Chromosome location of *Oryza sativa* recombination linkage groups

J. P. GUSTAFSON AND J. E. DILLÉ

U.S. Department of Agriculture Agricultural Research Service, Plant Genetics Research Unit and Plant Science Unit, University of Missouri, Columbia, MO 65211; and Department of Biology, Winthrop University, Rock Hill, SC 29733

Communicated by Peter H. Raven, May 27, 1992

**ABSTRACT**

In situ hybridization, a powerful tool for the molecular cytogenetist, can be used to physically map repetitive, low-copy, and unique DNA sequences in plant chromosomes. With the availability of a recombination map in *Oryza sativa* L. and an improved in situ hybridization technique, this study was designed to establish the relationship between the genetic and physical distances of the rice restriction fragment length polymorphism map. Analysis indicated that considerable variation can exist between genetic and physical maps. A 183-centimorgan linkage map for chromosome 2 covered <50% of the chromosome and did not include the centromere, whereas a 91-centimorgan linkage map for chromosome 1 covered ~80% of the chromosome. The results indicated that there are potential "hot" and "cold" spots of recombination and polymorphisms in rice, which involve both genes and restriction fragment length polymorphisms.

In situ hybridization (ISH), as defined by Gall and Purdie (1), cytologically locates labeled DNA to chromosomal sites. It was developed as a technique with the capability of physically mapping both repetitive and low-copy DNA sequences but has been modified since its introduction in order to detect unique-sequence DNA probes. The technique originally used autoradiographic labeling to map both highly repeated DNA sequences (2, 3) and unique-copy DNA sequences (4). However, problems associated with radioactive labels—i.e., short half-life, disposal and safety problems, and long exposure times—have hindered the widespread use of DNA hybridization in molecular mapping. Even though radio-labeled DNA probes provided considerable sensitivity for detecting target sequences, the amount of nontargeted signal has been large and has created some difficulty in identifying individuals hybridization events.

Harper and Saunders (5) significantly improved the ISH technique, enabling the detection of unique-sequence DNA probes on human chromosomes. The sensitivity was further improved such that sequences as small as 0.5 kilobases (kb) can now be detected (6). Unfortunately, use of the technique has been mainly limited to analyses of various animal species. Studies on plant species have severely lagged behind for several reasons. (i) Plants contain cell walls, which present a major problem in obtaining quality chromosome preparations; in addition, debris remains on the slides, which interferes with hybridization. (ii) Plant cytoplasmic debris can contribute to nonspecific binding of the probe (7). (iii) Plant chromosome condensation can vary during metaphase and is more pronounced than in animals. This variability can affect the accuracy of an analysis. (iv) The general scarcity of bands in plants as compared with those seen in mammalian G-band and the frequency of chromosome polymorphisms make karyotyping difficult. (v) Finally, the presence of low mitotic indices in plant chromosome preparations has limited the hybridization of low-copy probes (8). Success of ISH using autoradiography on low-copy and unique-sequence DNA probes in plants has been recently reported (7, 9–12). This advance was possible through the development of techniques for obtaining protoplast preparations instead of the standard mitotic chromosome root-tip squash preparations (7, 9–11) and meiotic pachytene preparations (12).

Recent advances in ISH techniques have also involved the use of nonradioactive systems. Langer et al. (13) described the synthesis and use of nucleotides with the attachment of the naturally occurring hapten biotin to the pyrimidine ring of deoxyuracil 5'-triphosphate (dUTP). More recently, biotinylated deoxyadenosine 5'-triphosphate (dATP) and deoxyxycytosine 5'-triphosphate (dTTP) analogs have also been described (14). All of the biotinylated analogs can be incorporated into DNA by substitution for the appropriate nucleotide in a nick translation (15) or random hexanucleotide priming reaction (16). The choice of which nonradioactive system to use depends on the sensitivity required, the type of DNA sequence to be analyzed, and the resolution of the signal necessary to obtain results. Of these three constraints, sensitivity has been the major problem. However, sensitivity has been improved through the development of better probe and target amplification methods. Biotin labeling with enzyme-conjugated reporters has been the most successful and widely used nonradioactive label for highly repeated, low-copy, unique-sequence DNA probes (7, 9, 13, 17–19).

