Theor Appl Genet (1988) 76:815–829

Molecular mapping of rice chromosomes

S.R. McCouch, G. Kochert*, Z.H. Yu, Z.Y. Wang, G.S. Khush, W.R. Coffman and S.D. Tanksley**

Department of Plant Breeding and Biometry, 252 Emerson Hall, Cornell University, Ithaca, NY 14853, USA

Received March 28, 1988; Accepted June 2, 1988
Communicated by F. Salamini

Summary. We report the construction of an RFLP genetic map of rice (*Oryza sativa*) chromosomes. The map is comprised of 135 loci corresponding to clones selected from a *Pst*I genomic library. This molecular map covers 1,389 cM of the rice genome and exceeds the current classical maps by more than 20%. The map was generated from F_2 segregation data (50 individuals) from a cross between an indica and javanica rice cultivar. Primary trisomics were used to assign linkage groups to each of the 12 rice chromosomes. Seventy-eight percent of the clones assayed revealed RFLPs between the two parental cultivars, indicating that rice contains a significant amount of RFLP variation. Strong correlations between size of hybridizing restriction fragments and level of polymorphism indicate that a significant proportion of the RFLPs in rice are generated by insertions/delections. This conclusion is supported by the occurrence of null alleles for some clones (presumably created by insertion or deletion events). One clone, RG229, hybridized to sequences in both the indica and javanica genomes, which have apparently transposed since the divergence of the two cultivars from their last common ancestor, providing evidence for sequence movement in rice. As a by product of this mapping project, we have discovered that rice DNA is less *C*-methylated than tomato or maize DNA. Our results also suggest the notion that a large fraction of the rice genome (approximately 50%) is single copy.

Key words: *Oryza sativa – Molecular markers – RFLP – Genetic map – Trisomics – DNA methylation*

* Permanent address: Department of Botany, University of Georgia, Athens, GA 30602, USA
** Department of Plant Breeding, International Rice Research Institute, Los Banos, Philippines
Offprint requests to S. D. Tanksley

Introduction

Rice, in many respects, is an ideal monocot species for genetic and molecular studies. It is diploid (2n = 24), has a relatively small genome (C = 0.6 pg), and can be regenerated from protoplasts and transformed by exogenous DNA (Bennett and Smith 1976; Yamada et al. 1985; Lorz and Gobel 1986; Uchimiya et al. 1986). In addition, it is the most widely-grown crop plant worldwide, being a staple food for more than half the world's population (Coffman and Juliano 1987). Finally, there is a vast reservoir of germplasm ($\geq$ 200,000 accessions) of both domesticated and wild rices maintained by rice researchers worldwide.

Genetic studies over the past 70 years have led to the establishment of a linkage map of rice chromosomes based on morphological markers and, more recently, isozyme markers (Kinoshita 1986; Kush and Singh 1986; Wu et al. 1987; Ranjhan et al. 1988). While these maps have provided important insights into the genetics and cytology of rice, they embody inherent limitations common to classical genetic maps. The majority of the markers consist of morphological mutant loci which are undesirable in breeding populations. In addition, alleles of these loci are normally recessive, seriously limiting the types of crosses in which they can be utilized successfully.

Molecular linkage maps have been, or are being, developed in several crop species including tomato (Bernatzky and Tanksley 1986a; Tanksley et al. 1987), maize (Helentjaris 1987; Burr et al. 1988), lettuce (Landry et al. 1987), and potato (Bonierbale et al. 1988). These maps, based largely on cloned DNA sequences, have a number of inherent advantages over classical maps and may provide new opportunities for applications in genetics and breeding. Since variation is detected directly at the DNA level, the alleles behave in a codominant manner and

normally have no visible effect on the phenotype. Because of the large number of these markers that can be isolated, it is possible to develop linkage maps that are, for practical purposes, saturated with genetic markers (Tanksley et al. 1987). Practical applications of these new maps include tagging and tracking major genes of economic interest, such as those for disease resistance and male sterility, as well as dissecting the genetics of more complex traits – those that are quantitatively inherited (Tanksley 1982; Burr et al. 1983; Paterson et al. 1988).

We have embarked on a project to develop a saturated linkage map of rice chromosomes. Such a map promises to be of immediate practical use in the breeding of this important crop, and to offer an opportunity to study, in greater detail, the organization and evolution of the rice genome. We report here the first version of the rice molecular map and describe what we have learned about the organization and behavior of the rice genome during the construction of that map.

Materials and methods

DNA extractions

Fresh-frozen or freeze-dried leaf tissue was ground to a fine powder with 0.5 mm glass beads in liquid nitrogen using a mortar and pestle. The frozen powdered tissue was mixed with urea-phenol extraction buffer (Shure et al. 1983), SDS added to 0.6%, and the solution incubated at 65 °C, inverting gently at 5 min intervals for 15 min. The mixture was then extracted with chloroform:isoamyl alcohol (24:1) to remove proteins and pigments, prior to phase separation by centrifugation (2,000 *g* for 15 min). The upper phase, containing the DNA, was pipetted off through a layer of Miracloth between extractions. DNA was precipitated with 2 volumes 95% EtOH, rinsed in 70% EtOH. dissolved in TE, and re-centrifuged. The supernatant was RNased (to 10μg/ml), re-precipitated with 0.1 volume 6 *M* LiCl and 2 volumes 95% EtOH, re-dissolved in TE, and quantified on the spectrophotometer.

Genomic library construction

Total plant DNA (cv 'IR36') was digested with the restriction enzyme *Pst*I. Digested DNA was size separated on a 10%–40% sucrose gradient and the 1–2 kb size fraction was selected for ligating into pUC8 plasmid (Maniatis et al. 1982). DH5-alpha bacterial cells were then transformed with ligated plasmid. Colonies of cells containing plasmids with rice inserts were selected based on X-gal and IPTG screening procedures. Individual colonies were isolated and plasmid mini-preps prepared according to Wilimzig (1985). Insert sizes were determined relative to *Hae*III-digested Φ X174 on 0.9% agarose gels. Random genomic clones were given consecutive numbers (RG-#) as they were prepared. The same numbers were then used to designate loci in the genome as detected by genetic mapping. Letter suffixes were added in cases where more than one genomic locus contained sequences homologous to a given clone.

Clones of known genes

The ribosomal genes, r5s and r45s, were from wheat (Gerlach and Bedbrook 1979; Gerlach and Dyer 1980). The maize waxy gene was provided by Dr. Susan Wessler, University of Georgia.

Table 1. Summary of polymorphism between sets of parental rice cultivars based on a survey of 12 clones from the *Pst*I library

Cultivars	Race	Origin	Polymorphic clones
IR39352-178-3-1-3	Indica	Philippines	5/12
Kuang Lu Ai 4	Indica	China	(42%)
5173	Indica	Colombia	3/12
Mahsuri	Indica/javanica derivative	Malaysia	(25%)
IR34583-19-3-3	Indica	Philippines	7/12
Bulu Dalam	Javanica	Indonesia	(58%)

Plant materials

Selection of parental material for the mapping experiments was based on a preliminary survey for maximum restriction fragment length polymorphism using 12 single copy probes from the *Pst*I library. Three different sets of parents (Table 1), for which F_2 seed was already available from the International Rice Research Institute (IRRI) in the Philippines, were assayed: (1) 'IR39352-178-3-1-3' × 'Kuang Lu Ai-4' (both indica types); (2) '5173' (indica) × 'Mahsuri' (indica × japonica derivative); (3) 'IR34583-19-3-3' (indica) × 'Bulu Dalam' (a Bulu or javanica rice).

A set of inbred primary trisomics in 'IR36' background (Khush et al. 1984), and a set of 10 different hybrid primary trisomics ('IR36' primary trisomics × 'Ma Hae' – a japonica cultivar) were used to assign linkage groups to chromosomes via dosage effects as described by Tanksley et al. (1987) and Young et al. (1987). Inbred and hybrid primary trisomics were grown at IRRI and lyophilized leaf tissue shipped to Cornell for DNA extraction and analysis.

Restriction digests, electrophoresis, and Southern analysis

DNA from each parental line and F_2 progeny was digested with the following 11 enzymes: *Eco*RI, *Eco*RV, *Dra*I, *Taq*I, *Hae*III, *Hinf*I, *Hind*III, *Msp*I, *Xba*I, (BRL), *Sca*I, and *Bst*NI (NEBiolabs). Spermidine (4 m*M*) was added to digests to promote complete digestion. Electrophoresis of plant DNA (5–6 μg DNA per lane), Southern blotting, and hybridization were according to Bernatzky and Tanksley (1986b). Whole plasmids, including rice inserts, were either nick translated or hexamer labeled with ^{32}P-dCTP to high specific activities ($1-10 \times 10^8$ cpm/μg) (Rigby et al. 1977; Feinberg and Vogelstein 1984), and used as probes on filters of rice DNA.

Filters for trisomic analysis were prepared similarly to those for parents and F_2's, with special care taken to calculate the concentrations of DNA for each triplo plant so that exactly 5 μg *Dra*I-digested DNA would be aliquoted to each lane in the gel. Further, large digestions were prepared, and several filters made from each digestion batch to assure uniformity among filters.

