Plant Molecular Biology **23**: 243–254, 1993.

PFGE analysis of the rice genome: estimation of fragment sizes, organization of repetitive sequences and relationships between genetic and physical distances

Kun-Sheng Wu [1] and Steven D. Tanksley *
*Department of Plant Breeding and Biometry, Cornell University, 252 Emerson Hall, Ithaca, NY 14853, USA (*author for correspondence); [1] Present address: Du Pont de Numours & Company, Central Research and Development, P.O. Box 80402, Wilmington, DE 19880-0402, USA*

Received 11 February 1993; accepted in revised form 15 June 1993

Key words: high-molecular-weight DNA, pulsed-field gel electrophoresis, genome organization, physical mapping, *Oryza sativa*

Abstract

Pulsed-field gel electrophoresis (PFGE) has been applied to analyze the rice nuclear genome. Probing 56 RFLP probes selected from the 12 rice chromosomes to PFGE blots of nine rare-cutting restriction enzymes revealed that there are relatively high numbers of 'rare-cutting' restriction sites in the rice genome. The average sizes of restriction fragments detected by single-copy probes are smaller than 200 kb for all of the rare-cutting restriction enzymes examined. Sizes of fragments detected by repetitive probes are variable, depending on the probes analyzed. By using PFGE, a tandemly repeated sequence, Os48, was found to be tightly linked to telomeric tandem repeats but not physically linked to r5s genes with which sequence homology had been observed. Relationships between genetic and physical distances have been established for three different chromosomal segments. In these regions 1 cm corresponds to ca. 260 kb on average. Analysis of a cluster of RFLP markers on chromosome 3 revealed that genetically clustered RFLP markers are also physically closely linked, suggesting that clustering of genetic markers may result in part from uneven distribution of single-copy sequences.

Introduction

Genetic mapping based on restriction fragment length polymorphisms (RFLP) in rice has rapidly progressed in recent years. A high-density map with more than 600 markers has been constructed in this laboratory [25, Causse *et al.* in preparation]. RFLP markers from this map have been successfully applied to tag genes controlling disease [38], insect [26] resistances and other important agronomic traits [1]. While RFLP have provided sufficient coverage of some chromosomal regions containing genes of agronomic interest, the physical relations of these markers remain untested. In order to utilize the high-density RFLP map in the study to genome organization and in the process of clone agronomically important genes by chromosome walking, it is essential to convert the genetic (linkage) map to a physical (base pair) map.

Pulsed-field gel electrophoresis, developed first by Schwartz and Cantor [30] and improved by others [7–9, 11], has allowed the separation of large DNA fragments up to 10 Mb [15]. When combined with rare-cutting restriction enzymes and Southern blotting procedures [34], this technique provides a powerful tool for physical mapping and has been successfully applied to construct complete physical maps of several relatively small genomes such as that of *Escherichia coli* [32], *Schizosaccharomyces pombe* [15] and *Saccharomyces cerevisiae* [24] as well as of large regions of the mammalian genome [6]. This technique has been recently adapted for use in a number of plant species [10, 13, 17, 35]. The use of PFGE analysis to study physical relationships between RFLP markers and agronomically important genes and to characterize repeated DNA sequences in plants has also been reported [14, 18, 19, 31]. These studies have demonstrated the feasibility of using PFGE to analyze genome organization in plants.

Some technical aspects of using PFGE in rice studies have also been described [33]. However, the source of DNA used was from protoplasts derived from cell suspension culture. Since somaclonal variations due to DNA rearrangements and mutations have been frequently observed in plant cell culture [5, 23, 28], the results obtained by using tissue culture materials may not represent the situation in intact plants. Therefore, it is important to use DNA isolated directly from native plant tissues, particularly when a physical map is being constructed. In addition, cell suspension culture may not be always available for a particular variety containing genes of interest. For this and other reasons, we developed a procedure for isolating high molecular weight DNA directly from rice plants and have used PFGE to determine the size range of fragments generated by different rare-cutting restriction enzymes. The relationship between genetic and physical distances has been determined for three different segments of the rice genome and some known repetitive sequences have also been investigated. Physical analysis of a cluster of genetic markers revealed that clustering was caused by tight physical linkage rather than recombination suppression.

Materials and methods

DNA preparation

Two rice cultivars, IR 36 and IS1188, both from the International Rice Research Institute at Los Baños, the Philippines, were used in this study. IS1188 is an isogenic line developed by G.S. Khush, bearing a gene for bacterial blight resistance, *Xa21*, in the IR24 genetic background.

Two- to three-month-old rice plants grown in a greenhouse served as a source of tissue for isolation of high-molecular-weight DNA. Before harvesting tissue, the leaf blades were cut off above the auricles and the plants were subjected to dark treatment for three days. This is essential to reduce starch granules in the cells. The whitish portions of the leaf sheaths were submerged in protoplast buffer (16% mannitol, 20 mM MES, 1% dextran sulfate, pH 5.6) and cut into 2–3 mm cross sections. The sections were transferred to enzyme solution containing 1.5% Cellulase Onozuka RS (Yakult, Tokyo, Japan), 0.15% pectolyase Y23 (Seishin, Tokyo, Japan) in protoplast buffer and incubated under vacuum for about 20 min and then under normal pressure for 4 more hours at room temperature. The tissue/enzyme mixture was filtered through one layer of Miracloth, 80 μM and 30 μM sieves sequentially. Two volumes of washing solution (230 mM NaCl, 166 mM $CaCl_2$, and 7 mM KCl pH 5.6) were added to the mixture before centrifugation at $200 \times g$ for 10 min to pellet the protoplasts. Two additional washings were performed. A final concentration of approximately 3×10^7 protoplasts/ml was mixed with an equal amount of low melting agarose in washing solution. This protocol routinely yields 2×10^7 protoplasts per gram of fresh whitish tissues. Solidified plugs were incubated in ESP solution (0.5 M EDTA pH 9.2, 1% sarkosyl, 1 mg/ml proteinase K) with one change for 60–70 h. DNA plugs were kept in the same solution in 4 °C until use.