Use of naturally occurring hapten, other than biotin, as a source of nonradioactive labeling has also been explored. Of all hapten tested, other than biotin, digoxigenin is the most promising (20, 21). Digoxigenin-dUTP can be incorporated into DNA by either nick translation or the random-primer reaction in the same way as described for biotin labeling. However, digoxigenin currently does not seem to offer any significant advantage over biotin.

Only recently has it become possible to physically map low-copy and unique-sequence DNA probes to plants (7, 9, 17–19) with nonradioactive biotin labeling. Detection of unique-sequence probes has been greatly enhanced by the use of root-tip protoplasts (8) and enzyme-conjugated reporters (7, 9, 18, 19). The production of protoplasts involves cellulose removal, thus increasing hybridization and detection (8, 22). However, the existing techniques for protoplast production might need slight modification for each plant species (18).

For mapping, the aim was to develop an ISH technique that would physically locate a unique-sequence gene on a chromosome. Therefore, resolution determined choice of the detection label. At present, the three main colorimetric detection systems are used to precipitate an insoluble colored product from an enzyme reaction, involving usually alkaline phosphatase or, in our study, horseradish peroxidase. (i) Bromochloroindolyl phosphate (BCIP) changes to indigo upon dephosphorylation and subsequent oxidation with al-

Abbreviations: ISH, in situ hybridization; RFLP, restriction fragment length polymorphism(s); cm, centimorgan(s); RAPD, randomly amplified polymorphic DNA.
kaline phosphatase. This compound is coupled to the reduc-
tion of a nitroblue tetrazolium salt to formazan (23). This
method was undesirable in our study because the rice chro-
mosomes were counterstained with Giemsa stain, which also
results in bluish-colored chromosomes. (ii) 3',5',5'-
Tetramethylbenzidine (TMB) is an alternative substrate with
horseradish peroxidase, but it produces a semisoluble blue
product. (iii) The polymerization of 1,2-diaminobenzidine
(DAB), in the presence of hydrogen peroxide and horseradish
peroxidase, results in a brown precipitate. The latter is
currently the most commonly used reaction on plants be-
cause it causes a color label different from that used to
counterstain chromosomes.

Numerous reports have suggested that genetic distances
seen on a genetic map differ greatly from those seen on a
physical map (18, 19, 24–28). Ganal et al. (28) estimated that
the tomato (Lycopersicum esculentum L.) Tm-2a gene
showed approximately a 3-fold decrease in recombination
based on the expected value of the estimated physical size of
the region. Chromosomal regions containing heterochromat-
in and regions near a centromere can suppress recombin-
ation, thus giving the illusion of very little spatial distance
between genes (18, 19, 24–26, 28–30). Estimates in DNA
amount per centimorgan (cM, unit of recombination measur-
genetic distance) can vary tremendously between and
within a genome. There are even genetic- and physical-
distance differences within areas on a single chromosome (19,
30–32). The intrachromosomal differences could be mainly
from heterochromatin and the proximity of the marker to the
centromere. In addition, inversions, translocations, and
other anomalies of chromosomes can affect recombination
(33). Meagher et al. (32) suggested that assuming the average
kb/cM ratio for any particular species would apply broadly
for all loci would create very large errors in estimating
physical distances.

With the creation of restriction fragment length polymor-
phism (RFLP) linkage maps in several crop plants (34–42)
and the use of nonradioactive ISH techniques, the extent to
which these linkage maps cover the physical length of the
genome can be determined. There are indications from look-
ing at linkage maps that tremendous differences could occur
between physical and genetic distances, as determined by
recombination map units, regardless of the presence or
absence of a centromere (18, 19, 31–33, 43, 44).