Filters were washed at 65 °C for 10 min each at 2 × SSC, 1 × SSC, and 0.5 × SSC (all washes contained 0.1% SDS), and exposed to X-ray film with intensifier screens. In certain cases, after the films had been developed (usually 24 h), filters were washed again to higher stringencies, 0.05 × or 0.02 × SSC, 65 °C, and re-exposed in order to verify copy number and polymorphism of clones.

Linkage analysis

Linkage analysis was performed on F_2 segregation data as described in Bernatzky and Tanksley (1986a). Recombination

values were calculated for pairs of markers which deviated from independent assortment ($P \leq 0.05$) using F_2 maximum likelihood algorithms published by Allard (1956). Three point linkage tests were used to confirm linear order in regions with several tightly linked markers.

The entire set of markers was then processed through the MAPMAKER computer program (Lander et al. 1987) using 2 point (LOD = 4.0), 3 point (LOD = 3.0), and multipoint (LOD = 3.0) analysis to confirm the order and recombination fractions for the markers. Markers showing segregation distortion (See Table 4) were initially separated and were not used to determine "framework" orders. The results generated by the MAPMAKER program strongly agreed with those derived without the benefit of multipoint analysis, but suggested rearrangement of markers in a few locations, supported by odds of 100:1 that the new orders were more likely. Distances between markers are presented in centimorgans derived using the Kosambi function (Kosambi 1944), and orders where the confidence level is below 99% are bracketed (Fig. 10).

Isozyme survey

The selected parental lines ('IR34583' and 'Bulu Dalam') were assayed for allozymic differences for different enzymes: acid phosphatase (ACP, E.C. 3.1.3.2), esterase (EST, E.C. 3.1.1.2), glutamate-oxaloacetate transaminase (GOT, E.C. 2.6.1.1), isocitrate dehydrogenase (IDH, E. C.1.1.1.42), malate dehydrogenase (MDH, E. C.1.1.1.37), 6-phosphogluconate dehydrogenase (6-PGDH, E. C.1.1.1.49), phosphoglucoisomerase (PGI, E. C.5.3.1.9), phosphoglucomutase (PGM, E. C.2.7.5.1), peroxidase (POX, E. C.1.11.17), shikimate dehydrogenase (SKDH, E. C.1.1.1.25), and triosephosphate isomerase (TPI, E. C.5.3.1.1), which cumulatively represent approximately 30 loci (Wu et al. 1987; Ranjhan et al. 1987; and unpublished results from this laboratory). Electrophoresis and enzyme staining were according to Shaw and Prasad (1970) and Vallejos (1983).

Polymorphism was detected for only PGI, where allelic differences were found for *Pgi-1* (*Pgi-A*), *Pgi-2* (*Pgi-B*) (Wu et al. 1987; Second 1982) and *Pgi-3*, a nuclear-encoded enzyme apparently localized in the chloroplast fraction (unpublished data, this laboratory). Unfortunately it was not possible to monitor segregation of *Pgi* alleles in the F_2 population.

Results

Parental screening and selection

Twelve single copy clones from the *Pst*I library were hybridized to filters containing DNA digested with eight restriction enzymes from the three sets of parents described in Materials and methods. Results from this experiment indicated that maximum polymorphism, 58% (7/12 clones), occurred between the 'IR34583' – 'Bulu Dalam' parents (Table 1). The remaining two sets of parents showed RFLP differences for 41% (5/12) and 25% (3/12) clones, respectively. The 'IR34583' – 'Bulu Dalam' F_2 was thus selected as the mapping population.

Library characterization

Clone copy number. A total of 300 genomic clones were surveyed for copy number and polymorphism by probing them onto filters containing total DNA from 'IR34583' and 'Bulu Dalam' that had been digested with 11 different restriction enzymes. The first 210 clones surveyed were considered a sample population, and all statistical analyses presented in the following sections were performed on this subset of clones. The clones were classified into three categories based on relative copy numbers. Categories were defined as follows: (1) single copy sequences (s.c.) where >90% of the signal in a lane was accounted for by one or two discrete bands (Fig. 1 a); (2) multiple copy sequences (m.c.) where 3 or more bands were present and no single band accounted for less than 20% or more than 80% of the signal (Fig. 1 c,d,f); and (3) repeated sequences where the signal in a lane appeared as a smear with no discrete bands, or if no single band within the smear accounted for more than 20% of the total signal in the lane (Fig. 1 e). After defining these categories with densitometric scans, most probes could be classified by eye based on autoradiograms.

When filters were washed at medium stringency (1 × or 0.5 × SSC, 65 °C), 58% of the clones were classified as single copy, 20% as multiple copy, and 22% as repeated (Fig. 2). At this stringency, it is estimated that the minimum level of sequence homology between duplexes is 80% (Beltz et al. 1983). At high stringency (0.05 × or 0.02 × SSC), where 99% sequence homology is required for maintaining duplex formation, 85% of the clones were classified as single copy, 12% multiple copy, and 3% repeated sequences (Fig. 2). Henceforth, references to copy number will refer to medium stringency unless specified otherwise.

Comparison of rice PstI library with similar libraries in tomato and maize – evidence for lower levels of C-methylation in rice. The library used in this study was constructed with the restriction enzyme *Pst*I, a methylation-sensitive enzyme which recognizes the sequence 5′-CTGCAG-3′ and is sensitive to cytosine methylation in the 5′ location (Nelson and McClelland 1987). Clones represented in the library used in this study are likely to be from regions of the genome which are not methylated at this site.

It has been observed in maize (Burr et al. 1988) and tomato (Tanksley et al. 1987) that such undermethylated areas are enriched for single copy sequences, possibly since they represent coding regions (Burr et al. 1988). It is of interest then to compare the percentage single copy versus multiple copy or repeated sequence probes generated by *Pst*I libraries in these three species. Rice and tomato have similar size genomes (0.6 and 0.7 pg of DNA, respectively) (Bennett and Smith 1976; Iyengar and Sen 1978; Galbraith et al. 1983), and both have a haploid complement of 12 chromosomes. In tomato, 40% of clones derived from libraries constructed from sheared DNA, or from the *C*-methylation-insensitive enzymes (e.g., *Eco*RI), are estimated to contain unique sequences (Tanksley et al. 1987; Zamir and Tanksley 1988). However, when clones are derived from a *Pst*I

