Theor Appl Genet (1992) 84:136–144

THEORETICAL AND APPLIED GENETICS

A novel repetitive DNA sequence in the genus *Oryza*

Tiyun Wu[1] and Ray Wu[1,2]*

Field of Botany[1] and Section of Biochemistry, Molecular and Cell Biology[2,]*, Cornell University, Ithaca, NY 14853, USA

Received May 17, 1991; Accepted Accepted October 17, 1991
Communicated by A. L. Kahler

Summary. Repetitive DNA sequences in the genus *Oryza* (rice) represent a large fraction of the nuclear DNA. The isolation and characterization of major repetitive DNA sequences will lead to a better understanding of rice genome organization and evolution. Here we report the characterization of a novel repetitive sequence, CC-1, from the CC genome. This repetitive sequence is present as long tandem arrays with a repeat unit 194 bp in length in the CC-diploid genome but 172 bp in length in the BBCC and CCDD tetraploid genomes. This repetitive sequence is also present, though at lower copy numbers, in the AA and BB genomes, but is absent in the EE and FF genomes. Hybridization experiments revealed considerable differences both in copy numbers and in restriction fragment patterns of CC-1 both between and within rice species. The results support the hypothesis that the CC genome is more closely related to the AA genome than to the BB genome, and most distantly related to the EE and FF genomes.

Key words: Rice – Repetitive sequences – *Oryza* – Tandemly repeated DNA

Introduction

A substantial fraction of higher plant genomes contains various families of repetitive DNA sequences located in both heterochromatic and euchromatic regions (Flavell 1980, 1986). Highly repetitive DNA sequences have been isolated and characterized from plants such as *Secale cereal* (Bedbrook et al. 1980; Appels et al. 1981, 1986), *Arabidopsis thaliana* (Martinez-Zapater et al. 1986), *Oryza sativa* (Wu and Wu 1987; Zhao et al. 1989), *Avena sativa* (Fabijanski et al. 1990), *Petunia hybrida* (Shepherd et al. 1990) and *Actinidia deliciosa* var. *'deliciosa'* (Crowhurst and Gardner 1991).

The rice nuclear genome contains approximately 50% repetitive DNA (Deshpande and Ranjekar 1980). Some highly repetitive sequences show a relatively rapid rate of sequence divergence among rice species (Zhao et al. 1989). Analysis of these repetitive sequences may provide an amplified view of the evolutionary relationship between species within the genus *Oryza*.

The genus *Oryza* includes 20 wild species and two cultigens, *Oryza sativa* and *Oryza glaberrima* steud (Chang 1984). Phylogenetic and evolutionary relationships between species of wild and cultivated rice have received increasing attention in recent years. Rice has been classified into six diploid genome types, AA, BB, CC, DD, EE and FF, and two tetraploid genome types, BBCC and CCDD (Chang 1976). The different genome types have been distinguished on the basis of cytogenetic studies and morphological variations. The evolutionary relationships among species of these different genome types have also been investigated using isozyme polymorphisms (Second 1982, 1985; Glaszmann 1987, 1988; DeKochko 1987), chloroplast DNA restriction fragment length polymorphisms (Ichikawa et al. 1986; Ishii et al. 1986, 1988; Dally and Second 1990), as well as ribosomal gene spacer-length variability (Cordesse et al. 1990). From these studies, hypothetical scenarios have been proposed for the evolution of the different types of genomes and for exchange between them. However, very few publications are available that pertain to the extent of variability of repetitive sequences in either cultivated or wild rice species.

Earlier we reported the isolation and characterization of genome-specific repetitive DNA sequences from the

* Correspondence to: R. Wu

genus *Oryza* (Wu and Wu 1987; Zhao et al., 1989). These sequences are useful markers for studying evolutionary relationships between species of rice within a given genome. However, highly repetitive DNA sequences, which are present in several genome types, are required for the elucidation of the evolutionary relationship between different genomes. In the study reported here we have isolated a repetitive sequence (CC-1) from *O. officinalis*, a member of the CC genome. This repetitive sequence is present in all but the EE and FF genomes. However, the copy number varies greatly, and the restriction patterns are different in the various genomes. Differences are also found in the methylation pattern an in the length of the repeat unit in different genomes.

Materials and methods

Plant materials and growth conditions

All wild rice and IR-derived domestic varieties listed in Table 1 were obtained from T.T. Chang at the International Rice Research Institute (IRRI) through R. Coffman (Cornell University, Ithaca, N.Y.). *O. sativa* var 'Calrose 76' was obtained from N. Rutger (USDA/ARS, Stoneville, Miss.). The growth conditions of all rice cultivars were as described previously (Zhao et al. 1989).

Isolation of total rice DNA and cloning of repetitive DNA

Total DNA was isolated from different rice varieties according to Zhao et al. (1989). For the cloning of the repetitive sequence from *O. officinalis*, total rice DNA was completely digested with *Hin*cII, size-fractionated electrophoretically on a 1% agarose gel, and stained with ethidium bromide. A prominent band corresponding to DNA fragments of approximately 200 bp in length was eluted from the gel and ligated into the *Hin*cII site of pUC13 by blunt-end ligation (Maniatis et al. 1982). The ligation mixture was used to transform *E. coli* JM101 cells, and plasmid pCC-1 containing a 194-bp repetitive sequence was selected using ^{32}P-labeled total DNA of *O. officinalis* as the probe. CC-1-related repetitive sequences from *O. sativa* var 'Labelle' and *O. meridionalis* (AA genome) and *O. alta* (CCDD genome) were cloned using the ^{32}P-labeled pCC-1 fragment as the probe.