There are several ways to physically map molecular mark-
ers and genes to chromosomes and analyze the differences
between genetic and physical distances. (i) One could employ
chromosome “walking” using pulsed-field gel electropho-
sis and DNA sequencing (28) or pulsed-field gel electro-
phoresis without DNA sequencing (45). Coupled with DNA
sequencing, pulsed-field gel electrophoresis is the most ac-
curate way for determining the physical location of one gene
to another. However, unless a telomere or a centromere is
involved, the two linked markers are located only relative
to each other, and they cannot be physically located along the
centromere. Another problem with this procedure is that
several megabases of DNA might be searched before locating
the desired gene. Also, a completely saturated linkage map is
critical to “walking” easily to the desired gene. At present,
saturated linkage maps for mapping are unavailable for the
cereals. However, saturated linkage maps for specific chro-
mosomal regions that can be used for chromosome walking are
being created. (ii) One could use molecular markers with deletion stocks where the involved chro-
mosomes contain deletions or translocations (46, 47). The mark-
ers or genes could then be mapped to specific deletions. This
procedure is also very time-consuming because the stocks
must be created before any genes or markers can be mapped,
and a complete set of deletion stocks that encompass an
entire genome could probably not be created because many
deletions will be lethal. In addition, the deletions or translo-
cations must be large enough to be visualized, which means
that the probe location is known only on a gross level. It is
also doubtful that deletion stocks could be created for many
diploid plant species because deletions in diploids would
usually cause the loss of genes required for plant function,
thus reducing viability of the plant. Polyploids differ because
the genetic buffering from the other genomes in the polyploid
allows survival with some genetic deletions. A saturated
linkage map would be necessary to precisely map the dele-
tions. A good chromosomal-banding-pattern karyotype of the
species involved would also be required to screen for and
detect the deletions. (iii) One could physically map genes in
species where stocks with known chromosomal rearrange-
ments (i.e., translocations) exist (48). (iv) ISH techniques
could be used to physically map the genes or markers to
locations on the chromosomes. This method appears the best
approach at present because probes can be mapped directly
to chromosomes of any germ-plasm base, variety, or species
without any stock-creation work. In addition, the technique
requires less time as compared with other methods. The
drawback to this procedure is the extremely low detection level (=5% with unique-sequence probes) (19).

Areas of almost no recombination, “cold spots,” and areas of
high recombination, “hot spots,” probably exist in almost
all species of animals and plants. These areas can be detected
by the gaps and clustering of markers, respectively, observed
in the numerous existing linkage maps. Recently published
plant linkage maps indicate that even the RFLP markers (37,
40) are clustering in various species. Randomly amplified
polyorphic DNA (RAPD) markers may also be involved in
clustering. With the differences noted between genetic and
physical distances, physical mapping will be necessary to
accurately locate cloned loci on chromosomes. Our study
was undertaken to ascertain whether the ends of the Oryza
sativa L. linkage groups could be physically mapped to
locations on chromosomes, thus allowing comparison of the
physical and genetic maps.

MATERIALS AND METHODS

Plants. O. sativa L. cv. IR36 supplied by G. Khush (the
International Rice Research Institute, Manila, Philippines)
was used as the varietal source of chromosomes in all
analyses.

DNA Sequences. The genomic DNA clones used were those
on the map of McCouch et al. (37) and were supplied by S. D.
Tanksley (Cornell University, Ithaca, NY). All clones were
mapped to specific chromosomes by trisomic analysis (37).
The genomic DNA clones ranged in size from 3.4–0.7 kb and
were present in plastid pUC19.

Chromosome Preparation. The protoplast technique was
that developed by Dillé et al. (8) for prometaphase and
metaphase chromosome preparations.

Biotin Labeling of Plasmids, in Situ Hybridization, Detec-
tion, and Visualization of Hybridized Probes. The technique
was that of Rayburn and Gill (49), which was subse-
quently modified by Gustafson et al. (18). Additional modi-
fications to the technique were used here. After counterstain-
ing with Giemsa stain, the slides were immediately placed
into xylene for 3 hr before mounting, thus eliminating air-
drying of the slides before mounting. This modification
seemed to give a better result in a clearer slide image. The slides were
analyzed with a Zeiss Photomicroscope III with a Ham-
matsu enhancement system and a television monitor. All
chromosome measurements were made directly from the
television screen with calipers. An average of the hybridiza-
tion-site measurements was taken by measuring the distance
from the centromere to the detection site and taking that
distance as a percentage of the length of the arm on which the

Genetics: Gustafson and Dillé
hybridization site was located. A percentage measurement was taken because it eliminated the problems associated with chromosome-condensation variation. Standard deviations were calculated for each group of measurements. The arm ratio of the chromosome showing a detection site was also measured, and standard deviations were calculated to determine whether the site was located on the same chromosome. This technique also checked for potential occurrence of any spurious hybridization sites. In addition, an ISH was done on pUC19 alone to determine whether any hybridizations occurred between pUC19 and rice DNA; no hybridizations between pUC19 and rice DNA were seen in >500 cells.

