

High-density linkage map of rice with expressed sequence tags

B.A. Antonio, A. Shomura, T. Shimano, Y. Kuboki, S.Y. Lin, T. Inoue, H. Kajiya, Y. Harushima, K. Yamamoto, Y. Nagamura, N. Kurata, M. Yano, Y. Minobe, and T. Sasaki

We have constructed a high-density linkage map of rice using an F_2 population derived from the cross between a japonica variety, Nipponbare, and an indica variety, Kasalath. A total of 1,383 markers, which consist of cDNA clones from callus and root, genomic clones as well as RAPD markers, have been mapped covering a distance of 1,575 cM. All cDNA clones have been sequenced and searched for similarities with known proteins and can be referred to as expressed sequence tags on the map. A majority of the genomic clones and RAPD markers was also sequenced to generate sequence-tagged sites. These extensive linkage analyses gave evidence on duplication of chromosomal segments, particularly in the distal region of chromosomes 11 and 12. Additional markers are being mapped using cDNA clones derived from other cDNA libraries such as green shoot, etiolated shoot, and developing seed. Ultimately, we would like to develop a saturated linkage map that will facilitate a more efficient utilization of molecular markers for rice improvement.

One of the most important advances in the field of biotechnology, which promises to revolutionize several areas of plant genetics and breeding, is the wide utilization of molecular markers. In conjunction with phenotypic and biochemical markers, these markers will have great impact in identifying and ultimately isolating genes for various agronomically important traits. In recent years, construction of RFLP linkage maps has been reported in a number of plants (Bernatzky and Tanksley 1986, Chang et al 1988, Rognli et al 1992, Da Silva et al 1993, Kleinhofs et al 1993). In rice, a molecular linkage map covering the entire genome was developed independently by McCouch et al (1988) and Saito et al (1991) with 135 and 322 markers, respectively. Such molecular maps may provide new opportunities for application in plant genetic manipulation, particularly in tagging genes for agronomically important traits with DNA markers. In addition, these maps could also serve as important tools in understanding the evolutionary relationships among different species as shown by the

synteny studies between such crops as wheat and rye (Rognli et al 1992), potato and tomato (Tanksley et al 1992), rice and maize (Ahn and Tanksley 1993), rice and wheat (Kurata et al 1994a), etc.

In the Rice Genome Research Program (RGP), we are constructing a high-density linkage map of rice with markers spaced at very close intervals throughout the genome. Most markers in this map have been sequenced to generate expressed sequence tags and sequence-tagged sites (STSs), and as such will be a model system for overall analysis of genome structure and function in plants. So far, a map with 1,383 DNA markers at an average interval of 300 kb and distributed along 1,575 cM on the 12 linkage groups has been reported by Kurata et al (1994b). Mapping of more DNA markers is currently in progress to generate a saturated map. This paper summarizes such results as well as some of the most recent findings in restriction fragment length polymorphism (RFLP) mapping at RGP.

Materials and methods

Plant materials

The parent strains consisted of a japonica variety, Nipponbare, and an indica variety, Kasalath. A single cross was made to obtain an F₂ population and 186 individuals were used for analysis of segregation of DNA polymorphism.

DNA manipulation

Total DNA was extracted from the green leaves of parental lines as well as the F₂ progenies by the CTAB method (Murray and Thompson 1980). Then 2 µg total DNAs were each digested with one of eight restriction enzymes, *Bam*HI, *Bgl*III, *Eco*RV, *Hind*III, *Apa*I, *Dra*I, *Eco*RI, and *Kpn*I, overnight at 37 °C. The digested samples were applied in 0.6% agarose gel, electrophoresed for 12 h and transferred in a positively charged nylon membrane by capillary blotting. These were used for hybridization with probes labeled with horseradish peroxidase according to the protocol of ECL direct nucleic acid labeling and detection system (Amersham).