Table 1. Rice entries, genome type and origin

Rice entries	Type	Genome	Origin	IRRI accession number
O. sativa				
Taipei 309	*japonica*	AA	China	42576
Labelle	*japonica*	AA	U.S.A.	
O. meridionalis	Wild	AA	Australia	101147
O. longistaminata	Wild	AA	Africa	100930
O. punctata	Wild	BB, BBCC[a]	Africa	103897
O. officinalis	Wild	CC	Asia	103286
O. alta	Wild	CCDD[a]	U.S.A.	100161
O. latifolia	Wild	CCDD[a]	U.S.A.	100165
O. australiensis	Wild	EE	Australia	101467
O. brachyantha	Wild	FF	Africa	101236

[a] Tetraploid (4n = 48). All other rice entries are diploid (2n = 24)

Slot-blot and genomic blot hybridization, and DNA sequence analysis

Preparation, of the labeled DNA probes by nick translation and slot-blot and genomic Southern blot hybridization, and quantitation of copy number of repetitive DNAs were carried out as previously described (Zhao et al. 1989). DNA sequencing was performed using the dideoxynucleotide chain termination procedure adapted to single-stranded M13 phage DNA.

Results

Isolation and characterization of a highly repetitive DNA sequence from rice

The clone CC-1 that contains a repetitive sequence about 190 bp in length was isolated from a size-selected library constructed from genomic DNA of *O. officinalis*. To determine the distribution of this repetitive DNA in species of other genome types, the cloned CC-1 DNA was used as a probe for slot-blot hybridization with 37 rice entries, including domestic varieties and wild species of various genome types (Zhao et al. 1989). DNA from species of the AA, BB, and CC genomes, including the tetraploid BBCC and CCDD, showed various degrees of hybridiza-

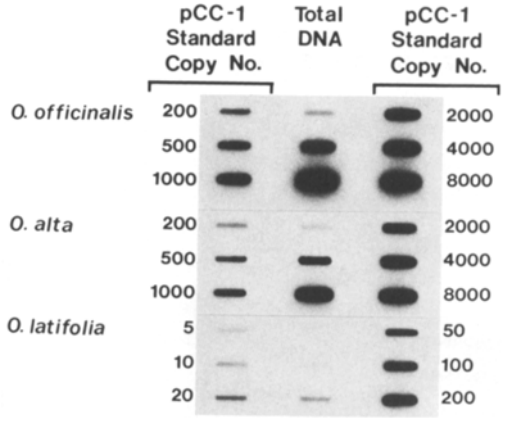

Fig. 1. Copy number determination of rice repetitive DNA sequences of the CC and CCDD genomes. DNA was loaded on nitrocellulose filters by using a slot-blot apparatus. A dilution series of the cloned repetitive DNA was used as a copy number standard (the *right* and the *left columns*). Three different amounts of rice genomic DNA were spotted in the *middle column: top lane* 10 ng, *middle lane* 100 ng, *bottom lane* 1000 ng. The filters were probed with 32P-labeled pCC-1 and exposed for 10 min for DNA samples from *O. officinalis* and *O. alta*, and 5 h for DNA samples from *O. latifolia*. Quantitation of the hybridization signal was done by tracing the X-ray film with a densitometer

tion, while no hybridization to DNA from the EE and FF genomes could be detected (data not shown).

The species with the CC and CCDD genomes which showed variation in the abundance of the CC-1 sequence were chosen for further quantitation by slot-blot hybridization. The copy number of the CC-1 sequence in *O. officinalis* was estimated to be 20,000 per haploid genome (Fig. 1). The most dramatic variation in the abundance of the CC-1-related sequences among species within a genome was found between *O. alta* and *O. latifolia* (CCDD genome). The two species showed a 1000-fold difference in copy number of the CC-1 related sequences: 15,000 copies in *O. alta* and only 15 copies in *O. latifolia* (Fig. 1).

Organization of the CC-1 and related sequences in various genome types

Southern blot analysis of performed to examine the arrangements of the CC-1 and its related sequences in the various genomes (Fig. 2A). The results suggest that a substantial amount of the CC-1 repetitive sequence in the AA, CC, and CCDD genomes is present as tandem repeats since monomer-size fragments (Fig. 2A, the lowest bands in lanes 1, 2, 6, and 7) were produced from *Hin*cII digestion. Only dimer-size fragments were produced from the two wild species of the AA genome (Fig. 2A, lanes 3 and 4). DNA sequence analysis revealed that the fragments indeed contained two tandem repeats of the CC-1 sequence, which showed 96.8% identity.

