Rohde W, Leclercq L, Mattes M,
nce of transposable elements on the
mbo. J. 6:287-294.
ete nucleotide sequence of a soybean

enetics of rice *Oryza sativa* L. Theor.

ity of the large spacer of *Vicia faba*
etitive sequence element. Mol. Gen.

H. Lörz, Max-Planck-Institut für
rmany [present address: Universität
18, D-2000 Hamburg 52, Germany];
noshida, Midori-Ku, Yokohama 227,

and critical reading of the manuscript,
nowledge the financial support of the
grant BDT 03902) and the Rockefeller

ite (1991) Rice genetics II. P.O. Box

Characterization of repeated DNA sequences specific to different rice genomes

F. Cordesse, A.S. Reddy, A. De Kochko, M.C. Kiefer, and M. Delseny

Repeated DNA sequences diverge very rapidly and often become species specific. We adopted two strategies to pick up such sequences in the AA and CC genomes of rice. In the first one, prominent bands on ethidium bromide-stained gels were cut out after preparative electrophoresis of restriction enzyme digests and cloned into plasmid vectors. For the AA genome, we isolated a tandemly repeated 352-bp fragment that shows the same general organization in all *Oryza sativa* accessions. Various rearrangements were observed in *O. rufipogon* and *O. longistaminata*. The presence of an additional Sau 3A site and a higher copy number distinguish indica types from japonicas. The sequence was not detectable in other genomes. Three tandem repeats (374, 367, and 193 bp long) were isolated from *O. officinalis* and were shown to be specific to the CC genome. These sequences permit analysis of the phylogenetic relationships between the CC species. In the second strategy, the intergenic spacer of a cloned ribosomal DNA sequence was dissected, and short sequences with various genome specificity were identified. All these specific sequences should be useful in analyzing AA/CC interspecific crosses.

The *Oryzae* tribe comprises several distinct genome types that have been defined mostly by their capacity to pair during meiosis. The cultivated species (*O. sativa* and *O. glaberrima*) belong to the AA type, as do a few related wild species (*O. rufipogon*, *O. longistaminata*, and *O. breviligulata*). Four other diploid genomes exist in wild species: BB (*O. minuta*), CC (*O. officinalis*, *O. collina*, and *O. eichingeri*), EE (*O. australiensis*), and FF (*O. brachyantha*). In addition, some species are allotetraploids with combinations of these genomes (BBCC: *O. malampuzhaensis* and *O. coarctata*) or of one of them with another genome that has not yet been identified at the diploid level (CCDD: *O. latifolia* and *O. alta*).

Besides elucidation of the phylogenetic relationships among these genomes, there is much interest in the study of wild species because rice breeders use them in interspecific crosses to introgress new favorable agronomic traits into cultivated varieties. Examples of characters that are possessed by the CC genome are resistance to blast, planthoppers, and drought.

The discovery that some repeated DNA sequences are highly species-specific (Grellet et al 1986) prompted us to search systematically for repeated sequences in the AA and CC genomes.

Materials and methods

A list of plant material used in this study as well as a description of growth conditions has recently been published (Cordesse et al 1990).

DNA was prepared from leaves by standard methods. It was digested with various restriction enzymes, and the resulting fragments were separated by agarose gel electrophoresis. Then the DNA was transferred from the gel to a nylon membrane and hybridized with the appropriate probe.

With several restriction enzymes, prominent DNA bands were readily detected on ethidium bromide-stained gels. These DNA fragments were recovered from the gel and used as hybridization probes after labeling by nick translation. Following this preliminary characterization, the fragments were ligated in vitro with the appropriate vector and amplified by cloning in *Escherichia coli*. Resulting colonies were screened by hybridization with the initial fragment eluted from a gel, and positives were further characterized by minipreparation of plasmids, restriction mapping, and sequencing (Maniatis et al 1982).

Results

In this report we describe the isolation and preliminary characterization of one AA- and three CC-specific repeated sequences. We also demonstrate that species-specific repeated sequences are found in the intergenic spacer sequence between adjacent ribosomal RNA genes.

A 352-bp tandem repeat specific to the AA genome

From the cultivar Cigalon, a 360-bp *Eco*RI fragment was cloned and sequenced (De Kochko et al 1991). The sequence is 94% homologous to that isolated from the cultivar Labelle (Wu and Wu 1987). This cloned fragment was used as a hybridization probe to look for identical or related sequences in various *Oryzae* having different genomes.

An example of the results is shown in Figure 1. Very clearly, the sequence is specific to the AA genome, being undetectable in the BB, CC, BBCC, CCDD, and EE genomes. As indicated by the typical ladder pattern, the sequence is generally organized as blocks of tandem repeats. However, during our survey of the AA accessions we observed a number of interesting variations both in copy number and sequence organization.

