigs to the most recent rice genome database FPC contig update and chromosome maps (September 12, 2003, build; www.genome.arizona.edu). RA and RC are both part of the larger rice genome database FPC contig 77 that maps adjacent to the centromere on rice chromosome 11 (www.genome.arizona.edu). FPC contig 77 spans about 6.3 cM (from 49.6–55.9 cM) corresponding to approximately 2.77 Mb of chromosome 11. Although the RA and RC contigs do not overlap on contig 77, they are separated by a relatively small sequence of about 100 kb (www.genome.arizona.edu; Fig. 3). Rice contig B is part of the larger FPC contig 12 that maps on rice chromosome 2 (www.genome.arizona.edu). Amplicons from the *ugt197* marker and the peptide transporter gene-like sequence (Supplemental Table III) also were used for microsynteny analysis by hybridization to *P. squamulatum*, buffelgrass, and rice contigs as shown in Table I and Figure 2.

The second gene identification approach relied on the use of publicly available sequence data from the above putatively orthologous rice contigs as templates to aid the identification of additional genes in the apomicts. We found on the rice genome database contig 12 (www.genome.arizona.edu) that the sequenced BAC OJ2055_H10 (AP005300) and our contig RB end-BAC a0017H11 overlapped (Supplemental Table I). BLASTX analysis of 10-kb segments of the AP005300 sequence resulted in identification of a Xa21 disease resistance RLK homolog named hXa21 (now annotated as feature OJ2055_H10.35 in the BAC sequence AP005300). The hXa21 sequence was PCR amplified (Supplemental Table III) and used for intra- and interspecies microsynteny analysis (Table I). In contrast, no sequenced BACs were identified at the time that overlapped with BACs from rice contigs A and C, and another approach for gene identification in these BACs was needed.

Table I. Distribution of gene-like sequences and markers among the apomicts and rice orthologous contigs

Gene-Like/Marker Sequences	Present in Contig						
	C1A	C1B	C1C ^a	P2	RA ^a	RB	RC
<i>ugt197</i> ^b	+	– ^d	+	+	+	–	–
AP2-domain transcription factor ^e	+	–	+	+	+	+	–
Xylosidase ^e	–	+	+	–	+	–	+
Peptide transporter ^e	+	+	+	+	+	–	–
Retrotransposon integrase ^e	+	–	+	+	n.a. ^f	n.a.	n.a.
Reverse transcriptase ^e	+	+	+	–	n.a.	n.a.	n.a.
hXa21 ^g	+	+	+	+	+	+	+
LRR RLK 128 ^g	+	+	+	+	+	+	+
S15.3 ^h	+	+	+	+	+	+	+

^aThe two most highly microsyntenic contigs. ^bSequence-characterized amplified region (Ozias-Akins et al., 1998). ^c+, Presence of the sequence in a given contig. ^d–, Absence of the sequence in a given contig. ^eIsolated by shotgun sequencing and BLASTX queries from ASGR-linked BACs. ^fNot analyzed. ^gIsolated by BLASTX queries with orthologous rice BAC sequences. ^hIdentified by interspecies hybridization.