Pulsed-field gel electrophoresis

DNA plugs were washed thoroughly as described by Ganal and Tanksley [17]. Each DNA plug (ca. 2 μg) was equilibrated with restriction enzyme buffer for one hour on ice. The buffer was then removed and replaced by 1 × reaction buffer containing 4 mM spermidine, 1 mM DTT and 10 to 30 units of restriction enzymes. The DNA plug was placed on ice at least 40 min before incubating at temperatures as recommended by the manufacturer for 6 h to overnight. For partial digestion, DNA plugs with restriction buffers and enzymes were incubated on ice at 4 °C overnight and then transferred to the temperature recommended by the manufacturer for 10 to 20 min. The reactions were stopped by adding EDTA to a final concentration of 50 mM and SDS to 0.2%. PFGE was carried out at 4 °C in 0.5 × TBE buffer in a CHEF gel apparatus as described by Chu *et al.* [11]. A pulse time of 15 s was used for separating 20–500 kb fragments in 1.5% agarose while a 2 min pulse time was employed for separating 50–1000 kb DNA fragments in 1% agarose gels. After electrophoresis, gels were exposed to short wavelength UV light for 5 min in the presence of ethidium bromide. Blotting to Hybond N + membranes (Amersham) was carried out in 0.4 M NaOH for 20–24 h. The membranes were neutralized in 2 × SSC for 10 min before use in hybridization. Hybridization procedures were basically the same as described by Bernatzky and Tanksley [4] and McCouch *et al.* [25].

DNA probes

DNA probes used in this study were from three different libraries. RG clones were from a rice random genomic library constructed in this laboratory [25]. RZ and CDO clones were from rice and oat cDNA libraries respectively. Inserts of RZ and CDO clones amplified by PCR directly from phagemids were used in hybridizations. For most of the RG clones inserts amplified by PCR using M13 or T7/Sp6 primers were used as probes, while in some cases clones with vector sequences prepared from plasmid minipreps were used directly. Probe Os48 was kindly provided by R. Wu [37] and the probe for 5S ribosomal genes was provided by W. Gerlach [21]. The inserts of these probes amplified by PCR were used in hybridization. An oligonucleotide (Tel) homologous to the telomeric sequence of *Arabidopsis* [29] was used as probe for rice telomeres. Probes were labeled by the random priming method of Feinberg and Vogelstein [16]. To reuse the filters, the probes were stripped sequentially with 0.12 M NaOH; 0.1 M Tris-HCl, 0.1 × SSC and 0.1% SDS; and 0.1 × SSC and 0.1% SDS, each for 20 min. Where necessary, the completion of stripping was monitored by reexposing the filters.

Results

Fragment sizes generated by different restriction enzymes as detected by single-copy probes

To identify suitable restriction enzymes for physical mapping in rice, we have tested a wide range of rare-cutting restriction enzymes, including *Bss* HII, *Csp* I, *Mlu* I, *Nar* I, *Not* I, *Nru* I, *Rsr* II, *Sfi* I, *Sma* I. At least two probes from each of the twelve chromosomes were selected and used to hybridize to PFGE blots containing DNA digested with these enzymes. Figure 1 shows an example of the hybridizations. The sizes of DNA fragments generated by these enzymes and detected with single-copy probes are summarized in Table 1. For most of the enzyme digests, each probe revealed a single major band. In some cases, single-copy probes recognized multiple bands with similar intensity. These could arise as a result of partial cleavage caused by methylation at CpG dinucleotides.

Average fragment sizes of different restriction enzymes are summarized in Table 2. All of the 'rare-cutting' restriction enzymes used in this study cut fairly frequently in rice DNA. No single enzyme gave an average fragment size larger than 200 kb. *Not* I is a methylation-sensitive enzyme containing CpGs and CpXpGs in its eight base pair recognition site. This enzyme usually creates

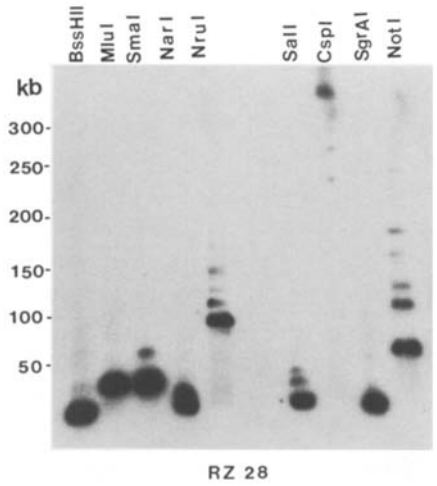

Fig. 1. Pulsed-field gel blots showing fragment sizes detected by a single-copy probe, RZ28. DNA from rice variety IS1188 was digested with enzymes indicated, fractionated by pulsed field gel electrophoresis, transferred to a Hybond N + membrane and probed with RZ28. Lambda DNA concatamers served as size standard.