As shown in Fig. 2A DNAs from different genomes displayed different restriction patterns of the CC-1 sequence. DNAs from two cultivars of the AA genome showed identical patterns (Fig. 2A, lanes 1 and 2). The two wild species of the AA genome also showed identical patterns (Fig. 2A, lanes 3 and 4). However, the two cultivars show very different *Hin*cII restriction patterns from the two wild species (compare Fig. 2A, lanes 1 and 2 to lanes 3 and 4). In addition, a CC and a CCDD species showed differences both in restriction patterns and in abundance of the CC-1-related sequence (Fig. 2A, lanes 6 and 7). The hybridization signal of one of the CCDD genome species, *O. latifolia,* was too weak to be visible in this photograph (Fig. 2A, lane 8), but was visible on the original X-ray film. This weak signal is consistent with the low copy number in this species as determined by slot-blot hybridization (Fig. 1). Also consistant with the results from slot-blot hybridization (data not shown) is the fact that species of the EE and FF genomes lack any CC-1-related sequences (Fig. 2A, lanes 9 and 10).

It is interesting to note that the size of the CC-1 repeat unit in the CCDD genome appears to be smaller than that in the CC species (compare Fig. 2A, lanes 6 and 7). This was shown more clearly in Fig. 2B. A Southern blot of *Hin*cII-partially digested genomic DNA from *O. offic-*

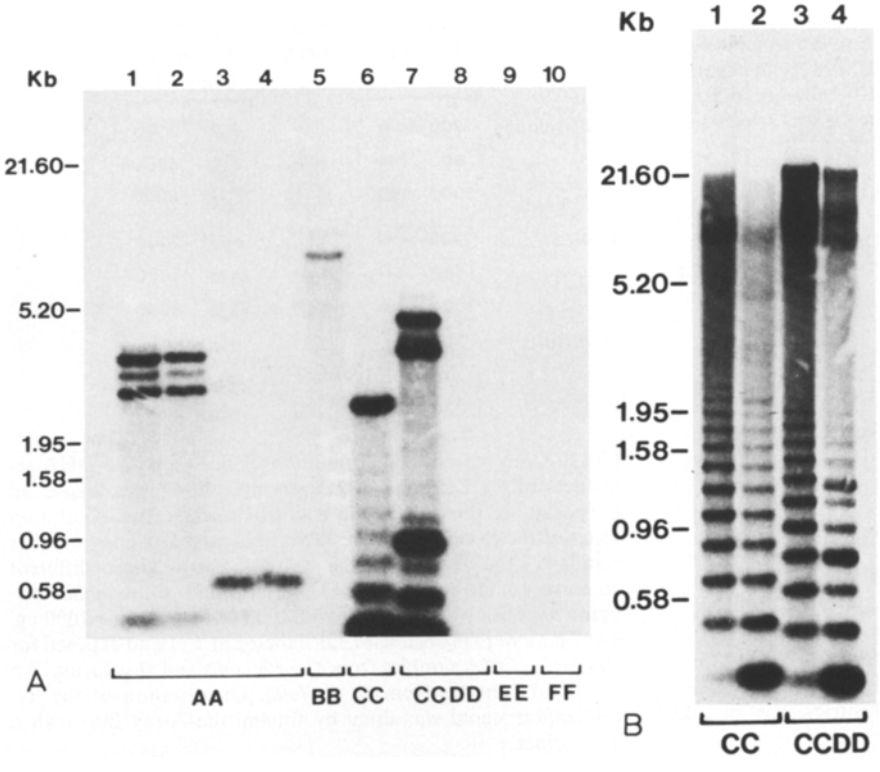

Fig. 2 A, B. Genomic blot analysis of different genomes of rice using pCC-1 as a probe. **A** Total rice DNA (5 μg/lane) from different genomes (as indicated at bottom) was digested to completion with restriction enzyme *Hin*cII and fractionated electrophoretically on a 1% (w/v) agarose gel. *Hin*dIII- and *Eco*RI-digested lambda DNA was run as size markers (*left-side margin*). The DNA samples were transferred to a NYTRAN filter and hybridized with ^{32}P-labeled CC-1. DNA from different rice species were loaded in different lanes as follows: *Lane 1 O. sativa* var 'Labelle', *lane 2 O. sativa* var 'Taipei 309', *lane 3 O. meridionalis, lane 4 O. longistaminata, lane 5 O. punctata, lane 6 O. officinalis, lane 7 O. alta, lane 8 O. latifolia, lane 9 O. australiensis, lane 10 O. brachyantha.* **B** The same amount of DNA as in 2A was partially digested with *Hin*cII for 2 min (*lanes 1 and 3*) and 30 min (*lanes 2 and 4*). The remaining steps were the same as described as in A. *Lanes 1 and 2* DNA from *O. officinalis, lanes 3 and 4* DNA from *O. alta*

inalis (CC genome, lanes 1 and 2) and *O. alta* (CCDD genome, lanes 3 and 4) was run for a longer time to better resolve the *Hin*cII fragments. The sequence is tandemly repeated in both types of genomes, but the size of the repeat unit in the CC genome (lanes 1 and 2) is larger than that in the CCDD genome (lanes 3 and 4).