**RESULTS AND DISCUSSION**

The present study physically mapped the linkage groups associated with 8 of the 12 rice chromosomes—1–6, 9, and 10 (Fig. 1). A total of 2443 metaphase spreads was analyzed, and 148 hybridization sites were detected (6.06%) (Fig. 2). All chromosomes mapped, except chromosome 6, were submetacentric, and the arm ratios could easily be distinguished. Chromosome 6 is the largest metacentric chromosome in the genome. Arm ratios of the chromosomes analyzed and RFLP location of the probes to the ends of the linkage groups clearly indicated that the measurements could separate and identify

---

**Fig. 1.** A karyotype of the recombination linkage map for *O. sativa* from McCouch et al. (37) and the corresponding chromosomes showing physical location of the linkage groups. Shaded clone numbers mark the ones physically mapped.
the chromosomes and probe locations (Table 1). Even when the probes were very close together (i.e., RG358 and RG386 on chromosome 9), there was little problem in assessing the physical order of the probes from the centromere. The remainder of the chromosomes are metacentric and very small in size (Fig. 1), which makes them extremely difficult to map. The large linkage groups mapped to chromosomes 1 and 5 were the only ones analyzed that spanned a centromere (Fig. 1). All other linkage groups were either completely on the long or completely on the short arm of the chromosome to which they were mapped. The linkage group for chromosome 9 contained six RFLPs covering 66 cM, whereas the linkage group on chromosome 10 contained only three RFLPs covering 43 cM (Fig. 1). Interestingly, the physical distance covered by both linkage groups appeared similar. The largest linkage group, consisting of 21 RFLPs covering 183 cM, was located on the short arm of chromosome 2. The chromosome 2 linkage group was notable in that it showed a significant difference when comparing the physical distances between RFLPs RG322 and RG139 (56 cM) vs. RG139 and RG83 (127 cM). This difference clearly suggests that the largest cM distance was not necessarily the largest in terms of physical distance covered on the chromosome. These differences could be from (i) the presence or absence of blocks of heterochromatin within one of the linkage groups, (ii) the presence of a cold spot of recombination within the linkage group between RG139 and RG83, (iii) Dillé et al. (8), Kurata (50), Fukui and Iijima (51), and others have shown that during mitosis, rice chromosomes appear to condense differently from one part of the chromosome to another during the early stages of mitosis (as shown in Fig. 2B). Perhaps different levels of recombination exist because of the various areas of condensation differences. The real answer remains unknown. These physical-mapping examples illustrate the problems that can arise when comparing cM distances even within a single chromosome. The cM/kb ratio for different locations along yeast (Saccharomyces cerevisiae) chromosome III has recently been shown to vary up to 10-fold; this ratio is lowest in the centromeric region and is highest midway in the arms (52).

Three of the RFLPs (RG136, RG246, and RG147) appear telomeric in their physical location (Fig. 1). However, it would be a mistake to conclude that they are telomeric because a biotin detection label can be large enough to cover a million base pairs. The physical location on the present map fairly estimates the actual distance, and considerable chromosome walking and sequencing is needed before one could ascertain the true location of the probes in question in their relationship to the telomere or centromere.

It was interesting to find that most of the detection sites were present as only a single spot instead of both chromatids being detected as two spots. As indicated, the detection level was ~6%; therefore, to detect both chromatids concomitantly, the percentage would be 6% of 6%. This product is a very low percentage, and only rarely was detection seen on both chromatids. Labeling of only one chromatid has been frequently seen in animal and plant studies (17-19, 53). When one looks at the animal and plant papers showing detection on both chromatids, the probe in question is generally 7-10 kb or larger in size or is moderately-to-highly repeated (3, 54, 55).