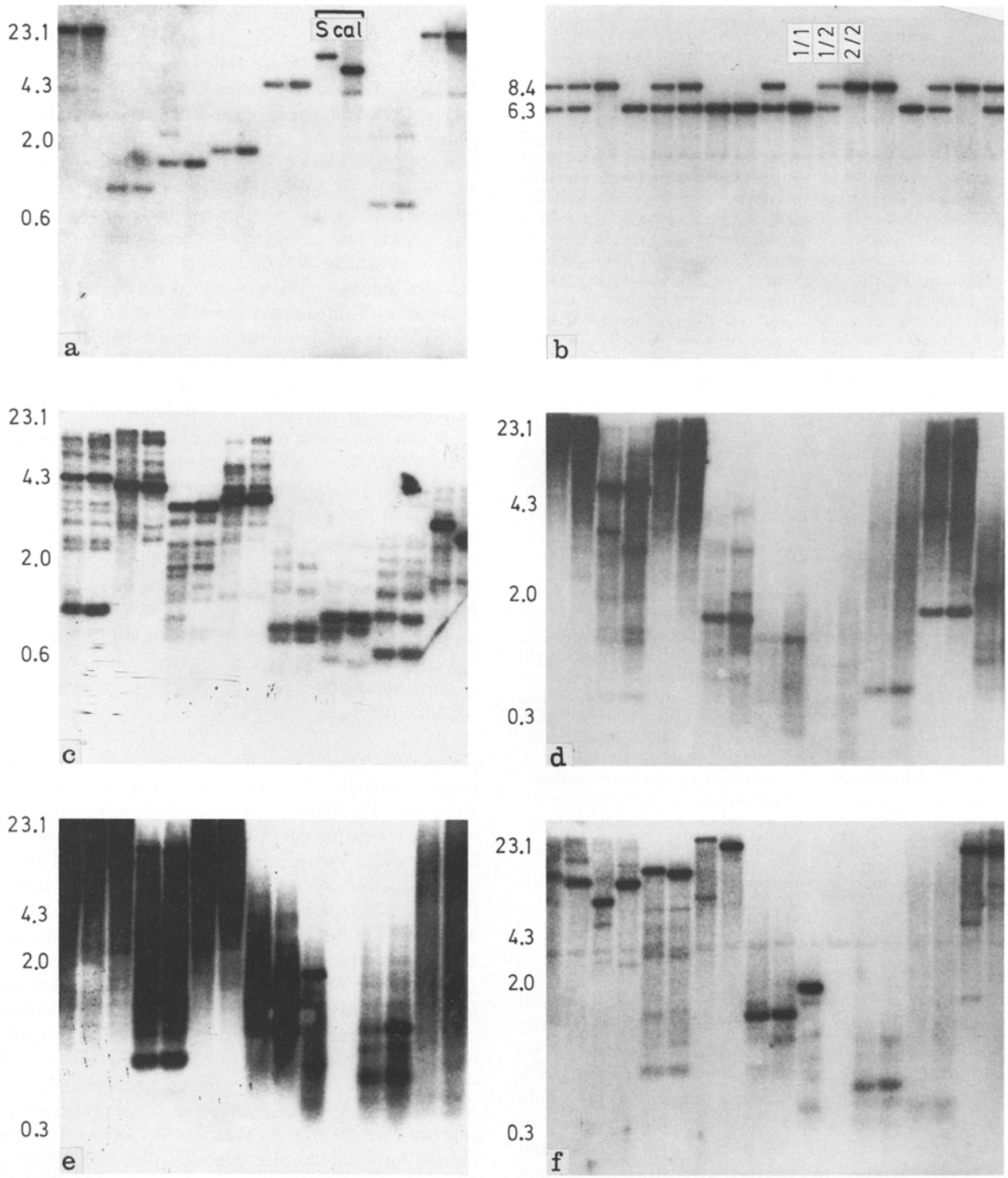

Fig. 1. **a** Autoradiogram derived from hybridizing a single copy clone, RG123, onto DNA from 'IR34583' (indica) and 'Bulu Dalam' (javanica). Indica and javanica DNA was digested in sets with different restriction enzymes, run out on an agarose gel, blotted and probed as described in Materials and methods. In each pair, *left lane* = indica, *right lane* = javanica. Left hand margin indicates molecular weight in kbp. For this probe, *Sca*I detected polymorphism and was subsequently used to map clone. **b** Autoradiogram from F_2 segregation probed with single copy clone. *1/1* = *indica* homozygote, *2/2* = *javanica* homozygote, *1/2* = heterozygote. **c** Autoradiogram derived from probing clone classified as multicopy (RG374). Composition of filter same as described in **a**. **d** Autoradiogram derived from probing another clone also classified as multicopy (RG196). In this case, part of the signal is present as an undifferentiated smear. **e** Autoradiogram derived from probing clone (RG343) classified as repetitive at moderate stringency (0.5 × SSC, 65 °C). **f** Same as **e** only filter washed to high stringency (0.05 × SSC, 65 °C). At this stringency, discrete bands are revealed

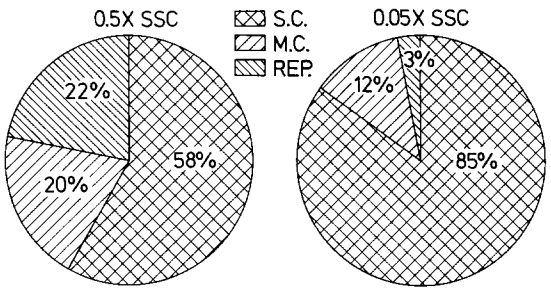

Fig. 2. Pie chart showing classification of *Pst*I genomic clones at moderate stringency (0.5 × SSC, 65 °C) and high stringency (0.05 × SSC, 65 °C). Statistics based on analysis of 210 clones

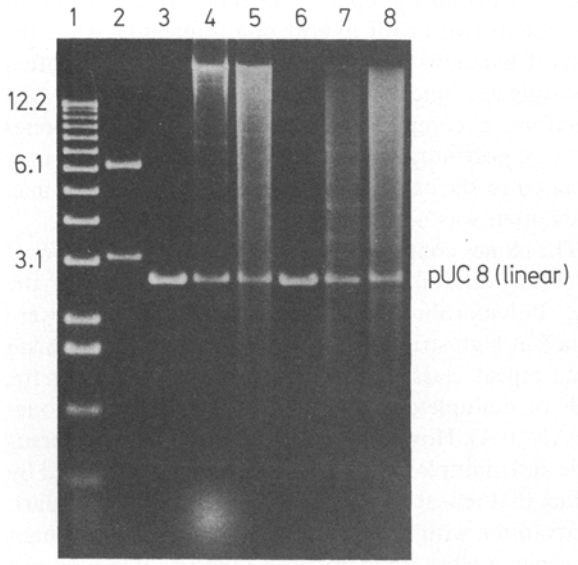

Fig. 3. Gel comparing molecular weight distribution of restriction fragments when tomato and rice DNA are digested with *Pst*I (*C*-methylation sensitive) and *Eco*RI (methylation insensitive). pUC8 plasmid (*uncut-lane 2*) was added to reactions as a control for complete digestion. Note that most of the tomato DNA remains high molecular weight when cut with *Pst*I (*lane 4*), but is cleaved to smaller fragments when digested with *Eco*RI (*lane 7*). In contrast, rice DNA is cleaved to smaller fragments with both *Pst*I (*lane 5*) and *Eco*RI (*lane 8*)

library of similar size to those isolated from rice,the single copy fraction rises to 92% (Tanksley et al. 1987). In maize, a significant increase in the proportion of single copy clones has also been observed when clones are derived from *Pst*I libraries (Burr et al. 1988).

Results of the survey with *Pst*I-generated genomic clones in rice estimate the single copy fraction of the library to be about 58%, significantly less than expected based on the tomato and maize results, suggesting that a *Pst*I library in rice is less biased toward single copy sequences than tomato or maize. To test this possibility, tomato and rice DNA were subjected to comparative digests with *Pst*I, as well as with the methylation-insensitive enzyme, *Eco*RI. In tomato, the majority of the DNA remains high molecular weight when digested with *Pst*I, while in rice, the DNA digested with *Pst*I shows a distribution of fragment sizes that is intermediate between that generated by *Pst*I in tomato and by *Eco*RI in either rice or tomato (Fig. 3). Comparative digests with other *C*-methylation-sensitive enzymes, including *Sal*I, *Cla*I, *Sma*I, *Xho*I, and *Hpa*II, produced similar results (data not shown). In all cases, rice DNA was cut more effectively by these enzymes than was tomato DNA. These data suggest that *C*-methylation is less frequent in rice than in tomato, and that a significant fraction of the repeated DNA in rice is not methylated as in tomato and maize.

Restriction fragment length polymorphism

Variation between rice cultivars. Despite a low frequency of isozyme polymorphism between the parents used for mapping, 'IR34583' and 'Bulu Dalam' (only PGI was polymorphic – see Materials and methods), 164 of the 210 genomic clones surveyed (78%) revealed allelic differences with at least 1 of the 11 restriction enzymes. This percentage is less than that found among most unrelated maize inbreds (Helentjaris et al. 1985; Murray et al. personal communication), but is significantly higher than that found in tomato (J. C. Miller and S. D. Tanksley, personal communication).

Variation within inbred parental lines. Occasional discrepancies in RFLP patterns between parental survey filters and F_2 filters pointed to the possibility of variation existing within the parental populations. The purity of the parental populations was tested by probing filters containing digested DNA from 12 individual plants of each parental line selected at random, with probes suspected of being polymorphic within the parental lines. Parental lines are expected to be pure lines in an inbred crop such as rice where outcrossing is estimated to occur naturally approximately 0.5% of the time (Poehlman 1979). However, cross-pollination events are possible, and if crossing occurred between individuals differing in genotype, some degree of heterozygosity would result, accompanied by continued genotypic segregation in subsequent generations. Polymorphism could also be the result of mixtures of pure line seed in one or both parental cultivars, in which case the 12 individuals tested per cultivar would show no heterozygous loci, but different alleles in the homozygous state would be present in varying proportions.

Of the four probes used to test for polymorphism within parental cultivars, three detected variation only within the javanica cultivar, while one detected polymorphism only within the indica cultivar. Single copy probes RG-140 (chromosome no. 1) and RG-149 (not mapped) produced banding patterns supportive of the hypothesis

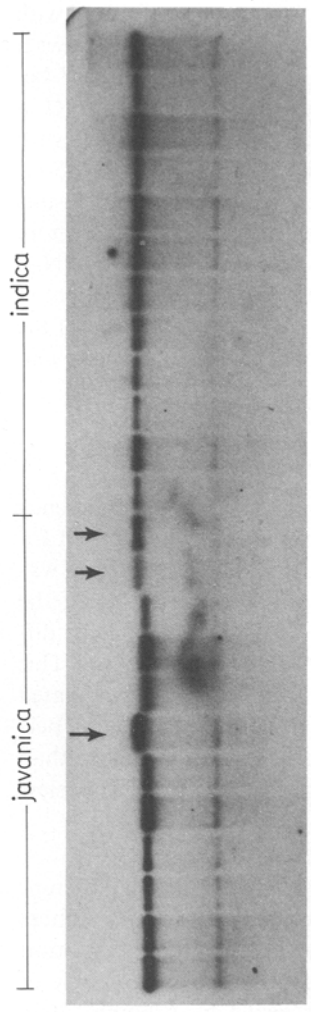

Fig. 4. Autoradiogram derived from probing *Hind*III-digested DNA from12 individuals of 'Bulu Dalam' (javanica parent, left) and 12 individuals of 'IR34583' (indica parent, right). Filter was probed with RG140, a clone suspected of detecting variation within the parental populations. *Arrows* point to individuals homozygous for variant allele in javanica variety

that the javanica parent is a mixed line. The same three individuals had variant alleles in the homozygous state for both probes (Fig. 4). These data suggest that the 'Bulu Dalam' cultivar used in this cross was not a pure line. Probe RG-152 uncovered heterozygosity within the cultivar in a single individual that showed the normal banding pattern for RG-140 and RG-149.