On the basis of copy number, the AA accessions can be separated into two groups. One comprises indica subtypes of *O. sativa*, *O. longistaminata*, and a fraction of *O. rufipogon* with a few thousand copies. The other group has a much lower copy number (a few hundred or less) and is composed of the japonica subtypes of *O. sativa*, *O. glaberrima*, *O. breviligulata*, plus the remaining *O. rufipogon*. Examples of reampli-

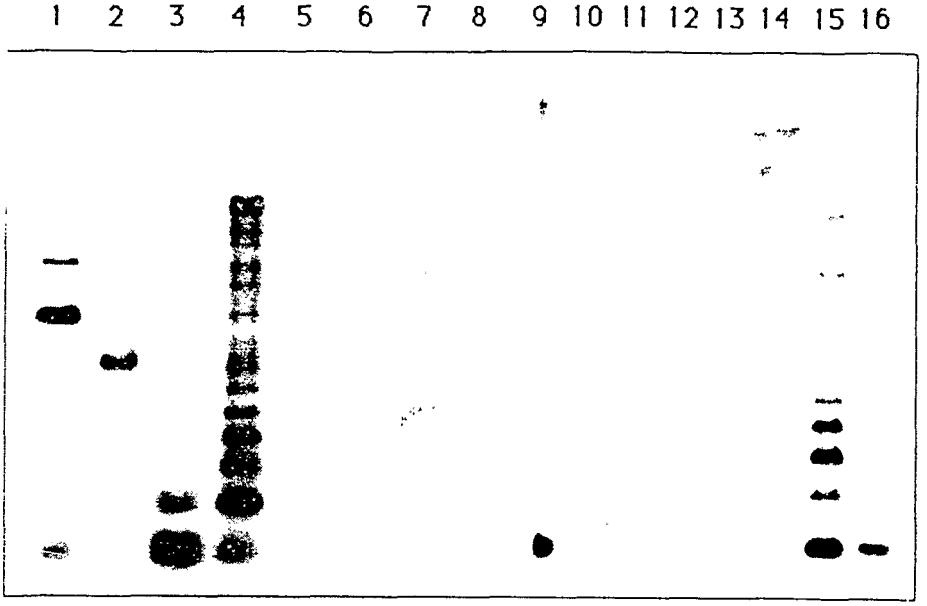
quences are highly species-specific
atically for repeated sequences in the

as a description of growth conditions
)).
ıethods. It was digested with various
nts were separated by agarose gel
rom the gel to a nylon membrane and

DNA bands were readily detected on
ıents were recovered from the gel and
k translation. Following this prelimi-
l in vitro with the appropriate vector
esulting colonies were screened by
ɔm a gel, and positives were further
estriction mapping, and sequencing

nary characterization of one AA- and
demonstrate that species-specific
spacer sequence between adjacent

nome
ıent was cloned and sequenced (De
ous to that isolated from the cultivar
t was used as a hybridization probe
s *Oryzae* having different genomes.
/ery clearly, the sequence is specific
C, BBCC, CCDD, and EE genomes.
ence is generally organized as blocks
f the AA accessions we observed a
nber and sequence organization.
ıs can be separated into two groups.
ɔngistaminata, and a fraction of *O.*
oup has a much lower copy number
japonica subtypes of *O. sativa*, *O.*
). *rufipogon*. Examples of reampli-



1. Hybridization with AA-specific 358-bp repeated fragment. DNA from various rice accessions was digested with *Eco*RI, blot-transferred, and hybridized with plasmid containing 358-bp AA fragment as an insert. Lane 1 = *O. rufipogon* (DN41), 2 = *O. rufipogon* (100968), 3 = *O. rufipogon* (W1655), 4 = *O. longistaminata* (EL34), 5 = *O. officinalis* (W65), 6 = *O. officinalis* (DO 4), 7 = *O. alta* (W17), 8 = *O. coarctata* (W551), 9 = *O. sativa* (58881), 10 = *O. officinalis* (101314), 11 = *O. minuta* (W1344), 12 = *O. officinalis* (W1306), 13 = *O. breviligulata* (WB35), 14 = *O. malampuzhaensis* (W1159), 15 = *O. longistaminata* (EL 15-17), 16 = *O. sativa* (Cigalon).

