Theor Appl Genet (1996) 92:702–708

X.-B. Li · C.-Z. Liang · H.-G. Wu · W.-X. Zhai
N. Huang · L.-H. Zhu

Isolation and identification of a non-specific tandemly repeated DNA sequence in *Oryza* species

Received: 15 May 1995 / Accepted: 8 September 1995

Abstract A tandemly repeated DNA sequence (RRS7) was isolated from *Oryza alta* (CCDD). RRS7-related sequences were also found tandemly arrayed in genomes AA, BB, BBCC, CC, and EE, and a small amount of RRS7-related sequences were detected in genome FF and the *Oryza* species with unknown genomes. DNA sequence analysis of the 1844-bp insert of RRS7 revealed that it contained six tandemly repeated units, of which five were 155 bp in length and one was 194 bp in length and contained an imperfect internal 39-bp duplication. Southern blot analysis showed that the boundary sequence contained in RRS7 is a single-copy sequence. A 155-bp consensus sequence derived from the six monomeric repeats contained no internal repeat and showed no significant homology to other currently known sequences. The results of Southern blot and sequence analysis revealed that there are at least two subfamilies present in the RRS7 family; these are represented by the *Dra*I site and the *Msp*I site, respectively. Restriction digestion with two pairs of isoschizomers *Mbo*I/*Sau*3A and *Msp*I/*Hpa*II demonstrated that most of the C residues in the GATC sites and the internal C in the CCGG sites of the RRS7 family in *O. alta* were methylated. The usefulness of the RRS7 family in determining the evolutionary relationship of the genome DD and other *Oryza* genomes is discussed.

Key words *Oryza* · Tandemly repeated DNA sequence · Restriction fragment length polymorphism (RFLP)

Communicated by M. Koornneef

X.-B. Li · C.-Z. Liang · H.-G. Wu[1] · W.-X. Zhai · L.-H. Zhu (✉)
Institute of Genetics, Chinese Academy of Sciences, Beijing 100101, China

C.-Z. Liang · N. Huang
International Rice Research Institute, P.O. Box 933, 1099 Manila, Philippines

Present address:
[1] 119 N-2 Hall, P.O. Box 110420, University of Florida, Gainsville, FL 32611, USA

Introduction

The genomes of eukaryotes are characterized by the presence of a large amount of repetitive DNA sequences that may be arrayed tandemly or dispersed throughout the genome (Singer 1982). Satellite DNAs, which consist of tandemly repeated sequences, are generally clustered at heterochromatic regions such as heterochromatic centromeres and telomeres (Brutlag 1980; Flavell 1980; Miklos 1985). While the size of the repeated unit remains conserved, both the copy number and the sequence of satellite DNAs have been determined to change rapidly during evolution (Brutlag 1980; Evans et al. 1983; Hemleben et al. 1992). It is therefore possible to investigate relationships between different species on the basis of tandemly repeated sequences (Xin and Appels 1988; Wu and Wu 1992; Ingham et al. 1993).

The genus *Oryza* consists of two cultigens and about 20 wild species. Morphological and cytological studies have indicated that *Oryza* species possess genomes such as AA, BB, CC, BBCC, CCDD, EE, and FF, while the genome constitution of some wild species is unknown (Vaughan 1994). Reassociation kinetic analysis was used by Deshpande and Ranjekar (1980) to determine that approximately 50% of the rice genome is composed of repetitive DNA sequences. Repetitive DNA sequences have been isolated and characterized from several species (Aswidinnoor et al. 1991; Zhao et al. 1989; Wu and Wu 1992). Some of these sequences are genomes AA-, CC-, EE-, or FF-specific, whereas others are shared by several genomes. However, none of these sequences have been isolated from the CCDD genome. The origin of the DD genome has yet to be determined because no diploid species with the DD genome has been found (Vaughan 1994). An analysis of the repeated DNA sequences of genome CCDD is thus important not only for understanding its structure and function, but it may also provide some clues to the origin of the DD genome as well as to its relationship with other genomes of *Oryza* species.