Repeat unit length variation and genome-specific modification of the CC-1 DNA

The CC-1-related repeats in *O. officinalis* and *O. alta* were cloned for further examination of their sequence and structure. Sequence analysis revealed that the basic repeat unit was 194 bp in *O. officinalis* (CC genome) and only 172 bp in *O. alta* (CCDD genome), which was consistent with the result from genomic Southern blot analyses (Fig. 2B). There is a 21-bp deletion between nucleotides 64 and 84 (Fig. 3A) in *O. alta*. Each CC-1 repeat unit from the CC genome contains an imperfect tandem repeat of the 21-bp sequences: positions 43–63 and 64–84 (Fig. 3B). It should be noted, however, that some of the repeat units contain single base pair insertions or deletions. Thus, the actual length may vary among repeat units. The 172-bp sequence in CCDD genome shares 85.5% identity with that of CC-1, excluding a copy of the 21-bp repeat (Fig. 3A). There are also several short imperfect, direct or inverted, repeats within the major repeat sequence of CC-1.

CC-1-related clones were also isolated and sequenced from size-selected libraries of other genomes. The results showed that the basic repeat unit was 192 bp in the AA genome due to a 2-bp deletion. A DNA fragment of 384 bp from the species *O. meridionalis* of the AA genome (Fig. 2, lane 3) was sequenced and found to be a dimer of the CC-1 sequence that contains a *Hin*cII site at the expected position (data not shown). This result suggests that every other *Hin*cII site within the tandem repeats of the 192-bp sequence is specifically modified in this wild species within the AA genome (and perhaps also in the other wild species, *O. langistaminata,* lane 4, Fig. 2A), since only dimer-sized fragments were found from the *Hin*cII-digested genomic DNA (Fig. 2A). Three individual clones from *O. officinalis* were sequenced. These CC-1-related sequences showed over 97% identity. The CC-1-related sequences showed 94.2% identity between AA and CC genomes, and 96.8% between the two repeats in the dimer from the AA genome.

The CC-1 sequence is comprised of 64.4% G+C residues, which is much higher than the average 44% G+C content of the rice genome, but is characteristic of other highly repeated tandem arrays in plants. No signif-

A

CC-RELATED SEQUENCES

```
                10         20         30         40         50
CC-1      GCTTTTAGGC AAGCCAGCCG GCTTGTGCCC CAGGGCACCC *TACCGAGGG
CCDD-1    -----G--C- -----G---- --G-T----- ----A----- C---------

                60         70         80         90        100
CC-1      GGTGGTGTAC TGTACCAAGG GGGTGGTGGA TTCTCCCGAA CCCCCGTATC
CCDD-1    -------G-T ---******* ********** ****T--A-- G---------

               110        120        130        140        150
CC-1      GGTCGGGGTA CGTGGGGATC GCAGGTCGGG AAGTGCTCAG CGTGTGTCAA
CCDD-1    -------A-G ---------- --G------- ---------- ----------

               160        170        180        190
CC-1      AAGCAAGCCG TCGGGCCGGA TGGCTC*GCA TGCCACAAGC CAGCTG
CCDD-1    -----G---- ---------C -----TT--- *----**--- ---T--
```

B

```
          43                  63
          |                   |
          ACCGAGGGGGTGGTGTACTGT
          ||| |||||||||| | | |
          ACCAAGGGGGTGGTGGATTCT
          |                   |
         64                  84
```

Fig. 3A, B. Comparison of the CC-1 sequences from *O. officinalis* and *O. alta*. **A** The sequence of the CC-1 from *O. officinalis* is given. For the sequence from *O. alta* (CCDD-1), only those nucleotides that differ from the CC-1 sequence are shown, and the identical sequences are represented by *dashes*. *Asterisks* indicate positions where deletions occurred in the CCDD DNA from *O. alta*. **B** Internal duplicated sequence in CC-1; *numbers* indicate the positions of CC-1 nucleotides

icant sequence similarity to the CC-1 sequence was found among sequences listed in the GEN EMBL database of 1990.

Distribution of CC-1 sequences in species of the AA and CC genomes

To examine more closely the relationship among rice species in different genomes, the cloned CC-1 was used as a probe for genomic Southern blot analyses of species of AA, CC, and CC-related (BBCC and CCDD) genomes. The restriction enzyme *Eco*RI, whose site was absent in the repeat units of CC-1, was chosen for the experiment. Genomic DNA from 39 species of rice within the AA genome was analyzed (Table 2), and the Southern blot results are shown in Fig. 4. Polymorphisms in restriction patterns and differences in copy number were found among different species. For example, the CC-1-related sequences were present at a lower copy number in an *indica* variety (lane 1) than in a *japonica* variety (lane 2). The size of the restriction fragments and abundance of the CC-1 sequences were similar for example between the two varieties shown in lane 1 and lane 3, and also between two varieties shown in lane 2 and lane 4. There were also variations in restriction patterns and abundance of the CC-1 sequences among wild species of the AA genome (lanes 5–39). The CC-1-related sequences were found to be less variable in 39 varieties of the same group; for example, the *indica* as a group or the *japonica* as a group (data not shown).

Differences in the length and number of restriction fragments were also found in species of the CC genome and CC-related tetraploid genome types (Table 3 and Fig. 5A). In contrast, larger variations were found in the abundance of the CC-1 sequences among species both between and within genome types (Fig. 4B). For example, the copy number is very low for rice DNA samples in lanes 4, 11, 27, 35, 36, 38 and 39; the copy number is very high for samples in lines 14, 15, 18, 23, 26, 31, 34, etc. Furthermore, the restriction patterns for samples in lanes 5, 18, and are 23 similar to one another but different from other samples.