large fragments in other plant species [17, 19]. The average fragment size of *Not* I in rice however is only 105 kb. The high frequency of occurrence of *Not* I restriction sites is clearly demonstrated in Fig. 1, where the RZ 28 probe hybridizes to a major *Not* I band of 60 kb and more than seven partial discrete bands over a distance of 250 kb. *Rsr* II, *Sfi* I and *Csp* I gave relatively large DNA fragments in rice as identified by single-copy probes, and *Bss* HII identified the smallest fragments on average among all the enzymes used even though the recognition sequence of this enzyme contains all G + C base pairs including two CpG dinucleotides. The average fragment sizes detected by genomic probes (RG clones) were slightly larger than those detected by cDNA probes (RZ and CDO clones). However, the variation was not significant as revealed by a *t* test. The fact that most of the methylation sensitive enzymes with CpGs and CpXpGs on their recognition sequences failed to yield larger fragment sizes implied either a lower overall methylation level or a higher frequency of CpG and CpXpG nucleotides or both in the rice genome relative to other organisms.

Hybridization patterns of multiple-copy and repetitive probes

The observation that most single-copy probes detected small fragments prompted us to investigate whether larger fragments ever exist in rare-cutting enzyme digests and whether multiple-copy or repetitive probes could recognize them. A number of multiple-copy and repetitive probes were selected based on previous surveys. The results of these experiments were variable, depending on the probes analyzed. Basically, three types of hybridization patterns were observed with repetitive probes: smears, multiple discrete bands and single bands. For example, with probes RG641, RG485 and RZ51, only smears were detected in all the rare-cutting enzyme digests indicating an interspersed distribution of the probe sequences (data not shown), while with probes Os48, Tel and r5s, multiple discrete bands were recognized (Fig. 2), indicating clustering of these probe sequences. With probes RZ163 and RG665, multiple bands were observed on conventional gels with methylation insensitive enzymes such as *Eco* RV, *Hin*d III, *Sca* I and *Xba* I, but a single fragment was detected on PFGE blots with rare-cutting restriction enzymes (data not shown). It was interesting to note that even repetitive probes recognized only relatively small fragments. Most of fragments detected by repetitive probes were smaller than 800 kb in complete digests using all the enzymes tested. Fragments larger than 1 Mb were rarely observed except for some rDNA fragments. This indicates that tandemly repeated blocks of the rice genome are relatively short compared to other plant species [17, 31].

Linkage of a tandemly repeated sequence, Os48, to telomeres

Os48 is a probe containing tandemly repetitive sequences specific to the AA genome of rice [37, 39]. The copy number has been estimated to be 2000 to 4000 per haploid genome. A 5S RNA pseudogenic origin of this sequence has been proposed based on the observation of sequence ho-

Table 1. Summary of fragment sizes detected by single-copy RFLP probes with nine restriction enzymes.

Probe	*Bss* HI	*Csp* I	*Mlu* I	*Nar* I	*Not* I	*Nru* I	*Rsr* I	*Sfi* I	*Sma* I
CDO962	–	95	50	40	130	60	–	–	75
RZ274	50	480	75	70	*330*	90	–	330	200
					70			*80*	
RG146	45	100	*100*	80	125	125	–	–	80
			40						50
RG229	–	–	170	–	–	–	–	300	170
RG317	120	–	150	120	75	150	–	*250*	220
RG345	40	–	80	–	50	170	–	*150*	120
								20	
RG492	<20	350	130	35	–	35	*470*	150	120
							350		
RG690	–	–	100	75	100	–	–	250	–
RG220	<40	–	50	<40	100	–	–	250	–
RG109	–	–	–	<20	*190*	80	90	120	40
CDO251	50	–	–	<20	–	70	*250*	*240*	45
						<10	*220*	200	
							180		
RG73	<40	–	80	<30	100	40	–	50	40
CDO196	35	–	65	<30	*250*	–	180	200	50
					140				
					80				
					35				
RG152	35	–	65	<30	*250*	40	180	50	50
					140				
					80				
					35				
CDO524	<30	–	30	<20	–	–	40	270	<20
RZ173	–	<60	<30	<30	70	–	<40	–	<20
RZ58	–	100	<30	<40	50	<40	–	–	<40
RG778	80	–	–	20	125	80	140	200	50
CDO497	75	–	100	–	*80*	80	*500*	*340*	40
					30		*400*	290	
							290		
RG650	<10	–	–	<10	*75*	80	*200*	*270*	15
						10	150		
RZ264	75	–	*120*	130	40	12	120	*340*	*55*
			80					*230*	35
								150	
RZ66	35	–	80	40	*250*	–	75	170	120
RZ28	<20	230	40	<20	*250*	90	275	40	40
					230				
					175				
					105				
					65				
RG757	40	390	550	110	*250*	40	440	–	90
					145				
CDO590	80	330	180	120	*250*	<30	*640*	250	200
					240		440	210	
							320		
CDO94	–	–	125	100	170	80	–	260	40
RZ163	–	230	110	50	–	150	–	–	*180*
RG665	50	–	70	40	70	–	–	–	<40

Note: Main partial bands are underlined while some minor partial are not listed. Fragment sizes of some other enzymes used only occasionally are not listed.