We report here the isolation of a tandemly repeated DNA sequence from *O. alta* (genome CCDD) and its characterization by Southern blot and sequence analyses. It is a non-specific DNA sequence that is present in all the 20 *Oryza* species examined in this study.

Materials and methods

Plant materials and DNA isolation

The *Oryza* species used in this study are listed in Table 1. The leaf samples of all of the *Oryza* species listed here, except for IR36 (*O. sativa*), were obtained from the PBGB division and IRGC, International Rice Research Institute (IRRI), Los Banos, Philippines. The accession of *O. alta* used to construct the genomic library (not shown) and IR36 were grown in the greenhouse at the Institute of Genetics, Academia Sinica, Beijing, China. Total DNA was isolated from leaves according to McCouch et al. (1988).

Library construction and repeated sequence selection

O. alta genomic DNA was partially digested with *Sau*3A and run in a 1% gel in TAE electrophoresis buffer. Fragments of 0.5–2 kb were recovered with GeneClean kit (Biolabs) and cloned into the *Bam*HI site of pUC19. The ligation mixture was transformed into JM83 *E. coli*. Recombinant plasmids were screened with the total DNA probe of *O. alta* by a standard colony hybridization procedure of Sambrook et al. (1989). The clones showing strong hybridization were screened again. A few clones showed very strong hybridization, suggesting that they might contain highly repetitive sequences.

PCR amplification

Two oligonucleotide primers (P1: 5′-TTACTGTAGCTCGGG CACC-3′, P2: 5′-GGTGCTTTTTGAGCGATCTC-3′) were synthesized at this laboratory. Plasmid DNA (10 ng) or *O. alta* total DNA (100 ng) were used per 50 µl of polymerase chain reaction (PCR) (10 m*M* Tris pH 8.3, 50 m*M* KCl, 1.8 m*M* $MgCl_2$, 0.1% Triton X-100, 200 µ*M* dNTPs, and 1 unit of *Taq* polymerase). PCR amplifications were performed in a Perkin Elmer Cetus Thermocycler Model 480

Table 1 Rice species, complexes, genome groups (Vaughan 1994), and origin

Species complex Taxa	Genome group	IRRI accession number	Origin
O. sativa complex			
O. sativa	AA	IR36	
O. glaberrima	AA	TOG6542	Africa
		103121	
O. barthii	AA	101937	Senegal
O. rufipogon	AA	106424	Vietnam
		106412	
O. longistaminata	AA	wspp89–364	Mali
O. officinalis complex			
O. punctata	BB	105980	Cameroon
	BBCC	104975	Kenya
O. minuta	BBCC	101141	Philippines
		p 90–18	Philippines
O. eichingeri	CC	101422	Uganda
O. officinalis	CC	101399	Vietnam
		105220	Indonesia
		100176	
O. rhizomatis	CC	105448	
O. alta	CCDD	105138	Suriname
		100161	Brazil
		101395	
O. grandiglumis	CCDD	101405	Brazil
		105155	Brazil
		105156	
O. latifolia	CCDD	100914	Mexico
		100955	
O. australiensis	EE	105269	Australia
		105272	Australia
O. meyeriana complex			
O. meyeriana	Diploid	104990	Malaysia
O. granulata	Diploid	106444	India
O. ridleyi complex			
O. ridleyi	Tetraploid	100820	Thailand
O. longiglumis	Tetraploid	100974	Indonesia
		105146	Indonesia
		105148	Indonesia
Species not in complexes			
O. brachyantha	FF	101232	Sierra Leone
O. Schlechteri	Tetraploid		

using the following profile: 94 °C for 5 min, followed by 35 cycles of 94 °C, for 30 s, 60 °C for 1 min, 72 °C for 1 min 30 s, and finally 72 °C for 10 min.

Southern blot hybridization

Restriction-digested DNA was fractionated by electrophoresis in 1% agarose gels and transferred onto Hybond N plus (Amersham) membranes. The conditions for DNA hybridization with [^{32}P]-labelled DNA probes and autoradiography have been previously described (McCouch et al. 1988).