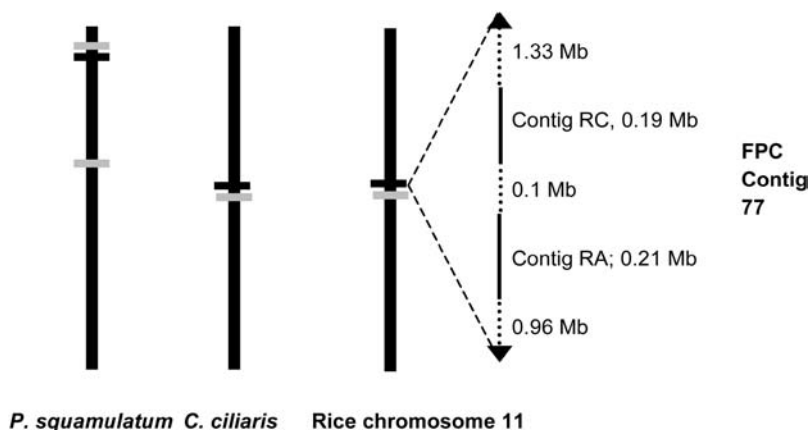


Figure 3. Comparative chromosomal location relative to centromeric sequences of the ASGRs from the two apomicts and their rice chromosome 11 orthologous genomic regions. Thick black vertical lines represent the relevant chromosomes in each species. Gray horizontal lines indicate chromosome centromeres and a centromere-like sequence on the distal side of the short arm of *P. squamulatum* chromosome, while black horizontal lines indicate the syntenic *ugt197*-containing ASGR segment in the apomicts and their orthologous rice genomic regions within FPC contig 77. The approximately 2.77-Mb FPC contig 77 is zoomed in next to the chromosome 11 scheme to show the relative orientation of contigs RA and RC (depicted by thin black vertical lines), and also to indicate in megabase (Mb) units the approximate sequence spanning (shown by dashed lines) in between these two contigs, and spanning from RA and RC to the centromere and to the distal end of FPC contig 77, respectively. The chromosome schemes for *P. squamulatum* and buffelgrass are representing the recently described localization of the ASGRs relative to centromeric sequences (Goel et al., 2003).

The third gene identification approach relied on the general conservation over time of function-encoding sequences and their cross-hybridization between species. A DNA pool of BACs showing a minimum tiling path on the *P. squamulatum* contigs P2 and on the buffelgrass contigs C1A, C1B, and C1C was labeled and hybridized to *Hind*III blots containing the rice contigs RA, RB, and RC. Strong cross-hybridization was observed on a single *Hind*III fragment (9.7 kb) of contig RA and was named 15.3. Several weakly cross-hybridizing fragments were registered on the other two rice contigs, RB and RC. All these fragments also hybridized either with one or both of the RLK probes hXa21 or the rice Leu-rich repeat (LRR)-containing RLK probe LRR-RLK128 (Fig. 2) that was identified by BLASTX to a region of BAC sequence (nucleotides 97,078–98,346 from AC098598). Therefore, we grouped these fragments in the family of related sequences named S15.3 (Table I). Despite us showing here that some of these DNA fragments contain additional gene(-like) sequences (see below and Fig. 2) and their further subcloning and sequencing might identify more genes, hybridization has already detected the presence in all of them of at least part of (LRR-) RLK-related gene sequences.

Contig Sequence Composition, Structure, and Microsynteny Analysis

A comparative contig structure analysis clearly reveals microsynteny patterns involving the conservation of both sequence composition and order in all analyzed contigs from the apomictic species and rice.

Two ancestral microsynteny patterns named I and II can be distinguished (blue and green double-vertical lines beside contigs P2 and C1C on Fig. 2, respectively). These two patterns show some divergence in sequence composition and order, or one of them may be absent, between compared contigs. These microsynteny patterns likely evolved via structural rearrangements (duplications, deletions, inversions, translocations) and nucleotide sequence divergence, as has been observed in other grass genomes (Bennetzen and Ramakrishna, 2002), and they pinpoint not only intracontig but also intra- and interspecies contig structure relatedness. For example, within a single contig, the microsynteny pattern II in P2 appears to have been duplicated and inverted, with the deletion of some distal sequences (Fig. 2).

Within a single species, the rice contig RC contains two pattern I-like segments separated by about 50 kb, and these two patterns are related to pattern I-like segment of contig RA (Fig. 2). Moreover, contigs RA and RB contain pattern II-like segments. The relatedness of these molecular patterns coupled with the collocation of contigs RA and RC on rice genome database contig 77 (www.genome.arizona.edu) suggests the occurrence of intrachromosome segmental cis-duplications close to the centromere of chromosome 11 during rice genome evolution (Fig. 3). In contrast, microsynteny pattern II similarity would indicate that rice contig RA from chromosome 11 is related in origin with a segment of contig RB from chromosome 2 through duplication and interchromosomal translocation processes. Analysis of microsynteny patterns further suggests that the buffelgrass contigs C1A, C1B,

and C1C (Fig. 2) could also originate via segmental duplication accompanied by rearrangements, and/or be the result of polyploid homologous/homoeologous chromosome regional allelisms/polymorphisms.