DNA probes

The probes used for hybridization consisted mainly of cDNA clones, genomic clones, and RAPD markers all derived from japonica cultivar, Nipponbare. The cDNA clones consisted of randomly selected clones from callus and root cDNA libraries. The nucleotide sequence from the 5' end for 300-400 bp was determined and translated into an amino acid sequence. Then a similarity search at the protein level was performed in the NBRF-PIR data base using the FASTA algorithm. Clones showing an optimized matching score of more than 150 with amino acid sequences in other organisms were considered as functionally identical clones. All sequenced clones are registered and deposited at the DNA Data Bank of Japan (DDBJ).

The genomic clones used for mapping consisted of random genomic clones, YAC-end clones, *Not*I linking clones, and telomere-associated sequences (TEs). The random genomic clones were prepared by ligating *Hind*III or *Pst*I DNA fragments in

pBluescriptII SK+ or pUC vector. The YAC-end clones were derived from both ends of a large size DNA fragment cloned in YAC, amplified by PCR as 200-1000 bp long DNA, and ligated into TA cloning vector PCRTM1000. The *NotI* linking clones consisted of *Sau3AI* partially digested 500-4000 bp fragments with *NotI* sites and cloned in pT7T318U vector at the *BamHI* site. The TELs were obtained using cassette ligation-mediated PCR of *Sau3AI* DNA digests and cloned in pCRII vector (Ashikawa et al 1994). For mapping of RAPD markers, 60 arbitrarily designed 10-nucleotide primers were initially subjected to RAPD analysis. Then, these primers were paired randomly and were used for detection of RAPD markers. Detection and mapping of RAPD markers and conversion of RAPD to STS markers were described by Monna et al (1994).

Linkage analysis

The segregation patterns and linkage relationships of RFLP in the F₂ population were analyzed using the MAPMAKER/EXP 3.0 software (Lander et al 1987). Multipoint analysis was performed to calculate the linkage of a large number of markers and produce a map of their order along the chromosomes. Recombination values between the markers were transformed into centimorgan (cM) distance by the Kosambi function (Kosambi 1944).

Results and discussion

RFLP map with 883 expressed sequences

To construct an RFLP linkage map of rice, we analyzed 2,950 cDNA clones from callus and root cDNA library. These clones showed various banding patterns such as single bands, double bands, as well as multiple bands with a smeared background in some cases, suggesting either single-copy sequences or repeated sequences in the genome. A total of 883 cDNA clones, which consisted of 465 clones from callus cDNA and 418 clones from root cDNA, showed distinct RFLP and were used for segregation analyses of the F₂ population derived from the cross Nipponbare/Kasalath. The positions of these clones represented by C-number and R-number for callus and root cDNA clones, respectively, are shown in Figure 1. A more detailed version of this map appeared in Kurata et al (1994b) and included such information as the accession number of the sequence data deposited in the DDBJ. In addition to cDNA clones, 265 genomic DNAs, 147 RAPD markers, and 88 other DNAs were also mapped for a total of 1,383 markers distributed along 1,575 cM on 12 linkage groups at an average interval of 1.14 cM.

A similarity search for proteins of other organisms showed that the cDNA clones have a high similarity to genes of a wide range of organisms including dicots, monocots, mammals, and yeast (Table 1). Most of these genes code for isozymes such as alcohol dehydrogenase (*adh*), aspartate aminotransferase (*got*), fructose bisphosphate aldolase (*ald*), glucose-6-phosphate isomerase (*pgi*), peroxidase (*pox*), etc. In the conventional linkage map, several isozymes have been mapped and assigned to specific chromosomes (Wu et al 1988). In our RFLP linkage map, we determined the loci of