Copy number estimation

Serial dilutions of the RRS7 plasmid DNA (0.01–50 ng) and the genomic DNA (100–1000 ng) of various species were transferred to Hybond N plus membranes (Amersham). [^{32}P]-labelled RRSS1.2 was used as the probe. Hybridization conditions were the same as those used in the Southern blot hybridization. Spots with the same hybridizing intensities were considered to have the same amount of target sequence.

DNA sequencing

The DNA sequence of RRS7 was determined with the standard dideoxy-nucleotide termination method using a 373A Automated DNA Sequencer (Applied Biosystem). The sequences of both ends were first determined, and two primers for PCR amplification were synthesized accordingly to subclone RRS7. Two fragments of about 1.2 kb and 0.6 kb were cloned and sequenced. Meanwhile, several subclones obtained by partial digestion of RRS7 with *Sau*3A were also sequenced. The sequence data were analyzed using the DNASIS program.

Results

Isolation and sequence analysis of RRS7

The putative clones containing repeated DNA sequences were used as probes for Southern blot analysis with the total DNA of *O. alta*. One clone, RRS7, was shown to be a tandemly repeated DNA sequence (not shown and see below). This clone was used for further analysis.

The DNA sequence of the 1844-bp insert of RRS7 defined by the *Sau*3A site is shown in Fig. 1. As can be seen there are six monomeric units arranged tandemly in RRS7. These repeat units are all 155 bp in length except for unit 6, which contains an imperfect internal duplication and is thus 39 nucleotides longer than the others. The 6 repeated units are not identical to each other, but share high sequence identity. They were aligned to demonstrate maximum similarity, and the 155-bp consensus sequence was subsequently derived (Fig. 2). The monomers were found to have between 92.9% and 96.8% homology to the consensus sequence, and the latter did not contain any internal repeats. A search within the EMBL data library showed no other DNA sequences having significant homology.

A 871-bp segment (nucleotides 1–871) precedes the tandemly repeated array, and was observed to have poor sequence homology to the repeats. This segment was shown to be a single-copy sequence (data not shown). It is noteworthy that there is a simple repeat sequence, $(TG)_7$, present within this segment.

Characterization of the RRS7 family by Southern blot analysis

The characteristic of the presence of a tandem repeat of the RRS7 family in *O. alta*, i.e., the restriction enzyme periodicity, was found in the total DNA digests of many enzymes—*Msp*I, *Mbo*I, *Hin*fI, *Taq*I, *Dra*I, *Alu*I, *Hae*III, and *Rsa*I (see below and data not shown). These enzymes can be classified into two groups: group I, including *Msp*I, *Mbo*I, *Hin*fI, and *Taq*I, whose recognition site was found in the repeat units of RRS7; group II, including *Dra*I, *Alu*I, *Hae*III, and *Rsa*I, whose recognition site was not found in the repeat units of RRS7. As in the case of the human alpha satellite DNA family (Waye and Willard 1987), smaller subfamilies of members within large satellite families, identified by restriction sites not present in the rest of the family, are usually present in a non-overlapping way. Thus, the RRS7 family may be classified into at least two subfamilies: one represented by *Msp*I and the other represented by *Dra*I.

The genomic organization of the RRS7 family in other *Oryza* species was also examined. Figure 3 shows the hybridization of RRS7 to *Msp*I and *Dra*I digests of DNA samples representing various genomes and species. The *Msp*I digestion produced similar ladder patterns in genomes AA, BB, BBCC, CCDD, and EE (Fig. 3A). In genome CC (Fig. 3A lane e), however, most of the signal appeared in the high-molecular-weight range and only three faint bands were below 2 kb. These results suggest that the *Msp*I subfamily is present in genomes AA, BB, BBCC, CCDD, and EE but not present in genome CC. On the other hand, *Dra*I digestion produced ladder patterns in all of these genomes (Fig. 3B). The 3 species (*O. alta*, *O. grandiglumis*, *O. latifolia*) containing genome CCDD had almost the same hybridizing pattern, with one characteristic being that the smaller the fragment, the stronger the intensity. This pattern is distinct from those of all the other genomes. The patterns in other genomes were also distinguishable from each other. The ladder patterns suggest the existence of a *Dra*I subfamily in these genomes. In *O. brachyantha* (genome FF) and the 5 species of unknown genomes (*O. schlechteri*, *O. meyeriana*, *O. ridleyi*, *O. granulata*, and *O. longiglumis*), only very faint hybridization was observed even in the longer exposed autoradiograph. Further studies are needed to elucidate the genomic organization of the RRS7 family in these genomes.