Among species, the presence of a complete pattern II both in P2 and C1A shows that these contigs are highly related. Buffelgrass C1A and C1C, and rice RA, contain some extent of both patterns I and II, indicating that they are most related from among the total of contigs of the three compared species. From these comparisons the highest hierarchical orthology between contigs in terms of microsynteny is shown between the buffelgrass contig C1C and the rice contig RA that share colinearity of seven gene-like sequences and marker *ugt197* over a chromosome segment spanning about 80 and 100 kb in each contig, respectively. Within this microsyntenic region an intervening colinearity disruption is given by a 20-kb segment of contig C1C involving the distal ends of patterns I and II that comprise homologs to the sequences RLK128 and S15.3, and RLK128, respectively (Fig. 2, colored dashed lines). This disruption would account for sequence duplications or insertions in C1C, or deletions in RA during the course of genome evolution. Despite this minor local microsynteny disruption, it can be concluded that from within the buffelgrass and the rice genomes the most orthologous genomic regions containing the ASGR-linked marker *ugt197* are the C1C and RA contigs, respectively. It is not discarded that further extension of P2 could reach a genomic region containing a pattern I-like segment; however, within the sequence stretch covered in this study, the P2 and rice orthologous contigs show relatedness only through pattern II-like segments.

Most importantly, it can be concluded that the rice contig RA, included in FPC contig 77 and located proximal to the centromere on chromosome 11, is the rice genomic region most orthologous to the ASGR segment containing the marker *ugt197*. This finding is highly significant for the use of microsynteny between the apomicts and rice as an approach to aid the positional cloning of the gene(s) underlying the apomixis trait.

Chromosome Localization of Orthologous Contigs

Evolutionary relationships between orthologous contigs of the three species are evident not only based on conserved microsynteny patterns, but also based on the similar chromosome location of contigs. FISH mapping showed that ASGR-linked contigs from buffelgrass were located close to the centromere (Goel et al., 2003), while the *P. squamulatum* ASGR-linked BACs localized to the distal end of the short arm of a single chromosome (Goel et al., 2003; Akiyama et al., 2004), where what appears as a secondary centromere-like sequence also was demonstrated (Goel et al., 2003).

Remarkably, both the buffelgrass ASGR and the rice orthologous contigs RA and RC (chromosome 11, FPC contig 77) share a centromere-proximal location, while

the presence of centromere-like sequences nearby the *P. squamulatum* ASGR may reflect its evolutionary origin from a related ancestral region of centromeric-proximal location, possibly through a translocation carrying a stretch of centromeric-like sequence to the new location (Fig. 3). Interestingly, rice and buffelgrass are most related both by the highest of all contig colinearity between rice contig RA and buffelgrass contig C1C, and also by the centromeric-proximal location of these two contigs. Thus, conserved chromosome location, as well as microsyntenic molecular patterns, could also account for a common evolutionary history of these contigs, reflecting ancient patterns of genome rearrangements involving the ASGR and its flanking centromeric regions.

Conclusions and Perspectives

Our finding of microsynteny between the *ugt197*-containing ASGR segment and the model cereal rice set an important precedent not only to continue constructing and dissecting the ASGR around this segment, but also to apply similar microsynteny studies to other ASGR segments. The highly advanced and continuing development of rice genome sequence and map databases represents a highly valuable tool to analyze the ASGR using rice as a reference species. A rice genome sequence based on ordered large-insert clones that covers 95% of the rice genome and has been assembled into pseudomolecules corresponding to the 12 rice chromosomes has now been published (International Rice Genome Sequencing Project, 2005). Use of the assembled and anchored rice sequence, in combination with microsynteny analysis and FISH mapping on the aposporous grasses, can significantly facilitate the chromosome walking extension of ASGR contigs, the generation of new landings within noncontiguous ASGR segments, and the jumping over BAC library gaps that may eventually occur at certain contig ends in these aposporous species. The rice sequence is expected to enhance microsynteny analysis and the isolation of new markers and genes from the ASGR. Assuming that one possible mechanism governing the apospory phenotype could be the alteration of the function of a gene that normally acts during sexual reproduction (Koltunow and Grossniklaus, 2003), i.e. through mutation or by a regulatory shift, the use of existing rice sequence data spanning regions of microsynteny with the aposporous grasses can well be a source of apospory candidate genes. Moreover, rice sequence data spanning these microsyntenic regions are an excellent resource for marker development throughout the ASGR in order to further delimit the region through genetic mapping or by genetic analysis of gamma radiation-deletion panel lines.