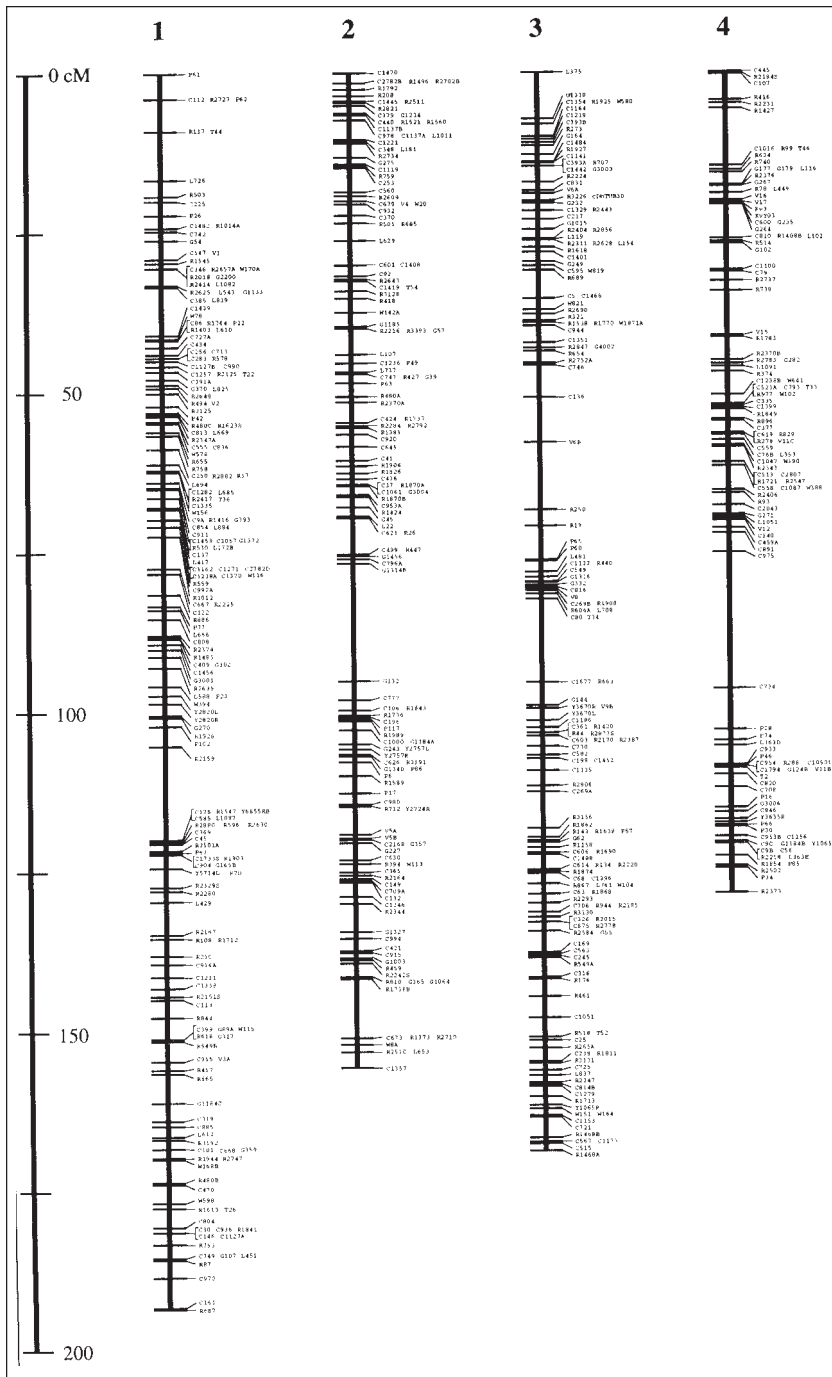


Table 1. Callus and root cDNA clones mapped in rice with similarity to known proteins.