Analysis of several rice RFLP linkage groups indicates that only a limited portion of the physical length of the rice genome is involved. This observation should not be unexpected, as several reports have suggested that the recombination location of genes vary considerably from their physical location (18, 19, 24-26, 29). What is interesting, however, is that RFLP polymorphisms appear capable of clustering, as found with morphological marker groups (37-42, 47). It appears possible that evenRAPDs will cluster. It would be interesting if the variations seen in the clustering of morphological markers, RFLPs, and RAPDs also depended on the

| Table 1. Arm ratios and RFLP location averages for the chromosomes in the rice genome |
|--------------------------------------|------------|-----------------|-----------------|
| RFLP | Detections, no. | Ratio of long-to-short arm ± SD | RFLP location, % |
| Chromosome 1 | | | |
| RG233 | 7 | 1.30 ± 0.094 | 62.80 ± 0.10 LA |
| RG316 | 5 | 1.35 ± 0.057 | 85.10 ± 0.07 SA |
| RG246 | 4 | 1.35 ± 0.057 | ~100.00 LA |
| Chromosome 2 | | | |
| RG83 | 13 | 1.37 ± 0.075 | 27.50 ± 0.10 SA |
| RG139 | 12 | 1.37 ± 0.065 | 42.80 ± 0.08 SA |
| RG322 | 11 | 1.38 ± 0.061 | 88.70 ± 0.07 SA |
| Chromosome 3 | | | |
| RG179 | 2 | 2.67 ± 0.36 | 52.40 ± 1.84 LA |
| RG227 | 5 | 2.81 ± 0.13 | 66.80 ± 13.4 LA |
| RG104 | 5 | 3.22 ± 0.80 | 91.80 ± 2.86 LA |
| Chromosome 4 | | | |
| RG163 | 3 | 1.54 ± 0.04 | 56.00 ± 1.65 LA |
| RG143 | 4 | 1.53 ± 0.01 | 70.00 ± 4.97 LA |
| RG91 | 2 | 1.50 ± 0.10 | 83.95 ± 5.44 LA |
| Chromosome 5 | | | |
| RG113 | 7 | 1.78 ± 0.08 | 33.20 ± 10.82 SA |
| RG207 | 7 | 1.78 ± 0.07 | 16.27 ± 7.17 LA |
| RG182 | 6 | 1.77 ± 0.07 | 39.82 ± 5.01 LA |
| RG344 | 8 | 1.82 ± 0.08 | 94.75 ± 5.92 LA |
| Chromosome 6 | | | |
| RG213 | 5 | 1.08 ± 0.05 | 18.27 ± 8.67 LA |
| RG147 | 10 | 1.03 ± 0.03 | ~100.00 LA |
| Chromosome 9 | | | |
| RG358 | 8 | 1.86 ± 0.04 | 79.56 ± 8.10 LA |
| RG386 | 4 | 1.86 ± 0.13 | 81.50 ± 2.89 LA |
| RG113 | 7 | 1.97 ± 0.06 | ~100.00 LA |
| Chromosome 10 | | | |
| RG241 | 5 | 2.29 ± 0.11 | 6.80 ± 2.39 LA |
| RG134 | 8 | 2.42 ± 0.19 | 7.75 ± 5.04 LA |

SA, short arm; LA, long arm.
*Percent distance from the centromere.
parents involved in creating the mapping populations. Only considerable mapping work on additional genetically different mapping populations can provide this answer. The exact physical location of any gene, RFLP, or RAPD, in reality, could depend on the individual organism (animal or plant) that donated the chromosomes or DNA to be sequenced. This dependence could be from the possibility of deletions, translocations, inversions, duplications, or insertions occurring between parents that would change the gene, RFLP, or RAPD location, as well as the physical location of hot and cold spots of recombination. The development of more markers and mapping populations will tell whether the genetic hot spots and cold spots of recombination also correspond to the clusters or gaps involving RFLPs and RAPDs and whether these occurrences are consistent between organisms. This information would have practical implications concerning the amount of usable genetic variation available to a plant breeder. If significant amounts of DNA within the genome of a species do not recombine, then the usable gene base (i.e., genetic variability) would be much smaller than envisioned. Any genes in these regions (gaps or cold spots) must have an extremely high selection pressure on them to remain homozygous, and any mutations causing variation would be at an extreme selective disadvantage. Maybe these regions contain housekeeping genes that, for survival of the organism, cannot tolerate any changes and that have evolved and concentrated into regions not commonly involved in recombination.

We express our appreciation for the support of the late Dr. E. R. Sears throughout this research. This research was supported by Rockefeller Foundation Grant RF 86059-3. This paper is a contribution of the U.S. Department of Agriculture, Agricultural Research Service, and Missouri Agricultural Experiment Station, Journal Ser. No. 11,633.