The availability of BAC libraries for both apomictic species, the capability of FISH mapping of BAC clones on both apomicts, the rapid development of the rice genome database, and microsynteny analysis between rice and the apomictic species are all key factors that,

combined, will facilitate the positional cloning of the gene(s) underlying apospory.

MATERIALS AND METHODS

Construction of Contig Maps and Chromosome Walking

Pennisetum squamulatum and buffelgrass (*Cenchrus ciliaris*) BAC clones hybridizing with the ASGR-linked marker ugt197 (Roche et al., 2002) were assembled into detailed contig molecular maps by complete *Hind*III-restriction and hybridization fingerprinting as described previously (Gualtieri and Bisseling, 2002; Gualtieri et al., 2002). Specific sequences were mapped onto the contigs by hybridization to fingerprinting blots (Gualtieri and Bisseling, 2002; Gualtieri et al., 2002).

Contig-end restriction fragments used as probes for chromosome walking were isolated by plasmid rescue of *Nsi*I- and *Cla*I-restricted BAC DNA as described previously (Gualtieri et al., 2002). Alternatively, amplicons of ASGR-specific sequence-characterized amplified regions generated from sequence analysis of *Bam*HI plasmid-rescued BAC ends were used as chromosome walking probes. High-density filters of each BAC library (Roche et al., 2002) were hybridized with contig-end probes following the protocol from the Clemson University Genomics Institute (<http://www.genome.clemson.edu/>). Positive clones were fingerprinted as described above to address whether they represented contig walking extensions.

BAC Subcloning and Sequencing

Starter cultures (5 mL Luria-Bertani [LB] media; 20 µg/mL chloramphenicol) were inoculated with single BAC colonies and incubated overnight at 37°C on a rotary shaker. Large-scale cultures (200 mL LB media; 20 µg/mL chloramphenicol) were inoculated with 200 µL of the starter culture and incubated with shaking for 14 h at 37°C. BAC DNA was purified using Tip100 columns (Qiagen), following the Low Copy Number Plasmid Isolation protocol provided by the manufacturer. Purified BAC DNA was treated overnight at 37°C with Plasmid-Safe DNase (Epicentre Technologies) according to the manufacturer's instructions to remove residual *Escherichia coli* genomic DNA.

BAC DNA (10 µg) was randomly sheared by passage through a GeneMachines Hydroshear DNA shearing device (Genomic Instrumentation Services) at a speed code setting of 9 to produce fragments of 1.5 to 3.0 kb. DNA fragments were blunt-end repaired using a combination of T4 DNA polymerase, Klenow DNA polymerase, and T4 polynucleotide kinase. The fragments were further size fractionated using a 1% (w/v) low-gelling point agarose gel (FMC Bioproducts). Specific size fractions were excised from the gel. DNA was purified using the Qiagen Gel extraction kit. Purified fragments were ligated into the dephosphorylated *Eco*RV site of pBluescript II (SK⁻) (Stratagene). Ligation reactions were used for transformation of XL10-gold competent cells (Stratagene).

Random transformants were transferred to LB media (1.4 mL LB media; 100 µg/mL ampicillin) in deep-well plates and incubated with shaking for 18 h at 37°C. Plasmid DNA was isolated using the Wizard MagneSil plasmid DNA purification system (Promega) with a Biomek 2000 workstation (Beckman).

Plasmids were sequenced using SK (5'-CGCTCTAGAACTAGTGG-ATC-3') or KS (5'-CTCGAGGTGCGACGGTATCG-3') primers. Sequencing was carried out using the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction kit (Applied Biosystems). Standard sequencing reactions included 1.5 µL plasmid DNA (typically 150–300 ng), 1 µL 10 µM primer, 1 µL BigDye reaction mix, 1 µL 5 × reaction buffer (5 × = 400 mM Tris, pH 9.0, 10 mM MgCl₂), and 5.5 µL distilled water. Sequencing reactions were carried out using an ABI 9700 thermocycler with the following cycling parameters: initial denaturation at 95°C for 1 min, followed by 99 cycles of 95°C, 10 s; 50°C, 5 s; 60°C, 4 min. Extension products were purified by isopropanol precipitation. Extension products were separated on an ABI PRISM 3100 genetic analyzer (Applied Biosystems).