Chr.	Position	Marker	Gene	Protein name	Organism	DBJ ID No.
1	30.4	R2657A	<i>ald2</i>	Fructose biphosphate aldolase	Rice	D28322
1	41.8	R1764	<i>got2</i>	Aspartate aminotransferase	<i>Bacillus sp.</i>	D24345
1	42.9	C727A	<i>gco1</i>	Glucan endo-1,3-beta-glucosidase	Common tobacco	D15500
1	44.5	C256	<i>rsc</i>	Reductase	Soybean	D15203
1	44.5	R578	<i>hbo</i>	(S)-tetrahydroberberine oxidase	<i>Coptis japonica</i>	D23922
1	49.3	R494	<i>nia</i>	Nitrate reductase (NADH)	Tomato	D23879
1	53.1	R1623S	<i>tub3</i>	Tubulin beta-2 chain	<i>Arabidopsis thaliana</i>	D24277
1	58.9	C250	<i>got1</i>	Aspartate aminotransferase	Proso millet	D23735
1	58.9	R37	<i>glt1</i>	Glutathione transferase 1	Maize	D32736
1	63.6	C9A	<i>elf3</i>	Elongation factor 2	<i>Caenorhabditis elegance</i>	D15078
1	64.5	C911	<i>gtl</i>	Glutamin:tRNA ligase	Human	D15594
1	69.2	R559	<i>ppp</i>	Phosphoprotein phosphatase	Human	D23910
1	70.3	C922A	<i>gbp</i>	GTP-binding regulatory protein beta chain	<i>Chlamydomonas reinhardtii</i>	D22667
1	71.9	R1012	<i>lcl</i>	Long-chain-acid:CoA ligase	Human	D24049
1	81	R886	<i>mdh</i>	Malate dehydrogenase, mitochondrial	Water melon	D24025
1	84.7	C808	<i>elf2</i>	Initiation factor eIF-4A	Curled-leaved tobacco	D22665
1	87.7	C409	<i>sip</i>	Stress inducible protein STI1	Yeast	D15287
1	90.4	R2635	<i>soi</i>	Spil hypothetical protein	Yeast	D24836
1	99.9	R1928	<i>vcp</i>	Vaseline-containing protein	Pig	D28306
1	119.1	C585	<i>secl</i>	SEC 7 protein	Yeast	D15403
1	119.4	R2630	<i>hud</i>	Elav/Sex-lethal related protein	Human	D24832
1	119.4	R596	<i>glt2</i>	Glutathione transferase 1	Maize	D28287
1	119.4	R2880	<i>osb</i>	Oxysterol-binding protein	Rabbit	D24980
1	119.5	C369	<i>gdh</i>	Glutamate dehydrogenase (NAD(P)+)	<i>Halobacterium salinarium</i>	D15259
1	121.3	C904	<i>sall</i>	SalT protein precursor	Rice	D28208
1	126.4	R476	<i>ams1</i>	S-adenosylmethionine synthetase 2	<i>Arabidopsis thaliana</i>	D28266
1	126.9	R2280	<i>ams4</i>	S-adenosylmethionine synthetase 2	<i>Arabidopsis thaliana</i>	D24629
1	133.7	R2167	<i>ams3</i>	S-adenosylmethionine synthetase 2	<i>Arabidopsis thaliana</i>	D28314
1	137	R210	<i>cad1</i>	Cathepsin D	Human	D23806
1	142.1	C1338	<i>ang</i>	58K antigen	<i>Rickettsia tsutsugamushi</i>	D22792
1	149.6	C399	<i>idh</i>	Isocitrate dehydrogenase (NADP+)	Alfalfa	D15280
1	155.3	R665	<i>rac1</i>	Rac1 protein	Human	D23963
1	165.7	R3192	<i>spk</i>	Serine/threonine-specific protein kinase	<i>Arabidopsis thaliana</i>	D25110
1	172.4	R480B	<i>ypt</i>	Transforming protein, ypt 1, homolog	Maize	D23874
1	180.3	C936	<i>mtn</i>	Metallothionein-like protein	<i>Arabidopsis thaliana</i>	D15602
1	180.3	C30	<i>tpi</i>	Triose phosphate isomerase	Maize	D15092
1	181.6	R753	<i>sds</i>	C-5 sterol desaturase	Yeast	D23996
1	184.1	R87	<i>tin</i>	Trypsin inhibitor	Rice	D23762

Table 1 continued.