The filter used for the experiment shown in Fig. 5A was washed and rehybridized with a CC genome-specific repetitive sequence, Oo2 (Zhao et al. 1989), and the results are shown in Fig. 5B. Consistent with our earlier results, the Oo2 sequence hybridized only to DNA from the CC genome and several species of the BBCC genomes. Apparently, the CC-1 and Oo2 sequences belong to two different repetitive sequence families. They show different patterns of hybridization and differential distribution among species of the CC genome. For example, *O. officinalis* (from Burma) contains a high copy number of CC-1 repeats, but a very low copy number of Oo2 repeats (compare Fig. 5A, lane 24 with 5B, lane 24).

Table 2. Rice entries of AA genome[a]

Lane number	Rice entries	Origin	IRRI accession number	IRRI collection number
1	*indica*		44-1	
2	*japonica*		67-1	
3	*O. glaberrima*	Africa	IR-8-4	102201
4	*O. sativa L.* sp.	Bangladesh	IR26-1	103831
5	*O. rufipogon*	West India, land race	26	DS14
6	*O. rufipogon*	West India	20	RD38
7	*O. rufipogon*	Sri Lanka	21	W555
8	*O. rufipogon*	Thailand	2–14	W162
9	*O. rufipogon*	Thailand	17	W1699
10	*O. rufipogon*	India	4–13	W133
11	*O. rufipogon*	India	4–14	W135
12	*O. rufipogon*	Malaysia		W593
13	*O. rufipogon*		4–8	103822
14	*O. rufipogon*		4–9	Plat88-791
15	*O. rufipogon*	India	16	W1669
16	*O. rufipogon*	China	18	W1654
17	*O. rufipogon*	China	13	W1655
18	*O. glumaepatula*	Paramaribo, Sourinam	IR9-1	100968
19	*O. rufipogon*	America	14	W1185
20	*O. rufipogon*		4–15	W1187
21	*O. rufipogon*	Australia	IR22-3	OR7
22	*O. rufipogon*	Australia	24	OR10
23	*O. rufipogon*	Australia	23	OR39
24	*O. rufipogon*	Australia	31	OR54
25	*O. rufipogon*	Australia	30	W1627
26	*O. meridionalis*	Australia	IR16-3	101147
27	*O. nivara*	India	IR20-1	101508
28	*O. barthii*	Gambia, Africa	IR3-4	100122
29	*O. breviligular*	Botswana		Wb01
30	*O. langistaminata*			TB82
31	*O. langistaminata*		2.7	CB23
32	*O. langistaminata*	Chad	35	TL81
33	*O. langistaminata*	Mali	IR14-1	101378
34	*O. langistaminata*	Mali	2–13	ILL116
35	*O. langistaminata*	Ivory Coast	36	IL52
36	*O. langistaminata*	Botswana	37	W102
37	*O. langistaminata*	Cameroun	32	UL12-6
38	*O. langistaminata*	Guinea	2–10	YL244
39	*O. langistaminata*	Zambia	2–14	ZL14

[a] DNA from the 39 rice entries was transferred to filters to be used for hybridization analysis. Lanes 1–39 correspond to those in Fig. 4

On the other hand, *O. eichigerri* (from Ouganda) contains a high copy number of Oo2 repeats, but a very low copy number of the CC-1 repeats (compare Fig. 5B, lane 17 with 5A, lane 17). In addition, Oo2 is absent in the CCDD (Fig. 5B, lanes 33-38) and some of the BBCC species that contain CC-1 repeats (Fig. 5B, lanes 5–8).

Discussion

We have shown that the rice repetitive sequence CC-1 is present in all but the EE and FF genomes. The results

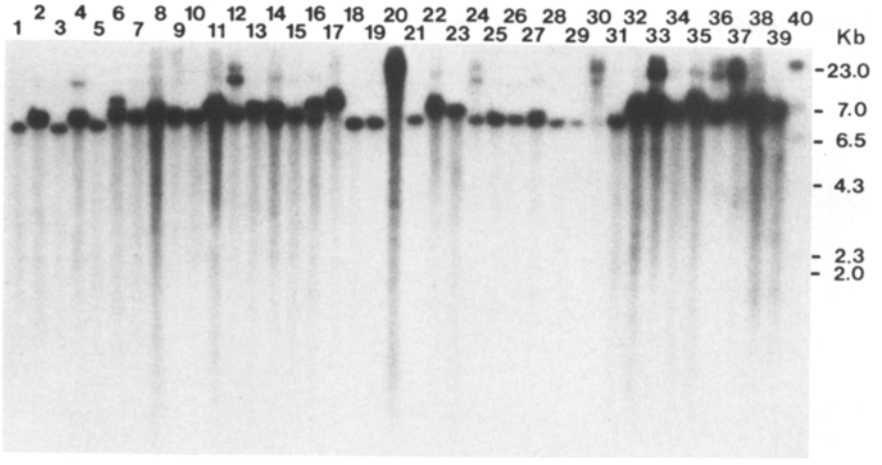

Fig. 4. Southern blot of rice DNA from the AA genomes. Total DNA (5 μg/lane) from different rice varieties of species having the AA genome was digested to completion with the restriction enzyme *Eco*RI and fractionated electrophoretically on a 1% agarose gel. *Hin*dIII-digested lambda DNA was run as size markers. The DNA samples were transferred to a NY-TRAN filter and hybridized with ^{32}P-labeled pCC-1 sequence. *Lanes 1–39* correspond to DNA samples from the species listed in Table 2