Table 2. Average fragment sizes (in kb) detected with probes from different libraries. The number in the parenthesis indicates the number of probes analyzed. In case of more than two bands with similar intensity, only the smallest band is used to estimate the average fragment size for the enzyme.

Enzyme	RG clones	RZ clones	CDO clones	All probes
Bss HII	46 (15)	51 (5)	38 (10)	43 (30)
Csp I	216 (6)	180 (9)	147 (4)	188 (19)
Mlu I	95 (15)	68 (11)	72 (10)	111 (36)
Nar I	40 (18)	55 (10)	46 (10)	46 (38)
Not I	116 (16)	58 (5)	113 (8)	105 (29)
Nru I	105 (13)	76 (9)	51 (8)	82 (30)
Rsr II	207 (9)	188 (5)	196 (9)	199 (23)
Sfi I	184 (15)	194 (5)	206 (8)	199 (28)
Sma I	79 (17)	75 (11)	59 (12)	79 (17)
Mean	116	98	96	109

Paired t test

RG clones/RZ clones	$t = 1.97$ ns
RG clones/CDO clones	$t = 1.88$ ns
RZ clones/CDO clones	$t = 0.22$ ns

ns = not significant.

mology between Os48 and wheat r5s genes [37]. However, in another study [12], a repeated element, pOs7, showing high homology (93%) to Os48 was also cloned. Sequence homologies of pOs7 with human alpha satellite DNA and maize knob heterochromatin specific repeats suggested that both Os48 and pOs7 belong to the alphoid satellite DNA family. To investigate the genome organization and possible roles of this major repeated sequence, PFGE analysis was thus carried out to establish the physical relationships between this repeated sequence and other major repeated sequences. When probe Os48 was used to hybridize to DNA blots of different rare-cutting restriction enzyme digests, discrete bands rather than a smear were observed, indicating that the repeat family is organized in clusters rather than dispersed throughout the genome. Each discrete band may represent a different locus on the rice chromosomes. The sizes of the fragments detected ranged from about 50 to 1000 kb. Hybridization patterns of the r5s probe, however, were completely different from those of Os48 (Fig. 2). No shared major fragment with Os48 was found with 11 restriction enzymes. This suggests that Os48 and r5s genes are not located in the same sites in the genome.

Experiments were further carried out to determine whether Os48 has an organization similar to satellite DNA. Satellite DNA has been found to be tightly linked to telomeres in tomato [18] and

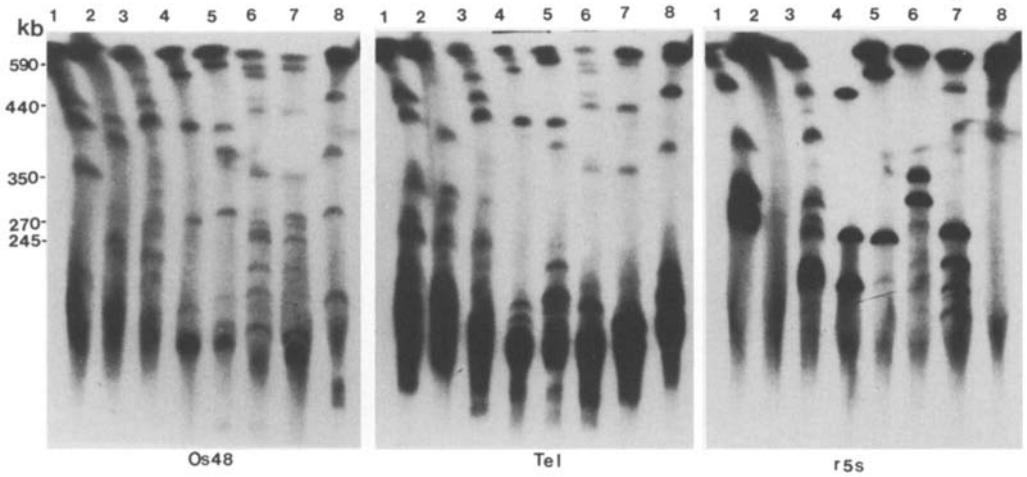

Fig. 2. Hybridization patterns of some known repetitive sequences. High-molecular-weight DNA was digested with *Not* I (1), *Rsr* II (2), *Sfi* I (3), *Bss* HII (4), *Sma* I (5), *Nar* I (6), *Sal* I (7), *Nru* I (8) and then separated by pulsed-field gel electrophoresis. The same filter was hybridized sequentially to each of the three probes, pOs48, r5s and Tel but for comparison purpose Os48 and Tel are put side by side in this figure.

rye [3]. To test this possibility, an oligonucleotide corresponding to the telomere tandem repeats of *Arabidopsis* $(TTTAGGG)_n$ was probed to the same filter after r5s probes had been stripped. The resulting hybridization patterns were very similar to those using the Os48 as probe (Fig. 2). A total of 27 identical fragments were found for eleven enzymes tested. In fact, most fragments larger than 300 kb were in common with these two probes. Some common fragments smaller than 50 kb were also observed, suggesting tight linkage between Os48 and telomere tandem repeats. This result suggests that Os48 be a subtelomere satellite-like DNA in the rice genome.