Two pairs of isoschizomers, *Sau*3 A/*Mbo*I and *Msp*I/*Hpa*II were used for Southern blot analysis to

Fig. 1 Nucleotide sequence of the 1844-bp *Sau*3A fragment of RRS7 (EMBL accession number X86001). The tandemly repeated units are designated *1–6* with a *small arrow* indicating their beginning. The two imperfect direct internal 39-bp repeats (*a* and *a'*) in monomer 6 are shown in *boldface* and *underlined* by *arrows*. The dinucleotide TG simple sequence repeat is shown in *boldface*

Sau3A

1 GATCTTGAAT CGCTTTGGCT ATATTGACAG CAAGTCTTCT CCAACGCCTT ATGATCTAAC

61 ATTAGATACA TTTTGATAAT ATCTCAAACT TCTTTATATT TGATTAGAAA TAAGTCAAAG

121 CAACTTAGAA TATGAAACGG AGAAAGTATT ATTCTTCTTA AAAAAAAGGA ACATTTAAAG

181 ACTTAGTTGA TATTTCATCC GCAAGATATA TATTCTGAAA GAAATACAGT CCCCAAATTT

241 TTTTCTTATG GGTGATTTTA TGGTTCGTGG GGAGATACTG AGAGGCTCCT TTGAAATTTC

301 GTATCTCTTC AGTATTTAAG TGGATGTACC AAAAATTACT GTAGCTCGGG CACCGTAAAA

361 TATATCTCCT CTCTTTTCTT CAATCTGAAT AAGCTAATAA CATCAAAAAA AAATTCCAAA

421 CTAACCGTTT TTGCGACTAG AGGAGTAGTA CGTAAGAGAC GGTGAGTTGG CGAATAGAAG

481 TCCCATTATC AACATCTCCA ATATCTAGTG ATGTTGTCTG TGTGTGTGTG TGCAGGTGGA

541 ACAAGCCAAG AAAAGACTGA ATCAATGGGA GGACAAGAAG CAACCGCTTC TTGCAGGCAC

601 GGATGACTAC TGATTCACAT GACAAATGAT GGGTAAAGTG CGGATGCGCG CTTATGCATA

661 TATACCCAAC GCCCAATGGC GCCGAAATGC ATGGTGCACG AGTTTCTTGG GTGAAATGCT

721 TGTTGGTGTT TGCACGTACA CCGTGTTTTT CCTTTCGTCC CTATATGCCC CTATATATGT

781 TGAGCTTGAA TTGCAATTAG AGGTGACCTC ATTGATTGGT TTTTCTTACG CGTTTTTATC

841 TTTTCTTCTT GAACATATTT ATTCATCATG AGATCGCTCA AAAAGCACCA AAACATGAGT

901 TTTGGACATA TTGGAGTGGA TTGGGTGCGT TCGTTTCGAA AAATCTCTCC GTGACTCGCG

961 CGGTGAACTT TTCACATTTA ATGCAGATAT TCGGTCAGTG CAGTGTGATG TTTCTCACCG

1021 GAACGAGATC GTTCAAAAAG CACCAAAACA TGAGTTTTGG ACATATTGGA GTGGATTGGG

1081 TGCGTTCGTT TCGAAAAATC TCTCCGTGAC TCGCGCGGTG AACTTTTCAC AATTAATGCC