Table 3. Rice entries, genome type and origin[a]

Lane number	Rice entries	Genome	Origin	IRRI accession number	IRRI collection number
1	*O. sativa*	AA	Senegal	34	CL7-2
2	*O. punctata*	BB	Cameroun		W1590
3	*O. punctata*	BB	Tanzanie	67	W1515
4	*O. punctata*	BB	Chad		TP43
5	*O. minuta*	BBCC[b]	Philippines	IR18-4	101089
6	*O. minuta*	BBCC[b]	Philippines	IR19-2	101125
7	*O. minuta*	BBCC[b]	Philippines	IR28-5	101141
8	*O. minuta*	BBCC[b]	Philippines	71	103865
9	*O. minuta*	BBCC[b]	Philippines	51	W1331
10	*O. punctata*	BBCC[b]	Ivory Coast	56	IP27
11	*O. punctata*	BBCC[b]	Nigeria	50	W1408
12	*O. punctata*	BBCC[b]	Labo Labo, Ghana	IR23-1	101409
13	*O. malampuzahensis*	BBCC[b]	Via NIG, Japan	IR15-2	100957
14	*O. collina*	CC	Sri Lanka	IR5-1	103410
15	*O. collina*	CC		4--2	103421
16	*O. eichingeri*	CC	Ganda	IR6-1	101422
17	*O. eichingeri*	CC	Ouganda	48	W1526
18	*O. eichingeri*	CC	Uganda	IR7-1	101425
19	*O. eichingeri*	CC	Ivory Coast	54	IP7
20	*O. officinalis*	CC	Thailand	IR21-3	100896
21	*O. officinalis*	CC	Sabah, East Malaysia	IR22-3	101150
22	*O. officinalis*	CC	India	45	D04
23	*O. officinalis*	CC	Sarawak	61	W1278
24	*O. officinalis*	CC	Burma	64	100181
25	*O. officinalis*	CC	Malaysia	44	100180
26	*O. officinalis*	CC	China	40	104618
27	*O. officinalis*	CC	China	58	CH83-3
28	*O. officinalis*	CC	China	4--3	105392
29	*O. officinalis*	CC	China	4--4	105393
30	*O. officinalis*	CC	China	4--5	105394
31	*O. officinalis*	CC	China	4--6	105395
32	*O. officinalis*	CC	China	4--7	105396
33	*O. alta*	CCDD[b]	Via USDA	IR1-2	101395
34	*O. grandiglumis*	CCDD[b]	Brazil	IR10-1	101405
35	*O. latifolia*	CCDD[b]	Campo cotaxtla, Mexico	IR12-3	100914
36	*O. latifolia*	CCDD[b]	Cuba	43	W1168
37	*O. latifolia*	CCDD[b]	South America	60	W1144
38	*O. latifolia*	CCDD[b]		73	100963
39	*O. australiensis*	EE	Australia	43	OA4

[a] DNA from the 39 rice entries was transferred to filters to be used for hybridization analysis. Lanes 1–39 correspond to those in Fig. 5

[b] Tetraploid (4n = 48). All other rice entries are diploid (2n = 24)