Relationship between genetic and physical distances

In our early attempts complete digestion was made with *Not* I, *Sfi* I, *Csp* I and *Rsr* II and the resulting DNA blots were probed with genetically tightly linked markers. However, no physical linkage was observed in all the markers tested even though some of them were located at less than 1 cm apart. It is obvious that 'rare-cutting' restriction enzymes cut so frequently in the rice DNA that fragments generated by this method are not large enough to bridge a centimorgan. To overcome this problem, we attempted a series of partial digestions with a limited amount of restriction enzymes and various incubation conditions. This enabled us to establish physical linkages between three pairs of genetically linked markers as shown in Table 3.

RG757 and CDO590 are 0.9 cM apart on chromosome 9. When these two markers were used to probe PFGE blots containing DNA partially digested with the enzymes listed in Table 1, both probes detected a partial *Not* I fragment of 250 kb (Fig. 3A). Physical linkage of these two probes was also indicated by the observation that a *Rsr* II band of 440 kb detected by RG757 was identical to that recognized by CDO590 (Fig. 3B). The *Not* I fragment in common between RG757 and CDO590 was slightly different in size as detected in gels run under different conditions (compare Figs. 3A and 3B). The fragment observed on Fig. 3A with 1.5% agarose seems to be better resolved. Thus, a maximum physical distance of 250 kb was assigned to 0.9 map units at this genetic interval RZ576 and RZ497 are located 0.9 cM apart on chromosome 3. In complete digests with a number of enzymes, no common fragment was found between these two probes.

Table 3. Physical linkage analysis of selected RFLP markers.

Genetic interval	Chromosome location	Map units (cM)	Physically linked fragments
RG757-CDO590	9	0.9	*Not* I (250 kb); *Rsr* II (440 kb)
RZ576-RZ497	3	0.9	*Sfi* I (290 kb)
CDO66-RZ28	8	0.9	*Not* I (230 kb, 250 kb)
CDO196-RG152	2	0.0	*Bss* HII (35 kb); *Mlu* I (65 kb); *Nar* I (30 kb); *Not* I (35 kb); *Rsr* II (180 kb); *Sma* I (50 kb)
RZ22-RZ25	3	0.0	*Csp* I (100 kb); *Mlu* I (50 kb); *Nar* I (60 kb); *Nru* I (65 kb); *Sma* I (45 kb)
RZ517-RZ399-RZ16	3	0.0	*Sfi* I (230 kb)
RG745-RZ574-RG227	3	0.0	*Csp* I (290 kb)
RZ173-RZ58	2	1.0	none
CDO962-RZ274	1	1.1	none
RZ274-RG146	1	1.0	none
RG109-RG727	1	1.1	none
RG470-CDO89	5	1.1	none
CDO78-RG778	6	1.3	none
CDO251-RG109	1	0.0	none
RG 4-RZ 264	7	0.0	none

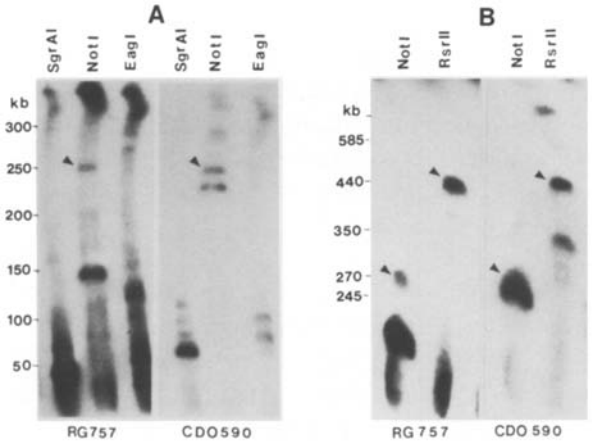

Fig. 3. PFGE blot showing common fragments detected by RG757 and CDO590. DNA from rice variety IS1188 was digested with enzymes indicated and fractionated by pulsed-field gel electrophoresis. The resulting filter was hybridized sequentially with probes RG757 and CDO590. Arrows indicate the common fragments between the two probes. Molecular weight standards are yeast chromosomes and lambda DNA concatamers.

However, in a partial digest with *Sfi* I, more than five partial bands could be identified and a partial band of 290 kb was identical between both probe sequences (Fig. 4). Thus, the largest physical distance between these two markers is about 290 kb, and one centimorgan is equivalent to at most 320 kb. Similarly, the physical distance was established between RZ28 and RZ66 which are located on chromosome 8 with a genetic distance of 0.9 cM. Figure 5 shows that RZ28 identifies up to twelve *Not* I partial digest fragments while RZ66 identifies up to five partial digest bands. Two clear partial digest bands of 230 and 250 kb were shared by RZ28 and RZ66, indicating that the physical distance between these two probes is ca. 230 kb. Thus, 1 cM corresponds to about 250 kb.

Table 3 lists the genetic intervals which have been analyzed by PFGE. Besides the three genetic segments mentioned above, physical linkages were also observed between CDO196 and RG152 and between RZ22 and RZ25 as well as among a cluster of markers on chromosome 3 (see next section for explanation). However, in six genetic intervals with genetic distances of approximately one centimorgan and two genetic intervals with zero recombination rate (CDO250-RG109 and RG4-RZ264), no physical linkages were found (Tables 1 and 3). Although the genetic recombinants between these tested markers were confirmed by scoring on a second set of mapping

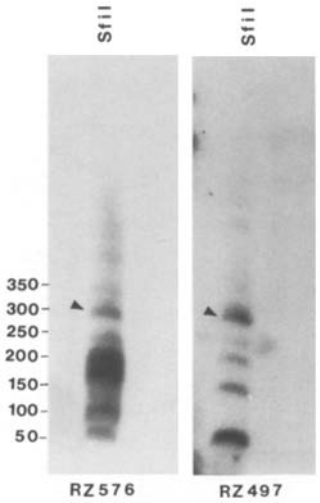

Fig. 4. Physical linkage of RZ576 and RZ497 by PFGE. High-molecular-weight DNA was partially digested with *Sfi* I. The gel was blotted onto a nylon membrane and hybridized successively to RZ576 and RZ497. Arrowheads indicate the common fragments.