RG-229, a multiple copy clone, was the fourth probe used to test for polymorphism within parental populations. In this case, polymorphism was evident within the indica cultivar; one fragment was present in all 12 plants, while a second fragment was present in only 5 plants. The banding pattern in the javanica was the same for all 12 individuals (each plant uniformly contained two hybridizing fragments). This clone is of interest in that unusual banding patterns were observed among individuals in the F_2 population as well (see following sections).

Differences among probes in detecting polymorphism. The percentage of clones detecting polymorphisms between the parents was similar for single copy (s.c.) and multiple copy clones (m.c.) (81% and 80%, respectively). This was surprising since a clone hybridizing to multiple fragments assays a larger number of restriction sites and thus a greater number of base pairs. A feature of the m.c. clones in this library was that two very different patterns of hybridization were evident. Eighty-five percent of the time, the hybridization pattern included a clear signal for only one or two major hybridizing fragments, while the array of bands with less signal lacked resolution, often appearing as a smear, and could not be scored (Fig. 1 d). Therefore, a comparison of the percentage of clones detecting polymorphism in the s.c. and m.c. clases may be biased to the extent that the full complement of m.c. bands often was not clearly distinguishable.

The clones containing repeated sequences uncovered detectable variability between the parents 64% of the time. Polymorphism was evident when filters were washed at high stringency. The degree of polymorphism in the repeat class was significantly less than that for single or multiple copy clones ($P = .05$, based on a one-way ANOVA). However, comparisons of repeated versus single and multiple copy sequences are strongly biased by the fact that repeat sequence hybridization produce dark smears, often with no discrete bands visible even at high stringency washes, thus obscuring the possible presence of RFLPs.

Differences among restriction endonucleases in detecting RFLPs. When DNA was cut with different enzymes, significant differences were observed in the proportion of clones hybridizing to polymorphic restriction fragments ($P < 0.01$, based on Chi-square analysis of a two-way contingency table). Enzymes that have a 6 bp recognition site uncovered the most variability, while those that recognized 4 bp sequences uncovered the least (Table 2).

The average size of restriction fragments hybridizing to clones was also significantly greater for enzymes recognizing a 6 bp sequence (7.24 kb) than for those recognizing a 4 bp sequence (1.33 kb) ($P < 0.01$, as determined by a one-way ANOVA). Restriction fragments generated by *Bst*NI, a 5 bp site, were slightly larger (2.1 kb) than those generated by the 4-cutter group (Table 2).

If all nucleotides occur independently and in equal frequency, a given 6 bp recognition sequence will occur every 4,096 bp within a genome, and 4 bp sites, every 256 bp. Therefore, the 6 bp enzymes were expected to cleave less often, generating larger fragments than 4 bp enzymes. This was confirmed by looking at stained gels

Table 2. Average size of restriction fragments hybridizing to genomic clones and % probes detecting polymorphism between indica and javanica parents ('IR34583' and 'Bulu Dalam') for 11 restriction enzymes

Enzyme	Recognition sequence	Expected MW of hybridizing restriction fragment	Average MW of hybridizing restriction fragment	% Probes polymorphic
*Eco*RV	GATATC	4.10	9.87	39.4
*Xba*I	TCTAGA	4.10	9.24	40.3
*Eco*RI	GAATTC	4.10	7.82	36.4
*Hind*III	AAGCTT	4.10	7.06	30.7
*Sca*I	AGTACT	4.10	5.70	29.8
*Dra*I	TTTAAA	4.10	3.75	30.7
*Bst*NI	$CC^{A}_{T}GC$	1.02	2.10	26.2
*Msp*I	CCGG	0.26	2.00	22.7
*Hae*III	GGCC	0.26	1.20	20.5
*Taq*I	TCGA	0.26	1.10	21.7
*Hinf*I	GANTC	0.26	1.00	15.6

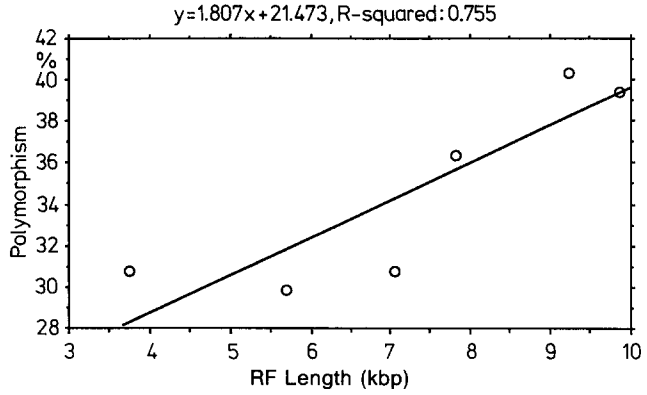

Fig. 5. Plot derived from regressing % polymorphic probes on average length of hybridizing restriction fragment. Each point represents the value for each of the 6 bp recognition enzymes

of digested rice DNA and by calculating molecular weights of hybridizing fragments on autoradiograms. However, fragments hybridizing to probes were considerably larger than expected. This result may be due in part to a GC content different than 50% or to our inability to detect very small hybridizing fragments (< 200 bp) on the Southerns. The average size of hybridizing restriction fragments for all enzymes was highly variable and, on average, twice the expected value (Table 2). Surprisingly, there were also significant differences in the average length of hybridizing restriction fragments among the 6 bp recognition enzymes ($P < 0.01$) even though the recognition sequences of all but one of the 6 bp enzymes contained permutations of the bases AATTGC. *Dra*I, which has no GC in its recognition site, produced the shortest average hybridizing fragments, 3,750 bp.

Correlation between fragment length and polymorphism – evidence for insertions/deletions

If RFLPs are generated largely by base substitutions, as opposed to insertions/deletions, one would predict no correlation between the amount of polymorphism detected by an enzyme and the average size of restriction fragments generated by that enzyme. This prediction is only valid for enzymes recognizing the same number of base pairs. The fact that there were significant differences in the average size of restriction fragments for the 6 bp enzymes in rice provided a test of this hypothesis. When the degree of polymorphism detected by 6 bp enzymes was regressed on average length of hybridizing restriction fragments, the regression was found to be highly significant (supporting the hypothesis that RFLPs in rice are caused by insertions/deletions) (Fig. 5).

Another way of testing this hypothesis is to ask whether the probability that a given enzyme detects polymorphism with a given probe is a function of the number of other enzymes that detected polymorphism with that same probe. If RFLPs are the result of base substitution, there should be no relationship between these two variables. This assertion is valid only if restriction sites for different enzymes along a piece of genomic DNA are independent and do not overlap. Based on the uniqueness of the recognition sites for the enzymes used in this study (Table 2), this assumption is likely to be correct. However, if an insertion or deletion is the cause of an RFLP, then RFLPs should be generated by any enzyme whose restriction sites encompass the insertion/deletion and the probability of different enzymes detecting polymorphism with a given probe would not be independent (Fig. 6). For each of the 6 bp recognition enzymes, probability of polymorphism was regressed on the number of other restriction enzymes detecting polymorphisms for all of the probes analyzed. The regression coefficient was positive and significant in every case ($R^2 > 0.66$), supporting the insertion/deletion hypothesis. The resulting regression plot for *Eco*RV is shown in Fig. 7.

CpG site variability. It has been reported that in mammalian and prokaryotic systems, high mutation rates are associated with CpG dimers (Barker et al. 1984). This possibility was investigated in rice by comparing the degree of polymorphism detected by 4 bp recognition enzymes containing a CpG dimer in their recognition sequences (*Msp*I, *Taq*I) to polymorphism detected by other 4 bp enzymes without the CpG in the cutting site (*Hae*III, *Hinf*I). Results in rice, like those in lettuce (Landry et al. 1987), suggest that CpG dimers are not hotspots for mutation in these plants since enzymes containing the dimer do not detect significantly higher frequencies of polymorphism than other enzymes.