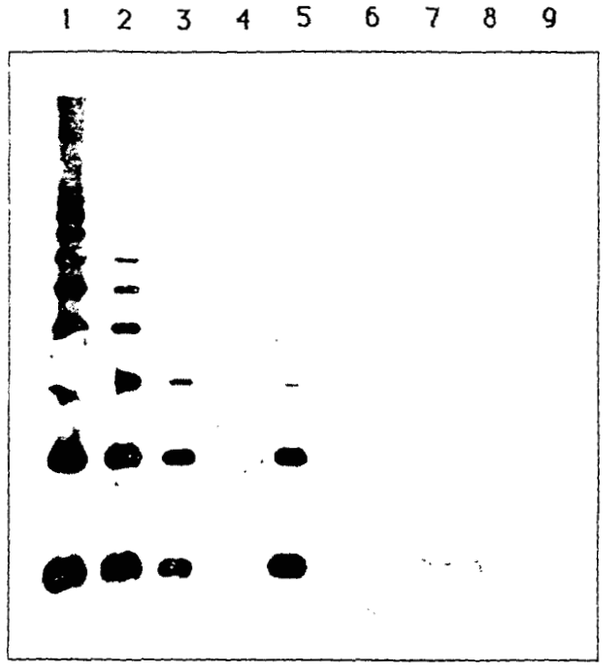
fication of multimers were observed in several high-copy-number *O. rufipogon*, while several *O. longistaminata* have a double ladder pattern, indicating the presence of several subfamilies within these species.

The use of 4-bp cutter enzymes allowed us to recognize two basic patterns discriminating between the japonica and indica types. These two patterns were also observed in different accessions of *O. rufipogon*. This lends additional support to the hypothesis that present-day *O. sativa* have a biphyletic origin (Second 1982).

Three tandemly repeated sequences specific to the CC genome

A similar strategy was used to isolate tandemly repeated sequences from *O. officinalis* accession W1278. Two distinct repeats of 374 and 367 bp as well as another of 193 bp were isolated by molecular cloning and were sequenced (unpubl. data). All these clones, when used as probes on the original *O. officinalis* DNA digested with the enzyme used for cloning, revealed typical ladder patterns (Fig. 2).

The 374- and 367-bp clones cross-hybridize, and sequencing revealed extensive homology (85%). These two sequences are also related to the *O. officinalis*-specific sequence recently described by Zhao et al (1989). Although the two repeated units



2. Hybridization with CC genome-specific 367-bp repeated fragment. DNA from different rice genomes was digested with SphI, electrophoresed, blot-transferred, and hybridized with in vitro-synthesized, labeled probe corresponding to 367-bp insert. Lane 1 = *O. officinalis* (CC), 2 = *O. eichingeri* (CC), 3 = *O. collina* (CC), 4 = *O. punctata* (BB), 5 = *O. minuta* (BBCC), 6 = *O. grandiglumis* (CCDD), 7 = *O. sativa* (AA), 8 = *O. australiensis* (EE), 9 = *O. brachyantha* (FF). Genome type is in parentheses.

clearly belong to two distinct subfamilies, it is not yet clear whether they are physically independent or interspersed in the genome.

The 193-bp repeat is a distinct one and does not cross-hybridize with the 2 others.

When the various genome types were analyzed with these three probes, it became clear that the three sequences are highly specific to the CC genome, since they do not hybridize with any of the others (Fig. 2). However, hybridization with the 193-bp probe is detected with the AA genome when less stringent conditions are used.

The copy numbers of the repeated elements were determined in W1278, from which they were isolated. The 374- and 367-bp repeats are present in 164,000 and 237,000 copies, respectively, and the 193-bp element is repeated approximately 200,000 times. Due to cross-hybridization, the copy numbers of the 374- and 367-bp repeats might be overestimated. Altogether, these 3 repeats account for 10-15% of the *O. officinalis* genome.

A survey of several CC genome accessions also revealed variations in copy number. We found few *O. officinalis* accessions in which both the 374- and 367-bp sequences were absent. We also obtained evidence of rearrangement. One of the most conspicuous situations is with the 193-bp repeat. In all the *O. officinalis* the typical ladder of 193-bp multimers was observed, but in *O. collina* only multimers of 750-bp (tetramers)

6 7 8 9

l fragment. DNA from different rice genomes
id hybridized with in vitro-synthesized, labeled
:s (CC), 2 = *O. eichingeri* (CC), 3 = *O. collina grandiglumis* (CCDD), 7 = *O. sativa* (AA), 8
type is in parentheses.