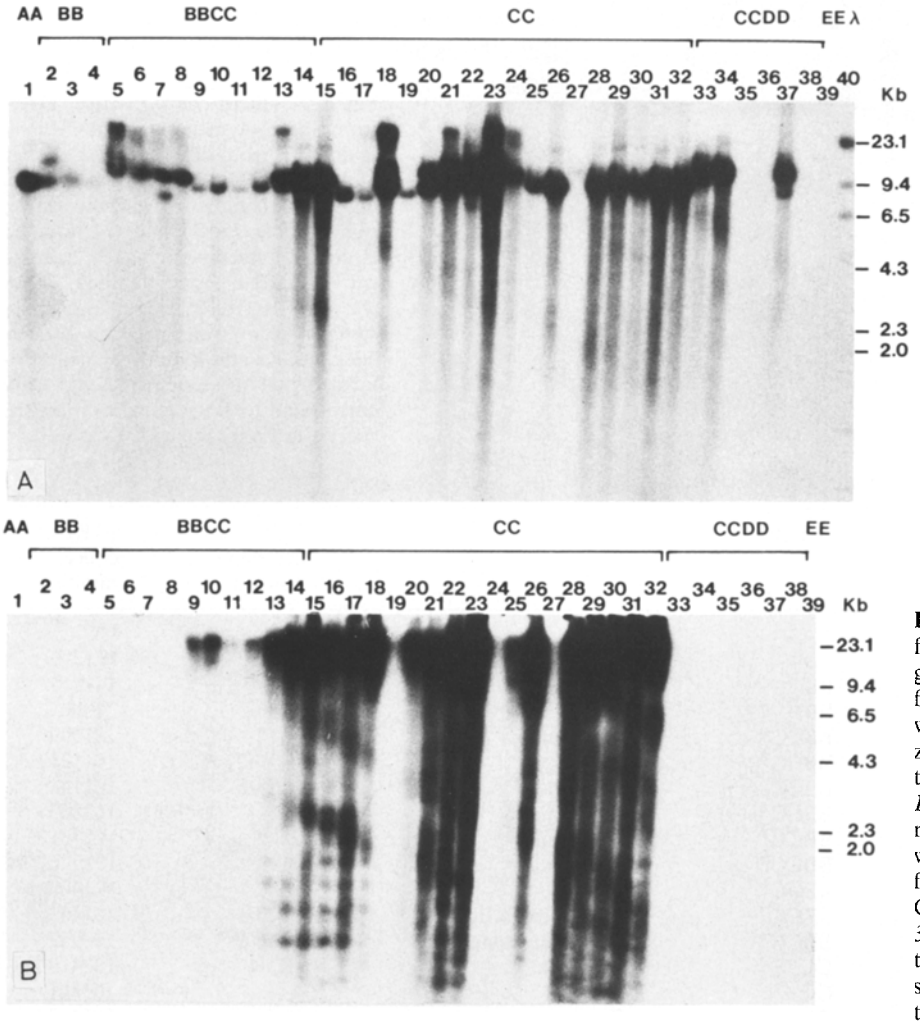

Fig. 5 A, B. Southern blots of DNA from BB-, CC-, BBCC and CCDD genomes. **A** Total DNA (5 μg/lane) from different rice varieties or species was digested with the restriction enzyme *Eco*RI and fractionated electrophoretically on an 1% agarose gel. *Hin*dIII-digested lambda DNA was run as size markers. The DNA samples were transferred to a NYTRAN filter for hybridization using the ^{32}P-labeled CC-1 sequence as the probe. *Lanes 1–39* correspond to DNA samples from the rice species listed in Table 3. **B** The same filter as in A was reprobed with they ^{32}P-labeled pOo2 sequence

suggest that the CC-1-related sequences might have been originally present in the ancestor of rice and subsequently amplified to various degrees in different genomes during evolution. The CC-1-related sequence might have been lost from the EE and FF genomes. Alternatively, the CC-1-related sequence might have been introduced into the other genomes after their divergence from the EE and FF genomes.

The isolation and characterization of both genome-general and genome-specific repetitive sequences may help establish the evolutionary relationships among rice genome types and species. Firstly, by using genome-general repetitive sequences as molecular markers, the evolutionary relationships among genomes may be deduced. Our data, which shows that all rice species of the AA and CC genomes contain the CC-1 sequence, suggest a close relationship between these two genomes. This is consistent with previous cytological evidence showing morphologically similar chromosomes in *O. sativa* (AA) and *O. officinalis* (CC) (Kurata and Omura 1984). Besides cytological studies, recent data from RFLP studies of chloroplast DNA diversity in the genus *Oryza* have also shown that the plastotype of the AA genome is closer to that of the CC genome (as well as the CCDD and BBCC genomes) than to those of the BB and EE genomes (Dally and Second 1990). The fact that the EE and FF genomes showed no hybridization to the CC-1-related sequences indicates a clear distinction at the nuclear repetitive DNA sequence level between AA/CC and EE/FF genomes. Our results suggest that the CC genome is more closely related to the AA genome than to the BB genome, and most distantly related to the EE and FF genomes.

Secondly, both genome-specific (Zhao et al., 1989) and genome-general repetitive sequences are useful in defining evolutionary relationships among species within a given genome. For example, the three tetraploid wild species of *Oryza*, i.e., *O. alta, O. grandiglumis,* and *O. latifolia,* have been found in Latin America, and all three species contain the CCDD genome. Based on cytogenetic relationship among the three species, Jena and Khush

(1988) suggested that the three Latin American tetraploid species belong to one species, the *O. latifolia* complex. Our results also show that these three species of the CCDD genome have very similar, if not identical, restriction fragment patterns and abundance of the CC-1 sequence (Fig. 5A, lanes 33, 34, and 37). In contrast, other varieties of *O. latifolia* (Fig. 5A, lanes 35, 36, and 38), which were classified into a different subgroup, essentially lack the CC-1 sequences. Indeed, subgroups of *O. latifolia* can be distinguished on the basis of the presence of absence of the CC-1 repetitive sequence. This is also consistent with recent results from RFLP studies of nuclear DNA (Wang et al. 1991).

Thirdly, rice species of the same geographic origin show similar restriction fragment patterns and abundance of the two repetitive sequences. For example, different varieties of *Q. minuta* (BBCC genome) from the Philippines (Fig. 5A, lanes 5–8) showed the same restriction fragment patterns and abundance of CC-1 sequences; similarly, varieties of *O. officinalis* (CC genome) from China (Fig. 5A and 5B, lanes 26, 28–32) showed the same restriction fragment patterns and abundance of the Oo2-related repetitive sequences. In contrast, varieties of *O. eichigerri* from different geographical locations show different restriction fragment patterns and abundance of the CC-1 sequence (for example see Fig. 5A, lanes 16–19). However, one variety of *O. officinalis* (lane 27) which was also from China showed no hybridization to the CC-1 sequence but hybridized to a probe that is CCDD genome specific (data not shown). These results may suggest that this variety in fact belongs to the CCDD genome rather than the CC genome. Since *Oryza officinalis* varieties are widely distributed in Asia and certain unusual varieties of the CC genome were also found, it has been speculated that *O. officinalis* from China may have the DD genome (Second 1985). A collection of both genome-specific and genome-general repetitive sequences, as we describe here, is clearly of great importance in establishing the evolutionary relationships both between and within genomes of rice.