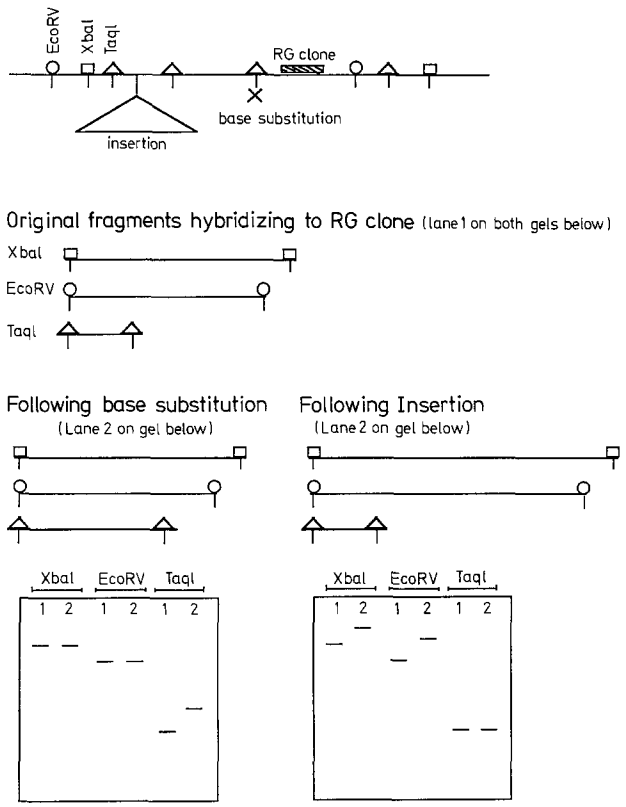

Fig. 6. Schematic depicting the expected restriction fragments generated by different restriction enzymes if polymorphism is due to base substitution versus sequence insertion. Note, with base substitution, the restriction pattern is affected for only one restriction enzyme whereas insertion is likely to affect more than one enzyme. Also note the enzymes which cut out larger fragments are more likely to include the insertion if it flanks the cloned region

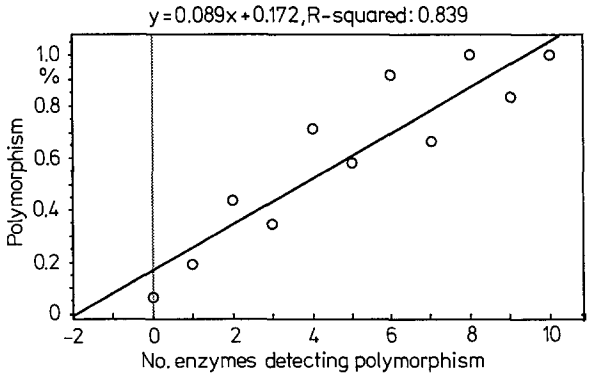

Fig. 7. Plot derived from regressing probability of polymorphism (between parental lines) detected by *Eco*RV on the number of other enzymes detecting polymorphism with each probe. Positive correlation would be expected if RFLPs are due to insertions/deletions. Similar results were obtained when similar plots were made for other 6 bp recognition enzymes (not shown)

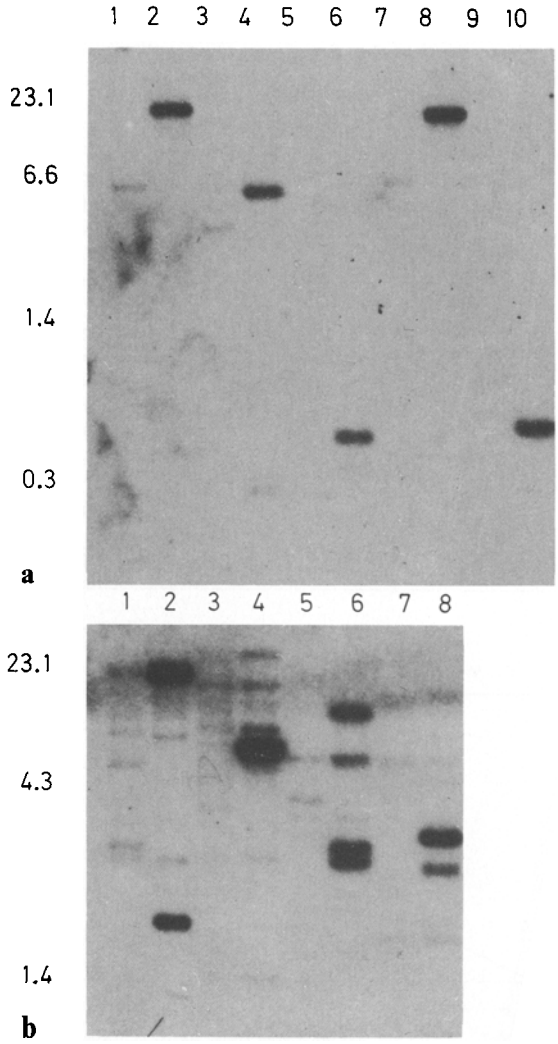

Fig. 8a and b. Autoradiograms from clones revealing null alleles in javanica parent. Each pair of *lanes* (i.e. *1–2, 3–4,* etc) represents digestion with a different restriction enzyme. *Left lane* in each set = javanica DNA, *right lane* = indica DNA. Left hand margin indicates molecular weight in kbp: **a** probed with RG358; **b** probed with RG333

Clones uncovering unusual polymorphisms

RG333 and RG358 – clones with null alleles. Two single copy clones, RG333 and RG358, were found to hybridize readily to DNA from the indica parent, but very faintly to DNA from the javanica parent (Fig. 8). Segregation in the F_2 was consistent with single locus inheritance (3:1, presence versus absence), and their corresponding loci could be mapped to the indica genome, chromosomes 8 and 9, respectively. A possible explanation of why these two clones hybridized only with indica DNA is that the clones were derived from an indica (or 'IR36') library, and that there have been major sequence rearrangement(s) in the corresponding regions of the javanica pa-

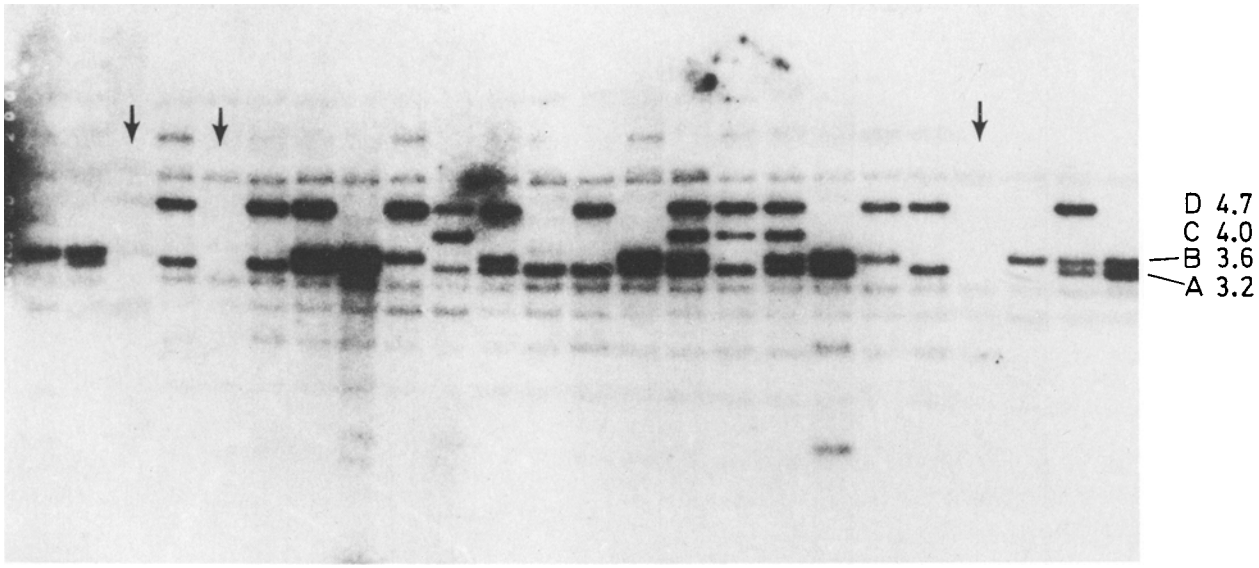

Fig. 9. F_2 segregation (from *Dra*I digest) for RG229. *Arrows* point to individuals that inherited none of the major hybridizing fragments from either parent. Null genotypes would be possible only if fragments (*a–d*) are not allelic. Fragments *a, b* come from javanica parent; fragments *d, c* from indica parent

Table 3. Presence/absence ratios of independently segregating alleles of RG229 fragments in F_2 population

Fragment	Present	Absent	χ^2 (3:1)
229A	32	10	0.03
229B	27	15	2.57
229C	8	34	72.12*
229D	31	11	0.03

* Significant deviation at $P<0.01$

rent relative to the indica parent such that the homologous counterparts to RG333 and RG358 no longer exist or perhaps never existed in the javanica genome.

RG229 – a clone with properties reminiscent of a transposon. The segregation of RG229 was unique. This clone hybridized to two principal restriction fragments in DNA from each parent, and both fragments showed a high degree of polymorphism (9 out of 11 enzymes) on survey films. By employing high stringency washes of 0.05 × SSC (indicative of >99% sequence homology), it was determined that the genomic sequences homologous to RG229 are highly conserved within and between the two genomes. However, neither of the homologous restriction fragments present in the two parents cosegregated, indicating that they do not reside at the same locus. In addition, the segregation patterns indicate that none of the fragments from the indica parent are allelic to any of the fragments from the javanica parent (Fig. 9). This point is exemplified by the presence of three F_2 individuals which have no fragments homologous to RG229 (Fig. 9). Such genotypes would not be expected if any of the fragments were allelic.