t yet clear whether they are physically

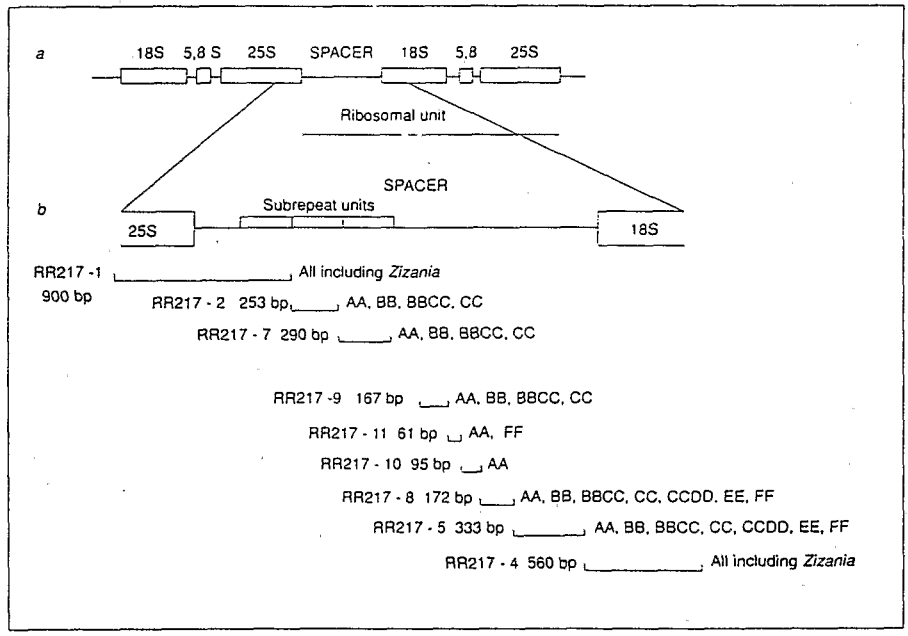
ıot cross-hybridize with the 2 others.
:d with these three probes, it became
to the CC genome, since they do not
, hybridization with the 193-bp probe
ɜent conditions are used.
re determined in W1278, from which
are present in 164,000 and 237,000
)eated approximately 200,000 times.
he 374- and 367-bp repeats might be
ınt for 10-15% of the *O. officinalis*

revealed variations in copy number.
)oth the 374- and 367-bp sequences
ıgement. One of the most conspicu-
officinalis the typical ladder of 193-
ly multimers of 750-bp (tetramers)

were observed, suggesting that a tetramer block has been amplified in this species. In *O. eichingeri*, a complex and irregular pattern was observed, suggesting that blocks of repeats are now dispersed in the genome.

A short sequence within the *O. sativa* ribosomal DNA spacer specific to the AA genome

Since we previously demonstrated that short repeats within the ribosomal DNA (rDNA) spacer are frequently species-specific (Tremousaygue et al 1988), we looked for a similar sequence in rice. Clone RR 217 (Takaiwa et al 1984; provided by F. Takaiwa) contains a full-length rDNA unit. A fragment containing all the intergenic spacer was subcloned and further dissected into nine subclones. Each was used in Southern blot hybridization experiments with DNA representatives of the various *Oryza* genomes and of that of the related genus *Zizania*. As shown in Figure 3, the genomes for these fragments showed a variable degree of homology. In the rDNA unit cloned in RR 217, there are three 260-bp short subrepeats, which account for size variation in the *O. sativa* rDNA (Cordesse et al 1989). These subrepeats hybridize with DNA from the AA, BB, BBCC, and CC genomes but not with DNA from the EE or CCDD genomes. Downstream from this repeat is a short, 95-bp sequence that hybridizes only with the AA genome. It hybridizes equally well with DNA from japonica or indica types.



3. Genome specificity of various rDNA spacer regions: a) general organization of rDNA units, b) enlargement and genome specificity of fragments subcloned from pRR217.

Fate of repeated sequences in allotetraploid genomes

The presence of the CC repeated sequences was examined in the allotetraploids BBCC and CCDD.

So far, the three CC-specific sequences (374, 367, and 193 bp) have been observed in all the BBCC accessions we have examined. Similarly, the 253-bp repeated fragment from the rDNA spacer, although not CC specific, recognizes a homologous sequence in the BBCC rDNA.

As far as the CCDD genome is concerned, the situation is completely different. None of the CC-specific repeated elements was detected in the CCDD genome. Neither did the subrepeat in the rDNA spacer that cross-hybridized with CC DNA recognize a fragment in the CCDD genome.

Discussion

We have isolated a set of repeated sequences from the AA and CC genomes. Hybridization of these cloned sequences to DNA from various cultivated and wild rices revealed that most of the sequences are highly species-specific. From a limited survey of accessions, we obtained evidence for rapid evolution of these elements involving several independent rounds of amplification and divergence as well as homogenization of each subfamily. The 352-bp AA sequence gave us further evidence for a biphyletic origin of *O. sativa* with an early separation of the indica and japonica subtypes in the presumed *O. rufipogon* ancestral population. This adds to the number of molecular methods that discriminate between these subtypes.