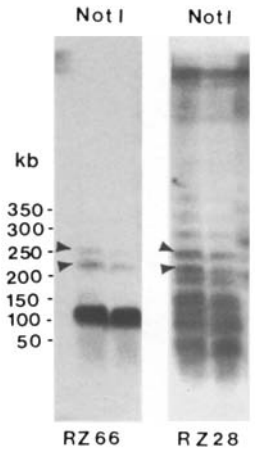

Fig. 5. Physical linkage between RZ66 and RZ28 detected by PFGE. High-molecular-weight DNA was partially digested with *Not* I and probed with RZ66 and RZ28 successively. Arrows indicate common fragments between the two probes. Molecular size standards in kb are shown to the left.

filters, the genetic distances in these intervals could be underestimated due to the relatively small population size (113 BC_1 progenies). A larger mapping population is needed to obtain a more accurate measurement of genetic distances between these markers.

Physical analysis of a cluster of markers on chromosome 3

In the rice genetic map, RFLP markers are not evenly distributed along the genetic linkage map (Causse *et al.*, in preparation). Markers are clustered in some regions while underrepresented in other regions. To investigate whether genetic recombination was suppressed in regions where RFLP markers are highly clustered, a region of chromosome 3 was selected to compare the genetic and physical linkage relationships. In the selected region, nine RFLP markers (RZ16, RG745, RZ399, RZ678, RZ574, RZ517, RG227, CDO33X and CDO260) are clustered with no recombination (in 113 BC plants) between any of them. The genetic linkage data of these markers were conformed by probing and scoring the progenies twice to make sure that no recombinants were missed while the physical relationship of six of these markers (excluding RZ678 CDO33X and CDO260) was analyzed by PFGE using two sets of filters. Figure 6 shows a PFGE Southern sequentially probed with RZ 517, RZ399 and RZ 16. It is interesting to note that all the probes tested showed tight physical linkage with a physical distance less than 300 kb (Table 3). This suggests that clusters of RFLP markers in the rice map be caused by close physical proximity rather than by recombination suppression which might be expected in the interspecific population used to construct the genetic linkage map.

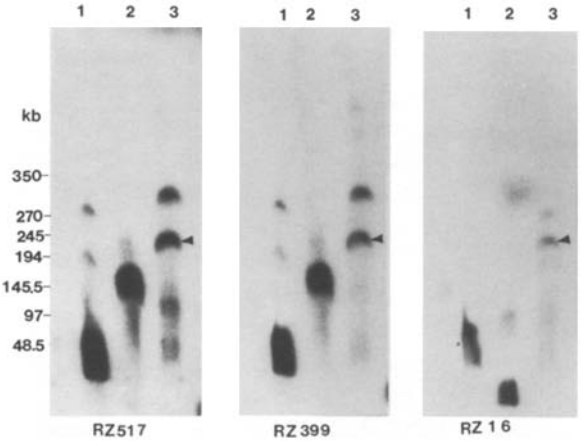

Fig. 6. Physical linkages of some genetically clustered RFLP markers. High-molecular-weight DNA of IS1188 was digested with *Not* I (1), *Rsr* II (2) and *Sfi* I (3) and probed sequentially by RZ517, RZ399 and RZ16. Arrows indicate the common fragments among the probes. Molecular size standards are yeast chromosomes and lambda DNA concatamers.

Discussion

We have developed a procedure to isolate high-molecular-weight DNA directly from intact plants and have applied PFGE to analyze the rice nuclear genome. Large scale physical mapping by PFGE is an important step towards understanding genome structure and function. To apply PFGE to genome analysis, one needs rare cutting restriction enzymes which are able to generate large fragments of DNA and sufficient RFLP markers as probes to identify the resulting fragments. Naturally occurring rare-cutting restriction enzymes usually have octameric recognition sites (such as *Not* I and *Sfi* I) or are hexameric enzymes which contain the dinucleotide sequence CpG and CpXpG. These DNA motifs are often underrepresented in the genomes of higher plants and frequently methylated at the cytosine residue [22]. Thus these enzymes produce large fragments in many species [17, 20, 31]. The results obtained with rice, however, are quite different from mammals and other plant species. The average fragment sizes for most of the rare-cutting enzymes are less than 200 kb when single-copy clones are used as probes. This apparently indicates a higher G + C content and low cytosine methylation, especially low 5-methyl cytosine at CpG or CpxpG in the genome. Rice has a total G + C content of 42% and 5-methyl cytosine content of 18.8% [27]. Combination of the G + C content and methylation will give rise to very small fragment sizes in this species. This can be dem-

onstrated by comparing rice with tomato. Tomato has a G + C content of 36% and 5-methyl cytosine content of 21% [27]. Based on the G + C content alone, an eight base pair recognition restriction enzyme, *Not* I, for example, would be expected to give an average fragment size of 910 kb in tomato while only 220 kb in rice. The fragment size of tomato is near five times larger than that of rice. Predicted fragment sizes of various rare-cutting restriction enzymes based on occurrences of mono-, di-, and tri-nucleotide sequences from sequenced rice DNA data have been given by Sobral *et al.* [33]. Our results showed a similar tendency. Since all the 'rare-cutting' restriction enzymes commercially available cut fairly frequently in rice and average fragment sizes detected by single-copy probes are not large enough to bridge a centimorgan unit, it will be very difficult to construct large-scale physical maps using the PFGE approach in rice.