By testing segregation of each RG229 fragment against all other fragments, it was determined that all hybridizing fragments were segregating independently (Table 3). Three of the four fragments, two from 'Bulu Dalam' (A, B) and one from 'IR34583' (D) also segregated in a 3:1 ratio (presence versus absence Table 3), indicative of discrete, independent genetic loci. Finally, three of the fragments (A, B, D) mapped to different chromosomes (9, 1, 5), ruling out the possibility of allelism (Fig. 10, see next section for details of mapping). The C fragment (from indica) did not segregate 3:1, nor were we able to map it to any of the established linkage groups. A survey of plants from the indica parent population provided an explanation for the aberrant segregation and probably the failure to map this fragment. The indica parent line was found to be polymorphic for presence/absence of the C fragment (data not shown) which likely led to a polymorphic F_1 population from which F_2 plants were derived. As a result, only a small fraction of the F_2 plants (19%) inherited the band, and thus the segregation ratio deviated significantly from the expected 3:1.

It has not escaped our attention that RG229 has some features indicative of transposition. All of the genomic copies are highly conserved suggesting that they all derived from a common ancestral sequence relatively recently (or were homogenized by gene conversion), yet

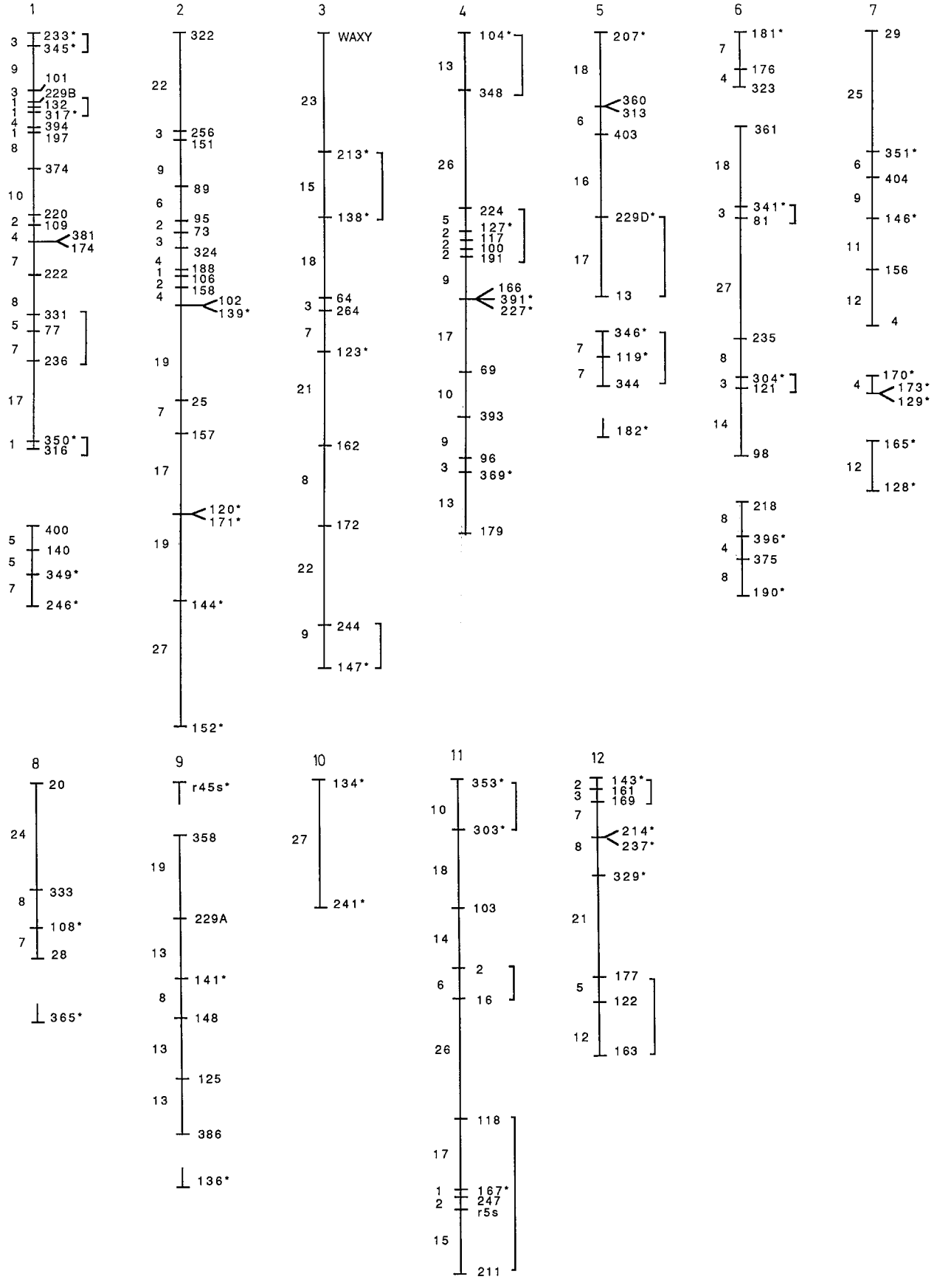
1
233*
345*
101
229B
132
317*
394
197
374
220
109
381
174
222
331
77
236
350*
316
400
140
349*
246*
2
322
256
151
89
95
73
324
188
106
158
102
139*
25
157
120*
171*
144*
152*
3
WAXY
213*
138*
64
264
123*
162
172
244
147*
4
104*
348
224
127*
117
100
191
166
391*
227*
69
393
96
369*
179
5
207*
360
313
403
229D*
13
346*
119*
344
182*
6
181*
176
323
361
341*
81
235
304*
121
98
218
396*
375
190*
7
29
351*
404
146*
156
4
170*
173*
129*
165*
128*
8
20
333
108*
28
365*
9
r45s*
358
229A
141*
148
125
386
136*
10
134*
241*
11
353*
303*
103
2
16
118
167*
247
r5s
211
12
143*
161
169
214*
237*
329*
177
122
163

Table 4. Segregation ratios for loci deviating from expected monogenic ratio (1:2:1). jav/jav = homozygous for javanica alleles, ind/ind = homozygous for indica alleles, jav/ind = heterozygous

Locus	Chromosome	jav/jav	jav/ind	ind/ind	χ^2 (1:2:1)	p
RG400	1	11	30	4	7.178	0.028
RG349	1	14	28	3	8.067	0.018
RG25	2	10	14	21	11.800	0.003
RG152	2	4	11	29	39.409	0.000
Waxy	3	4	20	21	13.400	0.001
RG213	3	2	22	18	12.286	0.002
RG138	3	7	17	21	11.400	0.003
RG64	3	5	20	19	9.273	0.010
RG264	3	5	18	21	13.091	0.001
RG123	3	4	21	20	11.578	0.003
RG127	4	6	21	19	7.696	0.021
RG166	4	5	17	22	15.409	0.000
RG391	4	5	16	23	18.000	0.000
RG227	4	6	17	19	9.571	0.008
RG69	4	2	20	18	12.800	0.002
RG393	4	2	21	20	15.093	0.001
RG96	4	3	25	17	9.267	0.010
RG369	4	3	27	13	7.465	0.024
RG179	4	1	22	17	13.200	0.001
RG98	6	3	24	9	6.000	0.050
RG4	7	13	20	3	6.000	0.050
RG134	10	3	28	12	7.698	0.021
RG211	11	7	17	22	12.913	0.002
RG329	12	9	16	20	9.133	0.010
RG122	12	5	21	17	6.721	0.035

copies of this element in two different genomes (indica and javanica) reside in different chromosomal positions. This result might be explained if RG229 had transposed in one or both genomes since the javanica and indica diverged from their last common ancestor. There is no evidence, however, that RG229 has been mobilized in the cross used for this mapping project since no new hybridizing restriction fragments were detected in the progeny – a result that would be expected if transposition had led to insertion into new genomic sites. The chromosomal distribution, nucleotide sequence, and phylogeny of RG229 is currently being investigated.

Linkage map construction

Selecting clones for mapping. Clones classified as single copy at high stringency, and for which an RFLP variant could be detected, were mapped. Enzyme(s) detecting variation between the two parents were used for digestion of progeny DNA. Resulting F_2 filters were probed with the corresponding clone, and autoradiograms scored for segregating genotypes (Fig. 1b). Linkage among segregating markers was determined as described in Materials and methods.