Acknowledgements. We thank T. T. Chang, R. Coffman, C. N. Bollich, and N. Rutger for rice seeds and X. Zhao for some of the rice DNAs used in this study. We are grateful to Z. Y. Wang and S. Tanksley for Southern filters (used in Figs. 4 and 5). We thank R. Second, P. K. Ranjekar, H. Xiao, Y. R. Mawal, Z. Y. Wang, M. Lei, and H. Moore for critically reading this manuscript. This research was supported by grant number RF91001, #136, from the Rockefeller Foundation.

References

Appels R, Dennis ES, Smyth DR, Peacock WJ (1981) Two repeated DNA sequences from the heterochromatic regions of rye (*Secale cereale*) chromosomes. Chromosoma 84: 265–277

Appels R, Moran LB, Gustafson JP (1986) Rye heterochromatin. I. Studies on clusters of the major repeating sequence and the identification of a new dispersed repetitive sequence element. Can J Genet Cytol 28: 646–657

Bedbrook JR, Jones J, O'Dell M, Thompson RD, Flavell RB (1980) A molecular description of telomeric heterochromatin in *Secale* species. Cell 19: 545–560

Chang TT (1976) The origin, evolution, cultivation, dissemination and diversification of Asian and African rices. Euphytica 25: 425–441

Chang TT (1984) Conservation of rice genetic resources: Luxury or necessity? Science 224: 251–256

Cordesse F, Second G, Delseny M (1990) Ribosomal gene spacer length variability in cultivated and wild rice species. Theor Appl Genet 79: 81–88

Crowhust RN, Gardner RC (1991) A genome-specific repeat sequence from Kiwifruit (*Actinidia deliciosa* Var. *deliciosa*). Theor Appl Genet 81: 71–78

Dally AM, Second R (1990) Chloroplast DNA diversity in wild and cultivated species of rice (Genus *Oryza*, and section *Oryza*). Cladistic mutation and genetic-distance analysis. Theor Appl Genet 79: 209–222

DeKochko A (1987) Isozymic variability of traditional rice (*O. sativa*) in Africa. Theor Appl Genet 73: 675–682

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

Fabijanski S, Fedak G, Armstrong K, Altossar I (1990) A repeated sequence probe for the C genome in *Avena* (oats). Theor Appl Genet 79: 1–7

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

Flavell R (1986) Repetitive DNA and chromosome evolution in plants. Philos Trans R Soc London 312: 227–242

Glaszmann JC (1987) Isozymes and classification of Asian rice varieties. Theor Appl Genet 74: 21–30

Glaszmann JC (1988) Geograpic pattern of variation among Asian native cultivated rice cultivars (*Oryza sativa* L). based on fifteen isozyme loci. Genome 30: 782–792

Ichikawa H, Hirai A, Katayama T (1986) Genetic analyses of *Oryza* species by molecular markers for chloroplast genomes. Theor Appl Genet 72: 353–358

Ishii T, Terachi T, Tsunewaki K (1986) Restriction endonuclease analysis of chloroplast DNA from cultivated rice species, *Oryza sativa* and *O. glaberrima*. Jpn J Genet 61: 537–541

Ishii T, Terachi T, Tsunewaki K (1988) Restriction endonuclease analysis of chloroplast DNA from A-genome diploid species of rice. Jpn J Genet 63: 523–536

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

Kurata N, Omura T (1984) Chromosome analysis. In: Tsunoda S, Takahashi N (eds) *Biology of rice*. Japan Sci Soc Press, Tokyo, and Elsevier, Amsterdam, pp 305–320

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbour, N.Y.

Martinez-Zapater JM, Estelle MA, Somerville CR (1986) A highly repeated DNA sequence in *Arabidopsis thaliana*. Mol Gen Genet 204: 417–423

Second G (1982) Origin of the genetic diversity of cultivated rice (*Oryza* spp). Study of the polymorphism scored at 40 isozymes loci. Jpn J Genet 57: 257

Second G (1985) Evolutionary relationships in the sativa group of *Oryza* based on isozyme data. Genet Sel Evol 17: 89–114

Shepherd AL, Anderson S, Smith SM (1990) Species-specific repeated DNA sequences from petunia. Plant Sci 67: 57–62

Wang Z, Second G, Tanksley S (1991) Polymorphism and phylogenetic relationships among species in the Genus *Oryza* as determined by analysis of nuclear RFLPs. Theor Appl Genet (in press)

Wu TY, Wu R (1987) A new rice repetitive DNA shows sequences homology to both 5S RNA and tRNA. Nucleic Acids Res 15: 5913–5923

Zhao X, Wu T, Xie Y, Wu R (1989) Genome specific repetitive sequences in the genus *Oryza*. Theor Appl Genet 67: 835–840