Isolation of Orthologous Rice Contigs

Gene-like sequences were identified in the ugt197-containing contigs by shotgun sequencing and BLASTX queries of the nonredundant database at GenBank. Sequences with a significant match (E value < e-06) were PCR amplified from their respective BACs (Supplemental Table III). Rice (*Oryza*

sativa) BAC library filters (OSJNBa) obtained from the Clemson University Genomics Institute were probed with these amplified *P. squamulatum* or buffelgrass sequences, and orthologous BAC clones were isolated. When orthologous rice clones had been sequenced but not fully annotated, gene-like sequences (in addition to the hybridizing probes) were identified by BLASTX. These gene-like sequences were amplified by PCR (Supplemental Table III) and used to study microsynteny with the buffelgrass and *P. squamulatum* BAC contigs by hybridization to contig map fingerprint blots. Alternatively, when orthologous rice clones had not been sequenced, conserved sequences (i.e. genes) were isolated by whole-contig interspecies cross-hybridizations using BAC inserts that represented a minimum tiling path across rice and orthologous contigs from the apomicts. Orthologous rice contigs identified in this way were further assigned to rice genome database contigs and mapped onto rice chromosomes using the Arizona Genomics Institute bioinformatics resources (www.genome.arizona.edu).

ACKNOWLEDGMENTS

We thank Anne Bell and Evelyn Perry for technical support, and Evelyn Folds and Kathy Mullinix for logistic assistance.

Received November 4, 2005; revised December 21, 2005; accepted January 4, 2006; published January 13, 2006.

LITERATURE CITED

- Akiyama Y, Conner JA, Goel S, Morishige DT, Mullet JE, Hanna WW, Ozias-Akins P (2004) High-resolution physical mapping in *Pennisetum squamulatum* reveals extensive chromosomal heteromorphism of the genomic region associated with apomixis. *Plant Physiol* **134**: 1733–1741
- Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis*. *Nature* **408**: 796–815
- Bennetzen JL (2000) Comparative sequence analysis of plant nuclear genomes: microcolinearity and its many exceptions. *Plant Cell* **12**: 1021–1029
- Bennetzen JL, Ma J (2003) The genetic colinearity of rice and other cereals on the basis of genomic sequence analysis. *Curr Opin Plant Biol* **6**: 128–133
- Bennetzen JL, Ramakrishna W (2002) Numerous small rearrangements of gene content, order and orientation differentiate grass genomes. *Plant Mol Biol* **48**: 821–827
- Choi HK, Mun JH, Kim DJ, Zhu H, Baek JM, Mudge J, Roe B, Ellis N, Doyle J, Kiss GB, et al (2004) Estimating genome conservation between crop and model legume species. *Proc Natl Acad Sci USA* **101**: 15289–15294
- Devos KM (2005) Updating the 'crop circle'. *Curr Opin Plant Biol* **8**: 155–162
- Devos KM, Gale MD (2000) Genome relationships: the grass model in current research. *Plant Cell* **12**: 637–646
- Goel S, Chen Z, Conner JA, Akiyama Y, Hanna WW, Ozias-Akins P (2003) Physical evidence that a single hemizygous chromosomal region is sufficient to confer aposporous embryo sac formation in *Pennisetum squamulatum* and *Cenchrus ciliaris*. *Genetics* **163**: 1069–1082
- Gualtieri G, Bisseling T (2002) Microsynteny between *Medicago truncatula* SYM2-orthologous genomic region and another region located on the same chromosome arm. *Theor Appl Genet* **105**: 771–779
- Gualtieri G, Kulikova O, Limpens E, Kim D-J, Cook DR, Bisseling T, Geurts R (2002) Microsynteny between pea and *Medicago truncatula* in the SYM2 region. *Plant Mol Biol* **50**: 225–235
- International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. *Nature* **436**: 793–800
- Koltunow AM (1993) Apomixis: embryo sacs and embryos formed without meiosis or fertilization in ovules. *Plant Cell* **5**: 1425–1437
- Koltunow AM, Grossniklaus U (2003) Apomixis: a developmental perspective. *Annu Rev Plant Biol* **54**: 547–574
- Ku H-M, Liu J, Doganlar S, Tanksley S (2001) Exploitation of *Arabidopsis*-tomato synteny to construct a high-resolution map of the ovate-containing region in tomato chromosome 2. *Genome* **44**: 470–475
- Limpens E, Franken C, Smit P, Willemsse J, Bisseling T, Geurts R (2003) LysM domain receptor kinases regulating rhizobial Nod factor-induced infection. *Science* **302**: 630–633