1141 GATATTCGGT CAGTGCGGTG CGATGTTTCT CACCGGAACG AGATCGTTCA AAAAACACTA

1201 AAACATGAGT TTTGGACATA GTGGGCTCGA TTGGGTGCGT TCGTTTCGAA AAATCTCACC

1261 GTGACTCGCG CGGTGAACTT TTCACATTTA ATGCCGATAT TCGGTCAGTG GAGTGGGATG

1321 TTTCTCACCG GAACGAGATC GTTCAAAAAG CACCAAAACA TGAGTTTTGG ACATATTGGA

1381 GTGGATTGGG TGCGTTCGTT TCGAAAAATC TCTCCGTGAC TCGCGCGGTG AACTTTTCTC

1441 AATTTATGCC GATATTTGGT AAGTGCAGTG CGATGTTTCT CACCGGAACG AGATCGCTCA

1501 AAAAGCACTA AAACATGAGT TTTGGACATT GTGGGCTCGA TTGGGTGCGT TCGTTTCGAA

1561 AAATCTCTCC GTGACTCGCG CGGTGAACTT TTTACATTTA ATGCAGATAT TCGGTCAGTG

1621 CAGTGGGATG TTTCTCACCG GAACGAGATC GTTCAAAAAG CACCAAAACA TGAGTTTTGG

1681 ACATATTGGA GTCGATTGGG TGCGTTCGTT TCGAAAAATC TCTGCGTGAG TGGATTGGGT

1741 GCGTTCGTTT CGAAAAATCT CTCCGTGACT CGTGCGGTGA ACTTTTCACA ATTTATGCCG

Sau3A

1801 ATATTTGGTA AGTTCAGTGC GATGTTTCTG ACCGGAACGA GATC

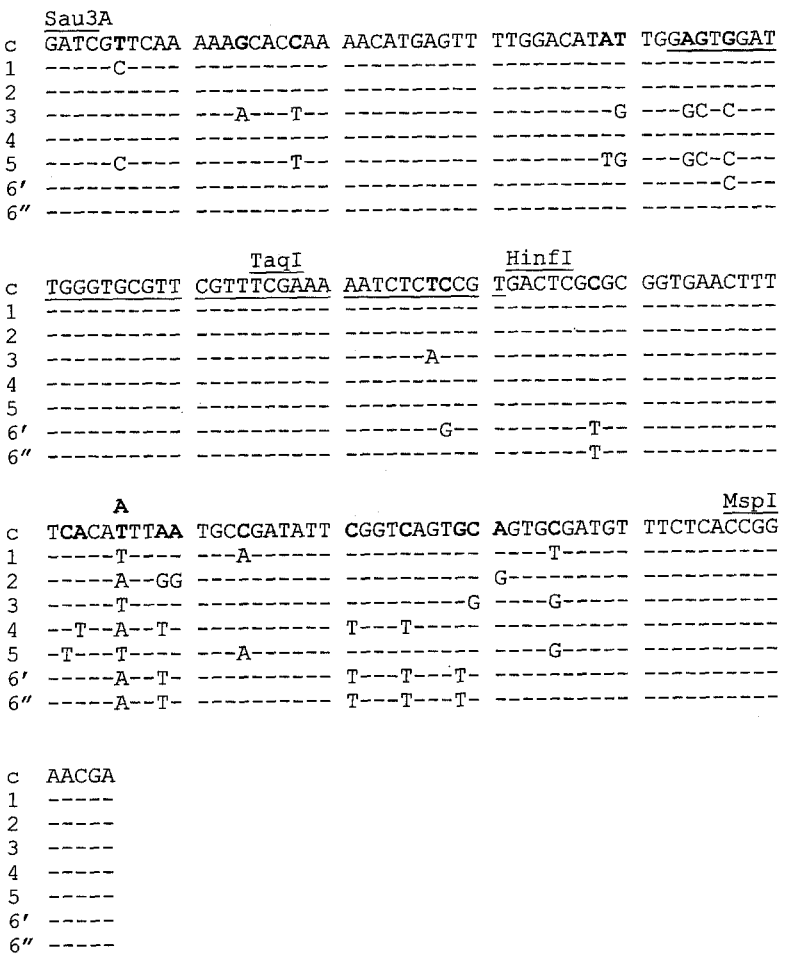

Fig. 2 Derivation of the 155-bp consensus sequence from the monomeric repeated units in RRS7. Six monomeric repeats were aligned to show maximum homology. The 155-bp consensus sequence from these six monomers was based on the most abundant nucleotide found at any given position (*underlined* bases were obtained from the 5 bases in each position in monomer 1–5 and the 2 bases in the duplication in monomer 6). Nucleotides in the monomer sequences that are identical to the consensus sequence are not indicated. The homology of monomer 6 to the consensus sequence was calculated by cancelling one of its internal repeats. *6′* Monomer 6 without its second internal repeat, *6″* monomer 6 without its first internal repeat, *c* consensus sequence. Recognition sites for *Sau*3A, *Hin*fI, *Taq*I, and *Msp*I are indicated *above* the sequence