Chr.	Position	Marker	Gene	Protein name	Organism	DBJ ID No.
2	1.6	R2702B	<i>hsp5</i>	Heat shock protein 70	Common tobacco	D23418
2	4.4	C1445	<i>aux</i>	Auxin-induced protein	<i>Arabidopsis thaliana</i>	D15870
2	6.3	C440	<i>dfc1</i>	Dihydrofolate-4-reductase	Garden petunia	D15312
2	6.6	C1137B	<i>dfc2</i>	Dihydrofolate-4-reductase	Garden snapdragon	D15715
2	7.4	C1137A	<i>dfc2</i>	Dihydrofolate-4-reductase	Garden snapdragon	D15715
2	32.3	C92	<i>ant</i>	Adenine nucleotide translocator	Rice	D22519
2	33.1	C1419	<i>thr</i>	Thioredoxin reductase (NADPH)	<i>Escherichia coli</i>	D13855
2	34.6	R3128	<i>eno2</i>	Enolase	Tomato	D25085
2	40.5	R3393	<i>clc</i>	Clathrincoat assembly protein	Rat	D24586
2	51.2	R480A	<i>ypt</i>	Transforming protein, ypt1, homolog	Maize	D23874
2	55.3	R1737	<i>prs</i>	Proteasome XC3 chain	African clawed frog	D24326
2	55.8	R2284	<i>ams5</i>	S-adenosylmethionine synthetase 2	<i>Arabidopsis thaliana</i>	D24632
2	63.3	R1826	<i>nab</i>	X16 protein	Mouse	D24389
2	65	C37	<i>gpd1</i>	Glyceraldehyde-3-phosphate dehydrogenase	White mustard	D15096
2	67	R1424	<i>ste1</i>	Regulatory protein STE7	Yeast	D24144
2	70.3	C621	<i>reg1</i>	14-3-3 protein	Barley	D15430
2	75.9	R447	<i>sac</i>	SAC1 protein	Yeast	D23860
2	103.4	C1000	<i>hsp3</i>	Heat shock protein 70	Maize	D15636
2	107.2	C626	<i>cyc</i>	cyc07 protein, S-phase specific periwinkle	Madagascar	D15433
2	120.6	C2168	<i>got3</i>	Aspartate aminotransferase	Proso millet	D16037
2	138.2	C915	<i>stk</i>	Kinase-related transforming protein	Mouse	D15597
2	139.3	R459	<i>gdc2</i>	Glycine-cleavage system protein H	Garden pea	D23865
2	139.8	R2242S	<i>tub4</i>	Tubulin beta-2 chain	Garden pea	D24606
2	142	R810	<i>ubq4</i>	Ubiquitin	Garden snapdragon	D25349 D25350
2	151.4	R2710	<i>urt2</i>	UTP:glucose-1-phosphate uridylyltransferase	Potato	D24887
3	14.7	R707	<i>qpc</i>	Ubiquinone binding protein QP-C	Bovine	D23977
3	18.5	C831	<i>rad6</i>	RAD6 DNA-repair homolog <i>Dhr6</i>	Fruit fly	D22670
3	20.1	R3226	<i>cof</i>	Cofilin	Yeast	D25113
3	21.7	R2443	<i>myb</i>	Transforming protein, myb, homolog	Maize	D24724
3	21.7	C1329	<i>pgi</i>	Glucose-6-phosphate isomerase	<i>Clarkia lewesii</i>	D15815
3	26.1	R2856	<i>cak</i>	Casein kinase II alpha chain	Maize	D24965
3	26.1	R2404	<i>EIF4</i>	Initiation factor eIF-5A	Common tobacco	D24702
3	26.3	R2628	<i>tpa</i>	Transplantation antigen P198	Mouse	D24830
3	35.4	C1468	<i>tub2</i>	Tubulin alpha-2 chain	Maize	D15886
3	37.9	R2690	<i>act</i>	Actin 1	Rice	D24576
3	39.2	R1538	<i>reg4</i>	14-3-3 protein	Barley	D24218
3	43.2	R2847	<i>gco2</i>	Beta-glucosidase	White clover	D24959
3	45.9	C746	<i>gri</i>	Glycine rich protein 2	<i>Arabidopsis thaliana</i>	D15512
3	79.7	C549	<i>hsp1</i>	Heat shock protein 70	Spinach	D22613

Table 1 continued.