- Mochida K, Yamazaki Y, Ogihara Y** (2004) Discrimination of homoeologous gene expression in hexaploid wheat by SNP analysis of contigs grouped from a large number of expressed sequence tags. *Mol Genet Genomics* **270**: 371–377
- Nogler GA** (1984) Gametophytic apomixis. In BM Johri, ed, *Embryology of Angiosperms*. Springer, Berlin, pp 475–518
- O'Brien SJ, Menotti-Raymond M, Murphy WJ, Nash WG, Wienberg J, Stanyon R, Copeland NG, Jenkins NA, Womack JE, Marshall Graves JA** (1999) The promise of comparative genomics in mammals. *Science* **286**: 458–481
- Ozias-Akins P, Akiyama Y, Hanna WW** (2003) Molecular characterization of the genomic region linked with apomixis in *Pennisetum/Cenchrus*. *Funct Integr Genomics* **3**: 94–104
- Ozias-Akins P, Roche D, Hanna WW** (1998) Tight clustering and hemizygosity of apomixis-linked molecular markers in *Pennisetum squamulatum* implies genetic control of apospory by a divergent locus which may have no allelic form in sexual genotypes. *Proc Natl Acad Sci USA* **95**: 5127–5132
- Paterson AH, Bowers JE, Burow MD, Draye X, Elsik CG, Jiang CX, Katsar CS, Lan TH, Lin YR, Ming R, et al** (2000) Comparative genomics of plant chromosomes. *Plant Cell* **12**: 1523–1540
- Roche D, Cong P, Chen ZB, Hanna WW, Gustine DL, Sherwood RT, Ozias-Akins P** (1999) An apospory-specific genomic region is conserved between buffelgrass (*Cenchrus ciliaris* L.) and *Pennisetum squamulatum* Fresen. *Plant J* **19**: 203–208
- Roche DR, Conner JA, Budiman MA, Frisch D, Wing R, Hanna WW, Ozias-Akins P** (2002) Construction of BAC libraries from two apomictic grasses to study the microcolinearity of their apospory-specific genomic regions. *Theor Appl Genet* **104**: 804–812
- Rosberg M, Theres K, Acarkan A, Herrero R, Schmitt T, Schumacher K, Schmitz G, Schmidt R** (2001) Comparative sequence analysis reveals extensive microcolinearity in the lateral suppressor regions of the tomato, Arabidopsis, and Capsella genomes. *Plant Cell* **13**: 979–988
- Schmidt R** (2000) Synteny: recent advances and future prospects. *Curr Opin Plant Biol* **3**: 97–102
- Schmidt R** (2002) Plant genome evolution: lessons from comparative genomics at the DNA level. *Plant Mol Biol* **48**: 21–37
- Shiu S, Mayer K, Li W, Tzeng Y, Karlowski W, Pan R** (2004) Comparative analysis of the receptor-like kinase family in *Arabidopsis* and rice. *Plant Cell* **16**: 1220–1234
- Stracke S, Sato S, Sandal N, Koyama M, Kaneko T, Tabata T, Parniske P** (2004) Exploitation of colinear relationships between the genomes of *Lotus japonicus*, *Pisum sativum*, and *Arabidopsis thaliana*, for positional cloning of a legume symbiosis gene. *Theor Appl Genet* **108**: 442–449
- Tian C, Li J, Chen M, Wan P, Sun S** (2004) Genome-wide analysis of the GRAS gene family in rice and *Arabidopsis*. *Plant Mol Biol* **54**: 519–532
- VandenBosch KA, Stacey G** (2003) Summaries of legume genomics projects from around the globe: community resources for crops and models. *Plant Physiol* **131**: 840–865
- Yan H, Mudge J, Kim D-J, Shoemaker R, Cook D, Young N** (2004) Comparative physical mapping reveals features of microsynteny between *Glycine max*, *Medicago truncatula*, and *Arabidopsis thaliana*. *Genome* **47**: 141–155
- Young ND, Mudge J, Ellis TN** (2003) Legume genomes: more than peas in a pod. *Curr Opin Plant Biol* **6**: 199–204
- Zhu H, Choi HK, Cook DR, Shoemaker RC** (2005) Bridging model and crop legumes through comparative genomics. *Plant Physiol* **137**: 1189–1196