show the degree of DNA methylation of the *Msp*I subfamily in *O. alta*. Figure 4 shows the results. The relatively faint bands in the *Sau*3A digest and *Hpa*II digest (compare lane a with lane b and lane c with lane d) suggest that most of the C residues in the GATC sites and the internal C in the CCGG sites of the RRS7 family are methylated.

Estimation of the copy number RRS7 family

Using a haploid genome size of 4.5×10^8 bp for the diploid species (Arumanagathan and Earle 1991) and 9.0×10^8 bp for the tetraploid species, we estimated the copy number of RRS7 family in some of the species (Table 2). The copy number of the RRS7-related sequence varies significantly among genomes. Considerable variation in copy number was also found within genomes, e.g., the copy number in the 3 species with genome CCDD changes up to ninefold. In all of the species examined, the highest copy number was found in *O. grandiglumis*, while the 3 species with genome CC have the lowest copy number.

Discussion

A novel non-specific tandemly repeated DNA family, RRS7, was identified in *Oryza* species. The RRS7 family is detectable in genome groups AA, BB, BBCC, CC, CCDD, and EE, the copy numbers varying from 2.9×10^2 to 1.1×10^5. RRS7-related sequences must have been present in the ancestor of *Oryza*, because a small amount of RRS7-related sequences are also detectable in genome FF and the unknown genomes. These have been amplified to various degrees in different genomes during the evolutionary course.

Non-specific repetitive DNA sequences can be used to investigate the evolutionary relationships among rice genomes and species (Wu and Wu 1992). It was shown in this study that genome FF, which has been found to be dissimilar to all other known *Oryza* genomes, contains only a very small amount of RRS7-related sequences. The 3 species with genome CCDD have the very similar hybridizing pattern produced by *Dra*I digestion, reflecting their very close relationship (Jena and Khush 1989). The amplification and homogenization of satellite DNA