Fig. 10. Molecular map of rice. *Numbers* at top indicate chromosomes. *Asterisks* indicate chromosomal location verified by trisomic dosage analysis. Orders where the confidence level is below 99% are *bracketed*

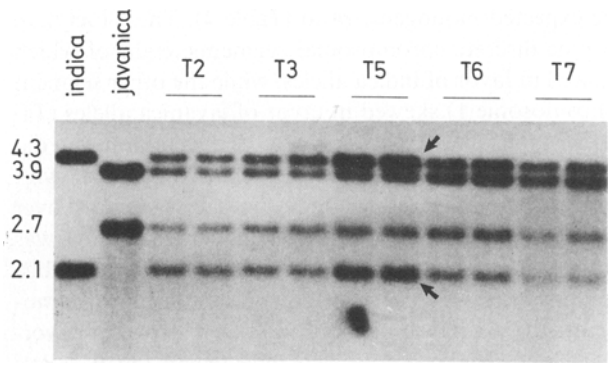

Fig. 11. Autoradiogram showing dosage effect in hybrid primary trisomics. Clone (RG182) is located on chromosome 5 as indicated by the increased signal for the indica fragments (*arrows*) in DNA from the triplo 5 plants (T5). DNA was digested with *Hind*III

Assignment of linkage groups to chromosomes. Once linkage groups had been tentatively identified, based on F_2 segregation analysis of RFLP markers, dosage analysis of trisomic DNA was performed to assign groups to their respective chromosomes. Both inbred ('IR36') and hybrid primary trisomics were used in this effort. All 12 inbred trisomics were available, but because of poor growth, only limited amounts of tissue could be obtained from several of these stocks. Hybrid trisomics were more vigorous and could be propagated more easily than the inbred trisomics, and therefore tissue was available in greater quantity. Ten different hybrid trisomics (missing Triplo 1 and 4) were used. Many clones were tested by both inbred and hybrid trisomics, and in all cases these results agreed with respect to chromosomal assignment.

Dosage comparisons were made within lanes by comparing the relative intensity of bands derived from hybridization with clones known to reside on different linkage groups. This was accomplished by either probing the same filter with several labeled probes or by re-probing the same filter and overlaying autoradiograms. Hybrid trisomics automatically provided convenient internal dosage controls where allelic differences between the indica and javanica parents were apparent. In the positive hybrids, the additional dosage showed in the fragment corresponding to the indica allele (Fig. 11).

Skewed segregation in favor of indica alleles. Twenty-five of the loci analyzed in the F_2 deviated significantly from

the expected monogenic ratio (Table 4). These loci map to nine discrete chromosomal segments, eight of which skewed in favor of indica alleles, while the other segment (chromosome 1) skewed in favor of javanica alleles (Table 4). When analyzed *individually*, the remaining loci did not deviate from the expected ratio (data not shown). However, when genotypic frequencies were summed over these remaining loci, a highly significant deviation was observed in favor of indica homozygotes (34% indica homozygotes, 45% heterozygotes, 21% javanica homozygotes: Chi-square = 59.48, $P < 0.01$). The overall favoring of indica alleles was also reflected in the F_2 allele frequency (54% indica alleles: 46% javanica alleles: Chi-square (1:1) = 80.50, $P < 0.01$).

Comparison on classical and RFLP linkage maps

The rice molecular map, derived from segregation data and trisomic analysis, is presented in Fig. 10. The 135 markers are distributed throughout the 12 chromosomes, defining a total of 1,389 centimorgans. The total map units covered by the RFLP map is greater by approximately 20% and 50%, respectively, than the classical linkage maps (both RFLP and classical maps compared on basis of recombination fraction) published by Kinoshita (1986) and Khush and Singh (1986), though the number of markers on all three maps is similar (119 for Kinoshita and 120 for Khush and Singh).

The distribution of the RFLP markers on the molecular map was compared with that of the morphological markers on the classical linkage map in terms of the numbers of markers per chromosome and the spacing of markers within chromosomes. There was a significant positive regression coefficient ($y = 0.263x + 2.493$, $R^2 = .67$) when the number of RFLP markers per chromosome was regressed on the length of chromosome as measured at the pachytene stage of meiosis (Shastry et al. 1960). In contrast, the regression of markers on the classical maps against chromosome length was not significant.

A Chi-square analysis was used to determine whether the number of RFLP or morphological markers on each chromosome differed from the expected, based on relative chromosomal lengths. Markers per chromosome on both morphological linkage maps differed from the expected ($P < 0.05$), while those on the RFLP map did not.

The mean map distance between markers, measured in recombination fractions, was not different among chromosomes for any of the three maps, based on a one-way ANOVA, suggesting that there is no chromosome-specific pattern of distribution. However, when the mean genetic distance per chromosome was regressed on number of markers per chromosome, a significant negative regression coefficient resulted from the RFLP map ($y = 0.857x + 23.394$, $R^2 = 0.656$). In order to make sure that this relationship was not confounded by the occurrence of unlinked segments in some chromosomes, the number of gaps per chromosome was checked for correlation with the number of markers per chromosome, and found not to be significant. These data suggest that markers on well-populated chromosomes tend to occur more tightly clustered than do markers on less populated chromosomes.

Discussion

DNA methylation

A comparison of cutting efficiencies of *C*-methylation-sensitive and insensitive restriction enzymes on rice and tomato DNA indicated that rice DNA is less methylated than that of tomato. Yet these two species have approximately the same C value. In tomato, most of the repetitive sequences are hypermethylated, whereas single copy sequences are undermethylated. This does not appear to be the case with rice where a significant fraction of repeated sequences was found in a library constructed with the *C*-methylation sensitive enzyme *Pst*I.

The role of DNA methylation in eucaryotes, and especially plants, is poorly understood. There is evidence that methylation plays a role in gene regulation, yet this cannot be a universal function as some eucaryotes, like *Drosophila*, have little or no methylated DNA (Doerfler 1983). The finding that two plant species (rice and tomato), of similar genome size and composition, vary so dramatically in level and disposition of methylation is a poignant reminder that we still have much to learn about the role of DNA methylation.

Genome organization

The view of genome organization that emerges from this study suggests that there are three classes of sequences based on copy number. Approximately 58% of the sequences surveyed in this study were single copy sequences, 20% multiple copy, and 22% repeated sequences. Previous studies using renaturation kinetics suggest that the rice genome is comprised of three fractions, 48% slowly reassociating or single copy, 25% intermediately reassociating, and 27% rapidly reassociating or repeated sequences (Deshpande and Ranjekar 1980).

The rice genome appears to be comprised of a relatively large proportion of single copy sequences which is in sharp contrast to other grasses (rye, wheat, or maize) which have large genomes and a higher proportion of repeated DNA (Flavell 1980). This observation is in keeping with the hypothesis that total genome size is

positively correlated with the size of the repeat fraction of DNA (Flavell 1980). The relatively small genome size and relatively large fraction of single copy DNA in rice makes it currently the best prospect among monocots for experiments involving "chromosome walking" – a technique now in use in other eucaryotic organisms for isolating genes of unknown gene product via genetic linkage to cloned RFLP probes (Orkin 1986).

Transposon activity may be a major source of RFLPs in rice

Restriction fragment length polymorphism can be caused by either base substitution or insertions/deletions. The latter may be the result of transposon activity which has been implicated as a major source of DNA variation in maize (Schwarz-Sommer et al. 1985). In rice, restriction enzymes which generate larger restriction fragments also tend to uncover more RFLPs (see Results section for details). Moreover, when one enzyme uncovers an RFLP, the probability that another enzyme will uncover an RFLP in the same clone, is significantly increased. The only plausible explanation for these observations is that a significant portion of RFLPs in rice is the result of insertions/deletions.

Another piece of evidence supporting an insertion/deletion hypothesis is the existence of two clones (RG333 and RG358) that hybridize strongly to DNA from the indica parent and very faintly to the javanica parent, but which segregate in the F_2 in a manner consistent with single locus inheritance. These probes most likely represent specific sequences in the indica genome that were either deleted from the javanica genome or inserted in the indica genome since these two lines diverged from their last common ancestor. Similar null alleles have been reported in lettuce (Landry et al. 1987) and in the nematode, *Caenorhabditus* (Herman and Shaw 1987).

Finally, we present evidence for sequence transposition in rice. One of the cloned sequences used for mapping, RG229, was found in two copies in both the indica and javanica parental genomes. However, genetic analysis indicated that each of the copies occupies a unique chromosomal position (i.e., none of the fragments are allelic). This result can most readily be explained by transposition of this sequence in either the indica or javanica genomes (or both) since their divergence from their last common ancestor.

Transposons have not been directly demonstrated in rice the way they have been in maize. However, there is recent evidence for high levels of gene inactivation and reversion as the result of interspecific hybridization in rice (G. S. Khush, personal communication). This phenomenon has many similarities to hybrid dysgenesis in *Drosophila* which is believed to result from transposon activation and provides additional circumstantial evidence that the rice genome may be host to one or more potentially active transposon families.

Molecular map of rice

The construction of a molecular linkage map in rice based on RFLPs offers new opportunities for applications in genetics and breeding. Genes governing agronomically important traits can be mapped based on linkage with RFLP markers, and segregating populations screened in the seedling stage for presence or absence of the chromosomal segments bearing the genes of interest. For traits that are difficult or time-consuming to score using conventional means, RFLP analysis, though expensive at this time, offers a hopeful alternative. A mapped array of DNA markers also offers the possibility of dissecting quantitatively inherited traits into their component genetic loci, pinpointing specific alleles that interact to give heterosis, as well as fingerprinting varieties, tracing the introgression of genes from wide crosses, and initiating a chromosome walk aimed at identifying and ultimately cloning a gene of interest.