The CC-specific sequences should prove invaluable to trace CC introgressions into the AA genome. They have already revealed some diversity among *O. officinalis*, *O. eichingeri*, and *O. collina*, which are the three major CC species, and even within *O. officinalis*. We have also recently isolated a specific dispersed repeat from *O. officinalis* that reveals similar and additional variability. By using these sequences as probes, we can evaluate more precisely the phylogenetic relationships within the CC genome and elucidate the origin of allotetraploids. A surprising result from our study was the complete absence of any of the currently identified specific sequences from the CCDD genome. Several explanations can account for this observation. The CC genome might derive from a CC genome that did not contain the sequences we isolated. Part of the CC genome, and particularly repeated specific sequences, might have been completely excluded from the allotetraploid. Finally, early cytogenetic experiments might have been misleading, and the CCDD genome might result from rapid rearrangement of other rice genomes following their introduction into America, as suggested by Second (1990).

Information on the AA-specific 352-bp tandemly repeated sequence will be published soon (De Kochko et al 1991).

d genomes
is examined in the allotetraploids BBCC

4, 367, and 193 bp) have been observed
Similarly, the 253-bp repeated fragment
fic, recognizes a homologous sequence

l, the situation is completely different.
detected in the CCDD genome. Neither
is-hybridized with CC DNA recognize

ces from the AA and CC genomes.
\ from various cultivated and wild rices
pecies-specific. From a limited survey
evolution of these elements involving
divergence as well as homogenization
ve us further evidence for a biphyletic
ne indica and japonica subtypes in the
This adds to the number of molecular
oes.
luable to trace CC introgressions into
me diversity among *O. officinalis*, *O.*
najor CC species, and even within *O.*
specific dispersed repeat from *O.*
iability. By using these sequences as
ogenetic relationships within the CC
ls. A surprising result from our study
dentified specific sequences from the
ount for this observation. The CC
iot contain the sequences we isolated.
specific sequences, might have been
ially, early cytogenetic experiments
ne might result from rapid rearrange-
uction into America, as suggested by

idemly repeated sequence will be

References cited

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

De Kochko A, Kiefer M C, Cordesse F, Reddy A S, Delseny M (1991) Distribution and organisation of a tandemly repeated 352-bp sequence in the *Oryzae* family. Theor. Appl. Genet. (in press)

Grellet F, Delcasso D, Panabieres F, Delseny M (1986) Organisation and evolution of a higher plant alphoid-like satellite DNA sequence. J. Mol. Biol. 187:495-507.

Maniatis T, Fritsch E F, Sambrook J (1982) Molecular cloning: a laboratory course. Cold Spring Harbor Laboratory Publication, Cold Spring Harbor, USA.

Second G (1982) Origin of the genetic diversity of cultivated rice (*Oryza* spp.): study of the polymorphism scored at 40 isozyme loci. Jpn. J. Genet. 57:25-57.

Second G (1991) Cytoplasmic DNA markers, phylogeny, and systematics in Oryzeae. Pages 475-486 *in* Rice genetics II. International Rice Research Institute, P. O. Box 933, Manila, Philippines.

Takaiwa F, Oono K, Sugiura M (1984) The complete nucleotide sequence of a rice 17S rDNA gene. Nucleic Acids Res. 12:5441-5448.

Tremousaygue D, Grellet F, Delseny M, Delourme R, Renard M (1988) The large spacer of a nuclear ribosomal RNA gene from radish: organisation and use as a probe in rapeseed breeding. Theor. Appl. Genet. 75:298-304.

Wu T Y, Wu R (1987) A new rice repetitive DNA shows sequence 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. 78:201-209.

Notes

Authors' address: F. Cordesse, A.S. Reddy, A. De Kochko, M.C. Kiefer, and M. Delseny, Laboratoire de Physiologie et Biologie Moleculaire Vegetale, Avenue de Villeneuve. URA 565 du CNRS, Universite de Perpignan, 66025 Perpignan-Cedex, France.

Citation information: International Rice Research Institute (1991) Rice genetics II. P.O. Box 933, Manila, Philippines.

Jean-Louis PHAM

Rice Genetics II

Proceedings of the
Second International Rice Genetics Symposium
14-18 May 1990

1991

IRRI
INTERNATIONAL RICE RESEARCH INSTITUTE
P. O. Box 933, Manila 1099, Philippines