Chr.	Position	Marker	Gene	Protein name	Organism	DDBJ ID No.
3	81.2	R1908	<i>acb</i>	Endozepine	Yeast	D28303
3	103.7	R2170	<i>uqn</i>	NADH dehydrogenase (ubiquinone) chain 2	<i>Paramecium tetraurelia</i>	D28315
3	107.2	C1452	<i>sod</i>	Superoxide dismutase	Rice	D15675
3	119.6	R1862	<i>prp</i>	Prp 16-1 protein	Yeast	D24417
3	121.2	R1158	<i>snr</i>	Small nuclear RNA-associated protein	Human	D24080
3	122	R1690	<i>elf3</i>	Initiation factor 2 alpha chain	Yeast	D24301
3	128.5	C63	<i>ubq1</i>	Ubiquitin fusion protein	Fruit fly	D15108
3	134.2	R2584	<i>cdh</i>	Cinnamyl-alcohol dehydrogenase	Kidney bean	D14802
3	150.7	R518	<i>elf1</i>	Elongation factor 1 alpha	Tomato	
3	160	R1713	<i>glt3</i>	Glutathione transferase III	Maize	D24311
3	166.5	R1468A	<i>cdc</i>	CDC2a protein	Rice	D24174
4	4.7	R416	<i>aox</i>	Amine oxidase	Rat	D23854
4	15.5	R634	<i>ocp</i>	Oryzain alpha chain	Rice	D23944
4	16	R740	<i>gyk</i>	Glycerol kinase	<i>Bacillus subtilis</i>	D23993
4	19.3	R78	<i>kin</i>	ncdD protein	Fruitfly	D23757
4	53.3	R1849	<i>art</i>	Arabinose transport protein	<i>Escherichia coli</i>	D24407
4	54.6	R896	<i>gpd2</i>	Glyceraldehyde-3-phosphate dehydrogenase	Maize	D28294
4	57.4	C559	<i>ppa</i>	Inorganic pyrophosphatase	Yeast	D15382
4	59	C1047	<i>reg3</i>	14-3-3 protein	Barley	D15663
4	109.2	R288	<i>ccp</i>	Cytochrome C peroxidase	Yeast	D23832
4	109.2	C954	<i>dds</i>	Dihydrodipicolinatesynthase	Wheat	D15614
4	109.2	C1794	<i>his1</i>	Histone H1	Wheat	D22924
4	121.3	C9B	<i>elf3</i>	Elongation factor 2	<i>Caenorhabditis elegance</i>	D15078
5	27.9	R1838	<i>dnj</i>	dnaJ protein homolog	Human	D24399
5	30.9	C259B	<i>ubq2</i>	Ubiquitin	Tomato, potato, oat	D22550
5	45	R569	<i>omc</i>	2-oxoglutarate/malate carrier protein	Bovine	D23915
5	55.5	R2059	<i>rbp</i>	Ribophorin	Human	D24495
5	55.5	C1388	<i>rab11</i>	GTP-binding protein rab11	Dog	D15842
5	55.5	R2558	<i>acc</i>	Acetyl-CoA carboxylase	Yeast	D24786
5	95.2	R3182	<i>hsp6</i>	Heat shock protein cognate 70	Tomato	D25105
5	95.2	C128	<i>ubc</i>	Ubiquitin conjugating protein	Wheat	D15130
5	96.8	C536	<i>pcd</i>	Pyruvate decarboxylase	Maize	D15369
5	102.2	C67B	<i>rif</i>	ADP-ribosylation factor 4	Human	D22513
5	102.2	C419	<i>cam</i>	Calmodulin	Wheat	D15295
5	109	C466	<i>mpp</i>	Processing peptidase catalytic chain, mitochondrial	Yeast	D15329
5	113.4	C686	<i>atp1</i>	H ⁺ -transporting ATP synthase beta chain	Rice	D15470
5	113.7	R2953	<i>dyl</i>	Dynammin-like protein	Fruit fly	D25026
5	118	R2754	<i>cad2</i>	Cathepsin D	Human	D24912
5	119.6	C1264	<i>kri</i>	Ketol-acid reductoisomerase chloroplast	Spinach	D27768
6	2.2	R2869	<i>pgd</i>	Phosphogluconate dehydrogenase	<i>Synechococcus</i> sp.	D24970
6	9.2	C688	<i>prt</i>	Transcription factor for E3	Human	D15472
6	9.8	R2291	<i>ste2</i>	Regulatory protein STE7	Yeast	D24636
6	10.1	R2749	<i>cys</i>	Cysteine synthase B	Pepper	D24907

Table 1 continued.