Os48 (also referred to as RC48) is a member of a tandemly repeated family, specific for the AA genome of rice. This probe showed homology to the wheat phenylalanine tRNA and 5S ribosomal RNA sequence [37]. PFGE analysis showed however no physical linkage between Os48 and r5s genes. Instead, tight physical linkage was found between Os48 and telomere tandem repeats. Recently, *in situ* hybridization has been carried out to map Os48 on rice metaphase chromosomes [36]. It was found that Os48 hybridized to eight chromosomes, all at the distal ends. Thus, both results from PFGE and *in situ* are in good agreement. In addition, r5s genes have been mapped to an internal region of chromosome 11 using PFGE with the restriction enzyme *Bss* HII (K.S. Wu, unpublished data) and only one locus was found in the rice mapping population. No Os48 sequence is known to be present at this chromosomal region. Physical linkage of Os48 to telomeres implies that this sequence, like pOs7, belongs to satellite DNA as suggested by De Kochko *et al.* [12]. Our results also suggest that organization of rice chromosomes may resemble that of tomato because satellite DNA is found to be tightly linked to the telomere tandem repeats in both species.

Establishment of the relationship between genetic and physical distances in the rice genome will provide general guidance for work on chromosome walking as well as on marker-assisted selection in breeding practice. The haploid genome of rice contains 450 Mb DNA [2] which corresponds to about 1400 maps units in the current RFLP map (Causse *et al.*, in preparation). This means that every centimorgan is equivalent to ca. 320 kb on averge. In this study, physical linkage was observed for three pairs of genetically closely linked probes. The maximum physical distance between RG757 and CDO590 is 250 kb which corresponds to 0.9 cM. One centimorgan unit thus equals 280 kb. In the genetic interval between RZ576 and RZ497, 0.9 cM corresponds to 290 kb, giving 1 cM equivalent to 320 kb. In both genetic segments, the observed physical distances were close to what was predicted. A relatively short physical distance was assigned to the genetic intervals between RZ28 and RZ66 where one centimorgan corresponded to only 250 kb.

Among the genetic intervals tested, no common restriction fragment was observed between CDO251 and RG109, and between RG4 and RZ264 each with zero genetic recombination rates nor between six pairs of markers with approximate 1% of recombination (Table 3). This means that the physical distances corresponding to 1 cM in these genetic segments could be larger than what were observed for the three genetic intervals mentioned above. However, underestimate of genetic distances for some of these markers due to the relatively small mapping population size could also lead to the same observation. To obtain a more accurate measurement of genetic distances, a larger mapping population is needed. However, since the mapping population currently employed was from an interspecific hybridization, and the original hybrid seeds had been exhausted, we were unable to extend our mapping population for such purpose.

Acknowledgements

We are grateful to Drs Martin Ganal and Pamela Ronald for critical reading of the manuscript.

This work was supported by a Rockefeller Foundation Fellowship to K.S.W. and a grant from Rockefeller Foundation to S.D.T.