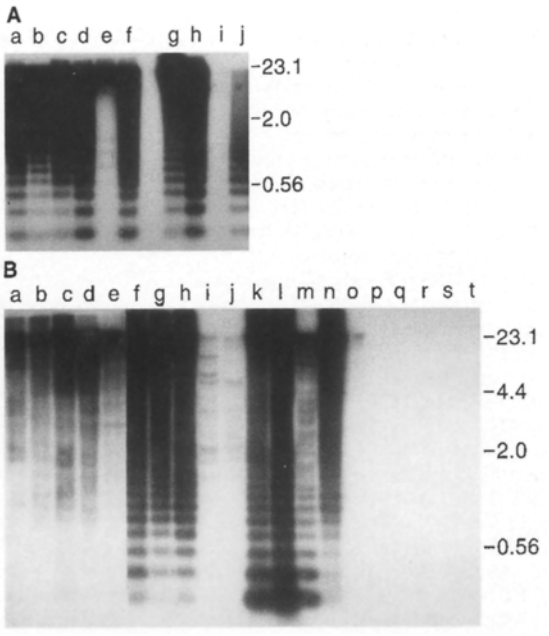

Fig. 3A, B Southern blot analysis of the DNAs of various rice species. Total DNAs were digested with *Msp*I (**A**) and *Dra*I (**B**) and hybridized with RSS7. DNA size markers are indicated in kb. All of the samples listed in Table 1 were analyzed (only part of them are shown here). No different hybridizing pattern was found between the different accessions of the same species. **A** *lane a O. punctata* 105980, *b O. punctata* 104975, *c O. minuta* 101141, *d O. grandiglumis* 101415, *e O. officinalis* 101399, *f O. latifolia* 100914, *g O. alta* 105138, *h O. australiensis* 105249, *i O. brachyantha* 101232, *j O. sativa* IR36. The strong signal in the high-molecular-weight region in *lane g* is caused by partial digestion, but the ladder pattern still exists in the completed digest (not shown). **B** *Lane a O. sativa* IR36, *b O. glaberrima* TOG 6542, *c O. barthii* 101937, *d O. rufipogon* 106424, *e O. longistaminata* wspp 89–364, *f O. minuta* p 90–18, *g O. punctata* 104975, *h O. punctata* 105980, *i O. eichingeri* 101422, *j O. officinalis* 105220, *k O. alta* 100161, *l O. grandiglumis* 105155, *m O. latifolia* 100955, *n O. australiensis* 105272, *o O. brachyantha* 101232, *p O. schlecheri*, *q O. meyeriana* 104900, *r O. ridleyi* 100820, *s O. granulata* 106444, *t O. longiglumis* 105148

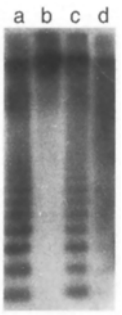

Fig. 4 Southern blot hybridization of RRS7 to restriction enzyme digests of *O. alta* total DNA. DNA (2 μg) was digested to completion with restriction enzymes *Mbo*I (*a*), *Sau*3 A (*b*), *Msp*I (*c*) and *Hpa*II (*d*) and hybridized with [^{32}P]-labelled RRS7

sequences can lead to the evolution of genomes, which may in part explain the origin of species discontinuities and biological novelty (Dover 1982, 1986). From our results it seems that the majority of the RRS7-related sequences in genome CCDD are present in the DD genome. Three species with genome CC, including *O. officinalis*, diploid *O. eichingeri*, and *O. rhizomatis*, contain much lower copy numbers of RRS7-related sequences than the species with genomes BBCC and CCDD. The low copy number of the RRS7 family in genome CC can be confirmed by the similar copy number detected in genomes BB and BBCC. If this hypothesis is correct, the RRS7 family will be very useful for studying the origin of genome DD.

The genomic organization of satellite DNA can reflect its age. The multimeric forms are expected to increase in amount as the satellite ages, and the restriction sites are lost through mutation (South 1975). The amplification of the *Dra*I subfamily has probably occurred relatively recently based on the large number of monomeric units defined by the *Dra*I recognition site. The dissimilar hybridizing pattern produced by *Dra*I digestion between genome CCDD and other genomes suggests that the amplification and homogenization of the *Dra*I subfamily in genome CCDD might have occurred during or after the formation of genome DD. Further studies on the sequences and genomic organization of the *Dra*I subfamily may be helpful for elucidating the evolutionary relationship of genome DD with other genomes.

Table 2 The copy number of RRS7-related sequences

Species	IRRI accession number	Genome groups	Copy number in haploid genome	Percentage of genome size
O. latifolia	100914	CCDD	1.3×10^{4a}	0.22
O. alta	101395	CCDD	7.8×10^{4a}	1.21
O. grandiglumis	105156	CCDD	1.1×10^{5a}	1.97
O. officinalsis	100176	CC	9.3×10^{2}	0.032
O. eichingeri	101424	CC	4.3×10^{2}	0.015
O. rhizomatis	105448	CC	2.9×10^{2}	0.01
O. minuta	101141	BBCC	2.3×10^{4a}	0.40
O. puntata	104975	BBCC	3.0×10^{4a}	0.52
O. puntata	104980	BB	2.3×10^{4}	0.80
O. australiensis	105272	EE	2.6×10^{4}	0.90
O. glaberrima	103721	AA	1.2×10^{4}	0.40
O. rufipogon	106412	AA	7.8×10^{3}	0.27
O. sativa	IR36	AA	9.3×10^{3}	0.32