The availability of such a map and of the vast array of germplasm preserved by rice workers around the world, in conjunction with current advances in transformation and regeneration systems in rice, promise to make this plant an attractive organism for genetic and evolutionary studies at the molecular level. The implications of productive interaction of basic and applied research in rice are particularly noteworthy for a crop which serves as a dietary staple for more than half of the world's population.

The RFLP mapping project is continuing with the goal of a more fully saturated map, aimed especially at adding markers to poorly represented chromosomes and closing the gaps that occur due to a paucity of markers in certain regions of chromosomes. The physical location of markers on the chromosomes and the relationship between the genetically defined distance and actual physical distance separating markers is being pursued through in situ hybridization techniques.

A second map is also in the early phases of preparation, based on a population of doubled haploids derived from another culture. This new map will be more versatile because the mapping population, like a recombinant inbred population, can be propagated eternally in many different environments, facilitating additions of markers and studies of traits with low heritability by many workers simultaneously.

Acknowledgements. This research was supported by a grant from the Rockefeller Foundation. Thanks to S. David and J. Prince for typing and editing the manuscript, respectively, and to F. Loaiza, J. Chen, and T. Fulton for technical assistance. We also wish to acknowledge S. Lincoln and E. Lander for their invaluable help with MAPMAKER.

References

Allard RW (1956) Formulas and tables to facilitate the calculation of recombination values in heredity. Hilgardia 24:235–278

Barker D, Schafer M, White R (1984) Restriction sites containing CpG show a higher frequency of polymorphism in human DNA. Cell 36:131–138

Beltz G, Jacobs K, Eickbush T, Cheerbas P, Kafatos F (1983) Isolation of multigene families and determination of homologies by filter hybridization methods. Methods Enzymol 100:266–285

Bennett MD, Smith JB (1976) Nuclear DNA amounts in angiosperms. Philos Trans R Soc London, Ser B 274:227–274

Bernatzky R, Tanksley SD (1986a) Toward a saturated linkage map in tomato based on isozymes and random cDNA sequences. Genetics 112:887–898

Bernatzky R, Tanksley SD (1986b) Majority of random cDNA clones correspond to single loci in the tomato genome. Mol Gen Genet 203:8–14

Bonierbale M, Plaisted R, Tanksley SD (1988) Construction of comparative genetic maps of potato and tomato based on a common set of cloned sequences. Genetics (in press)

Burr B, Evola SV, Burr RA, Beckman JS (1983) The application of restriction fragment lenth polymorphism to plant breeding. In: Setlow SK, Hollaender A (eds). Genetic engineering: principles and methods, vol 5. Plenum Press, New York, pp 45–59

Burr B, Burr RA, Thompson KH, Albertsen MC, Stuber CW (1988) Gene mapping with recombinant inbreds in maize. Genetics 118:519–526

Coffman WR Bienvenido OJ (1987) Rice. Nutritional quality of cereal grains: genetic and agronomic improvement. Agronomy Monograph No. 28, ASA-CSSA-SSSA, Madison, WI, USA, pp 101–131

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

Doerfler W (1983) DNA methylation and gene activity. Annu Rev Biochem 52:93–124

Feinberg AP, Vogelstein B (1984) A technique for radiolabelling DNA restriction fragments to a high specific activity. Anal Biochem 132:6–13

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

Galbraith DW, Harkins KR, Maddox JM, Ayres NM, Sharma D, Firoozabady E (1983) Rapid flow cytometry analysis of the cell cycle in intact plant tissue. Science 220:1049–1051

Gerlach WL, Bedbrook JR (1979) Cloning and characterization of ribosomal RNA genes from wheat and barley. Nucleic Acids Res 7:1869–1885

Gerlach WL, Dyer TA (1980) Sequence organization of the repeating units in the nucleus of wheat which contain 5s rRNA genes. Nucleic Acids Res 8:1851–1865

Helentjaris T (1987) A genetic linkage map for maize based on RFLP's. Trends Genet 3:217–221

Helentjaris T, King G, Slocum M, Siedenstrang C, Wegman S (1985) Restriction fragment polymorphisms as probes for plant diversity and their development as tools for applied plant breeding. Plant Mol Biol 5:109–118

Herman RK, Shaw JE (1987) The transposable genetic element in the nematode *Caenorhabdibs elegans*. Trends Genet 3:222–225

Iyengar GAS, Sen SK (1978) Nuclear DNA content of several wild and cultivated *Oryza* species. Environ Exp Bot 18:219–224

Khush GS, Singh RJ (1986) Relationship between linkage groups and cytologically identifiable chromosome of rice. In: Rice genetics, Proc Int Rice Genet Symp, May 1985. Los Banos, Philippines, pp 239–248

Khus GS, Singh RJ, Sur SC, Librojo AL (1984) Primary trisomics of rice: Origin, morphology, cytology, and use in linkage mapping. Genetics 107:141–163

Kinoshita T (1986) Standardization of gene symbols and linkage maps in rice. In Rice Genetics, Proc Int Rice Gen Symp, May 1985, IRRI, Los Banos, Philippines, pp 215–228

Kosambi DD (1944) Ann Eugen 12:172–175

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Landry BS, Kessell R, Leung H, Michelmore RW (1987) Comparison of restriction enconucleases and sources of probes for their efficiency in detecting restriction fragment length polymorphisms in lettuce (Lactuca sativa L.). Theor Appl Genet 74:646–653

Lorz H, Gobel E (1986) In vitro culture of rice and direct gene transfer to cereal cells. In: Rice genetics, Proc Int Rice Genetics Symp (May 1985). IRRI, Los Banos, Philippines, pp 849–857

Maniatis T, Fritsch E, Sambrook J (1982) Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory Press Cold Spring Harbor, NY

Nelson M, McClelland M (1987) The effect of site-specific methylation on restriction-modification enzymes. Nucleic Acids Res 15 Suppl: r219–230

Orkin SH (1986) Reverse genetics and human disease. Cell 47:845–850

Paterson AH, Lander ES, Hewitt ID, Peterson S, Lincoln SE, Tanksley SD (1988) Resolution of quantitative traits into Mendelian factors by using a complete RFLP linkage map. Nature (in press)

Poehlman JM (1979) Breeding field crops, 2nd edn AVI Publishing Co, New York

Ranjhan S, Glaszmann JL, Ramirez DA, Khush GS (1988) Chromosomal location of four isozyme loci by trisomic analysis in rice (*Oryza sativa* L.). Theor Appl Genet 75:541–545

Rigby P, Dieckmann M, Rhodes C, Berg P (1977) Labeling deoxyribonucleic acid to high specific activity in vitro by nick-translation with DNA polymerase I. J Mol Biol 113:237–251

Schwarz-Sommer Zsuzsanna, Gierl A, Cuypers H, Peterson PA, Saedler H (1985) Plant transposable elements generate the DNA sequence diversity needed in evolution. EMBO J 4:591–597

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

Shastry SVS, Ranga Rao DR, Misra RN (1960) Pachytene analysis in *Oryza* – 1. Chromosome morphology in *Oryza sativa*. Indian J Genet 20:15–21

Shaw CR, Prasad R (1970) Starch gel electrophoresis of enzymes – a complication of recipes. Biochem Genet 4:297–320

Shure M, Wessler S, Federoff N (1983) Molecular identification of the *Waxy* locus in maize. Cell 35:225–233

Tanksley SD (1982) Molecular markers in plant breeding. Plant Mol Biol Rep 1:3–8

Tanksley SD, Miller J, Paterson A, Bernatzky R (1987) Molecular mapping of plant chromosomes. In: Gustafson JP, Appels RA (eds) Chromosome structure and function. Plenum Press, New York, pp 157–173

Uchimiya H, Fushimi T, Hashimoto H, Harada H, Syono K, Sugawara Y (1986) Expression of a foreign gene in callus derived from DNA-treated protoplasts of rice (*Oryza sativa* L.). Mol Gen Genet 204:204–207

Vallejos E (1983) Enzyme activity staining. In: Tanksley SD, Orton TJ (eds) Isozymes in plant genetics and breeding, part A. Elsevier, Amsterdam, pp 469–516

Wilimzig (1985) LiCl-boiling method for plasmid minio-preps. Trends Genet 1:158

Wu KS, Glaszmann JC, Khush GS (1988) Chromosomal location of ten isozyme loci in rice (*Oryza sativa* L.) through trisomic analysis. Biochem Genet 26:311–328

Yamada Y, Yang ZQ, Tang DT (1985) Regeneration of rice plants from protoplasts. Rice Genet Newsl 2:94–95

Young ND, Miller JC, Tanksley SD (1987) Rapid chromosomal assignment of multiple genomic clones in tomato using primary trisomics. Nucleic Acids Res 15:9339–9348

Zamir D Tanksley SD (1988) Tomato genome is comparised mainly of fast evolving single copy sequences. Mol Gen Genet 213:254–261