References

1. Ahn SN, Bollich CN, Tanksley SD: RFLP tagging of a gene for aroma in rice. Theor Appl Genet 84: 825–828 (1992).
2. Arumuganathan K, Earle ED: Nuclear DNA content of some important plant species. Plant Mol Biol Rep 9: 208–218 (1991).
3. Bedbrook JR, Jones J, O'Dell M, Thompson RD, Flavell RB: A molecular description of telomeric hereochromatin in Secale species. Cell 19: 545–560 (1980).
4. Bernatzky R, Tanksley SD: Methods for detection of single or low copy sequences in tomato on Southern blots. Plant Mol Biol Rep 4: 37–41 (1986).
5. Brown PTH, Kyozuka J, Sukekiyo Y, Kimura Y, Shimamto K, Lorz H: Molecular changes in protoplasts-derived rice plants. Mol Gen Genet 223: 324–328 (1991).
6. Burmeister M, Monacco AP, Gillard EF, Van Ommen LM, Lehrach H: A 10-Megabase map of human Xp21, including the Duchenne muscular dystrophy gene. Genomics 2: 189–202 (1988).
7. Cantor CR, Warburton P, Smith CL, Gaal A: Voltage ramp pulsed field gel electrophoresis separation of large DNA molecules. In: Dunn M (ed) Electrophoresis '86, pp. 161–171. VCH Publishers, Weinheim (1986).
8. Carle GF, Olson MV: Separation of chromosomal DNA molecules from yeast by orthogonal-field-alternation gel electrophoresis. Nucl Acids Res 12: 5647–5664 (1984).
9. Carle GF, Frank M, Olson M: Electrophoretic separation of large DNA molecules by periodic inversion of the electric field. Science 232: 65–68 (1986).
10. Cheung WY, Gale MD: The isolation of high molecular weight DNA from wheat, barley and rye for analysis by pulsed-field electrophoresis. Plant Mol Biol 14: 881–888 (1990).
11. Chu G, Vollrath D, Davis RW: Separation of large DNA molecules by contour-clamped homogeneous electric fields. Science 234: 1582–1585 (1986).
12. De Kochko A, Kiefer MC, Cordesse F, Reddy AS, Delseny M: Distribution and organization of a tandemly repeated 352-bp sequence in the *oryza* family. Theor Appl Genet 82: 57–64 (1991).
13. Devos KM, Vercruysse-Dewitte D: Preparation of plant DNA for separation by pulsed field gel electrophoresis. Electrophoresis 10: 267–268 (1989).
14. Ellis THN, Lee D, Thomas CM, Simpson PR, Cleary WG, Newman M-A, Burcham KWG: 5s rRNA genes in *Pisum*: sequence, long range and chromosomal organization. Mol Gen Genet 214: 333–342 (1988).
15. Fan JB, Chikashige Y, Smith CL, Niwa O, Yanagida M, Cantor CR: Construction of a *Not* I restriction map of the fission yeast *Schizosaccharomyces pombe* genome. Nucl Acids Res 17: 2801–2818 (1988).
16. Feiberg AP, Vogelstein B: A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132: 6–13 (1983).
17. Ganal MW, Tanksley SD: Analysis of tomato DNA by pulsed field gel electrophoresis. Plant Mol Biol Rep 7: 17–28 (1989).
18. Ganal MW, Lapitan NLV, Tanksley SD: Macrostructure of the tomato telomeres. Plant Cell 3: 87–94 (1991).
19. Ganal MW, Bonierbale MW, Roder MS, Park WD, Tanksley SD: Genetic and physical mapping of the patatin genes in potato and tomato. Mol Gen Genet 225: 501–509 (1991).
20. Ganal MW, Young ND, Tanksley SD: Pulsed field gel electrophoresis and physical mapping of large DNA fragments in the *Tm-2a* region of chromosome 9 in tomato. Mol Gen Genet 215: 395–400 (1989).
21. Gerlach WL, Dyer TA: Sequence organization of the repeating units in the nucleus of wheat which contain 5s rRNA genes. Nucl Acids Res 8: 1851–1865 (1980).
22. Gruenbaum Y, Naveh-Many T, Cedar H, Razin A: Sequence specificity of methylation in higher plant DNA. Nature 292: 860–863 (1981).
23. Larkin J, Scowcroft W: Somoclonal variation – a novel source of variability from cell cultures for plant improvement. Theor Appl Genet 60: 197–214 (1981).
24. Link AJ, Olson M: Physical map of the *Saccharomyces cerevisiae* genome at 110-Kilobase resolution. Genetics 127: 681–698 (1991).
25. McCouch SR, Kochert G, Yu ZH, Wang ZY, Khush GS, Coffman WR, Tanksley SD: Molecular mapping of rice chromosomes. Theor Appl Genet 76: 815–829 (1988).
26. McCouch SR, Tanksley SD: Development of and use of restriction fragment length polymorphism in rice breeding and genetics. In: Khush GS, Toennissen G (eds) Rice Biotechnology, pp. 109–133. C.A.B International, Wallingford, UK/International Rice Research Institute, Manilla (1991).
27. Messeguer R, Ganal MW, Steffens JC, Tanksley SD: Characterization of the level, target sites and inheritance of cytosine methylation in tomato nuclear DNA. Plant Mol Biol 16: 753–770 (1991).
28. Muller E, Brown PTH, Hartke S, Lorz H: DNA variation in tissue-culture-derived rice plants. Theor Appl Genet 80: 673–679 (1990).
29. Richards EJ, Ausubel FM: Isolation of higher eukaryotic telomere from *Arabidopsis thaliana*. Cell 53: 127–136 (1988).
30. Schwartz DC, Cantor CR: Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. Cell 37: 67–75 (1984).
31. Siedler H, Graner A: Construction of physical maps of the *Hor1* locus of two barley cultivars by pulsed field gel electrophoresis. Mol Gen Genet 226: 177–181 (1991).

32. Smith CL, Econome JG, Schutt A, Klco S, Cantor CR: A physical map of the *Escherichia coli* K12 genome. Science 236: 1448–1453 (1987).
33. Sobral WS, Honeycutt RJ, Atherly AG, McClelland M: Analysis of rice (*Oryza sativa* L.) genome using pulsed field gel electrophoresis and rare-cutting restriction endonucleases. Plant Mol Biol Rep 8: 253–275 (1991).
34. Southern EM: Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98: 503–517 (1975).
35. Van Daelen RAJ, Jonkers JJ, Zabel P: Preparation of megabase-sized tomato DNA and separation of large restriction fragments by field inversion gel electrophoresis (FIGE). Plant Mol Biol 12: 341–352 (1989).
36. Wu HK, Chung MC, Wu TY, Ning CN, Wu R: Localization of specific repetitive DNA sequences in individual rice chromosomes. Chromosoma 100: 330–338 (1991).
37. Wu T, Wu R: A new rice repetitive DNA shows sequence homology to both 5s RNA and tRNA. Nucl Acids Res 15: 5913–5923 (1987).
38. Yu ZH, Mackill DJ, Bonman JM, Tanksley SD: Tagging genes for blast resistance in rice via linkage to RFLP markers. Theor Appl Genet 81: 471–476 (1991).
39. Zhao X, Wu T, Xie Y, Wu R: Genome-specific repetitive sequences in the genus *Oryza*. Theor Appl Genet 78: 201–209 (1989).