[a] 2n (CD or BC) as the haploid genome

Acknowledgements We thank Drs. D.S. Brar, B.R. Lu and Ms. Ma Socorro Almazan (IRRI) for the rice leaf samples. We thank Drs. B.R. Lu and D.S. Brar for critically reading the manuscript and valuable suggestions. This research was supported by the Rice Genome Research Program of China and the Rockefeller Foundation's International Program on Rice Biotechnology.

References

Arumanagathan K, Earle ED (1991) Nuclear DNA content of some important plant species. Plant Mol Biol Rep 9:229–241

Aswidinnoor H, Nelson RJ, Dallas JF, Mcintyre CL, Leung H, Gustafson JP (1991) Cloning and characterization of repetitive DNA sequences from genomes of *Oryza minuta* and *Oryza australiensis*. Genome 34:790–798

Brutlug DL (1980) Molecular arrangement and evolution of heterochromatic DNA. Annu Rev Genet 14:121–144

Deshpande VG, Ranjekar PK (1980) Repetitive DNA in three *Gramineae* species with low DNA content. Hoppe-seyleris Z Physiol Chem 361:S 1223–1233

Dover G (1982) Molecular drive: a cohesive mode of species evolution. Nature 299:111–117

Dover G (1986) Molecular drive in multigene families: how biological novelties arise, spread and are assimilated. Trends Genet 168: 159–165

Evans IJ, James AM, Barnes SR (1983) Organization and evolution of repeated DNA sequences in closely related plant genomes. J Mol Biol 170:803–826

Flavell R (1980) The molecular characterization and organization of plant chromosome DNA sequences. Annu Rev Plant Physiol 31: 569–596

Hemleben V, Zentgraf U, King K, Borisjuk N, Schweizer G (1992) Middle repetitive and highly repetitive sequences detect polymorphisms in plants. In: Kahl G, Appelhans H, Kompf J, Driesel AJ (eds) DNA polymorphisms in eucaryotic genomes. Huthig, Heidelberg, pp 157–170

Ingham LD, Hanna WW, Baier JW, Hannah LC (1993) Origin of the main class of repetitive DNA with selected *Pennisetum* species. Mol Gen Genet 238:350–356

Jena KK, Khush GS (1989) Cytogenetic relationships among the three species of *Oryza latifolia* complex. Rice Genet Lett 5:74–75

McCouch SR, Kocket G, Yu ZH, Wang ZY, Khush GS, Coffman WR, Tanksley SD (1988) Molecular mapping of rice chromosomes. Theor Appl Genet 76:815–829

Miklos GLG (1985) Localized highly repetitive DNA in vertebrate and invertebrate genomes. In: MacIntype (ed) Molecular evolutionary genetics. Plenum Press, New York, pp 241–319

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Singer MF (1982) Highly repeated sequences in mamalian genomes. Int Rev Cytol 76:67–112

Southern EM (1975) Long range periodicities in mouse satellite DNA. J Mol Biol 94:51–69

Vaughan DA (1994) The wild relatives of rice: a genetic resources handbook. International Rice Research Institute, Los Banos, Philippines

Waye JS, Willard HF (1987) Nucleotide sequence heterogeneity of alpha satellite repetitive DNA: a survey of alphoid sequences from different human chromosomes. Nucleic Acids Res 15: 7549–7569

Wu T, Wu R (1992) A novel repetitive DNA sequence in the genus *Oryza*. Theor Appl Genet 84:136–144

Xin ZY, Appels R (1988) Occurrence of rye (*Secale cereale*) 350-bp family DNA sequences in Agropyron and other Triticeae. Plant Syst Evol 160:65–76

Zhao X, Wu T, Xie Y, Wu R (1989) Genome-specific repetitive sequences in the genus *Oryza*. Theor Appl Genet 78:201–209