Chr.	Position	Marker	Gene	Protein name	Organism	DDBJ ID No.
6	11.2	C764	<i>hca</i>	ClassII histocompatibility antigen	Human	D15525
6	12.6	C1032	<i>ag12</i>	Floral homeotic protein AGL2	<i>Arabidopsis thaliana</i>	D15657
6	13.1	R845	<i>ctf</i>	Cystathionine gamma-lyase	Yeast	D28293
6	17.9	R1966	<i>sus</i>	Sucrose synthase	Barley	D24462
6	34.8	R2147	<i>sal2</i>	SalT protein	Rice	D24547
6	57	C235	<i>hmg2</i>	High mobility group-like protein NHP2	Yeast	D15191
6	69.8	R111	<i>fdh</i>	Formate dehydrogenase	<i>Pseudomonas</i> sp.	D23770
6	69.8	C58	<i>srp</i>	Signal recognition particle 19K	Human	D15105
6	112	C556	<i>gdc1</i>	Glycine-cleavage system protein H	Garden pea	D15379
6	112.1	R2403	<i>pgk</i>	Phosphoglycerate kinase, cytosolic	Wheat	D26320
6	115.2	C259C	<i>ubq2</i>	Ubiquitin	Tomato, potato, oat	D22550
6	121.5	C69	<i>eif1</i>	Initiation factor eIF-4A	Curled-leaved tobacco	D15109
6	126.2	R1888	<i>ams2</i>	S-adenosylmethionine synthetase 2	<i>Arabidopsis thaliana</i>	D24436
6	127.3	R1394B	<i>nod</i>	Nodulation protein	<i>Rhizobium leguminosarum</i>	D24124
6	128.9	R1167	<i>cat</i>	Catalase chain I	Maize	D24082
6	128.9	C607	<i>hmg1</i>	High mobility group protein	Wheat	D28196
7	40.3	R2401	<i>thx</i>	Thioredoxin	<i>Arabidopsis thaliana</i>	D24700
7	46.5	R1488	<i>hxx</i>	Hexokinase P1	Yeast	D24182
7	49.2	C67A	<i>rif</i>	ADP-ribosylation factor 4	Human	D28199
7	54.2	R610	<i>mak</i>	MAK16 protein	Yeast	D23935
7	54.2	C479	<i>sps</i>	Spermidine synthetase	Human	D22594
7	55.4	C492	<i>gcw3</i>	Glycine-rich cell wall structural protein	Garden petunia	D22596
7	88	R2394	<i>cpk</i>	Protein kinase, calcium dependent	Soybean	D24697
7	98.5	C1412	<i>elf2</i>	Elongation factor1 beta chain	Rice	D15852
7	101.9	R3349	<i>cyt</i>	Cystathionine gamma-lyase	Potato	D25146
7	105.3	C507	<i>cpn</i>	Probable chaperonin	<i>Synechococcus</i> sp.	D26192
7	108.4	C1340	<i>par</i>	Par gene protein	Common tobacco	D22794
7	124.1	C213	<i>odh</i>	Oxoglutarate dehydrogenase	<i>Escherichia coli</i>	D15178
7	124.6	R411	<i>tab</i>	Tat-binding protein	Human	D23852
7	125.4	C586	<i>gcw1</i>	Glycine-rich cell wall structural protein	Garden petunia	D22623
8	1.1	R1963	<i>map</i>	Membrane alanyl aminopeptidase	<i>Escherichia coli</i>	D28310
8	1.8	R662	<i>hyp2</i>	Hypothetical protein 1 (sul 3' region)	<i>Bacillus subtilis</i>	D23961
8	2.6	R1880	<i>acl</i>	Acyl carrier protein 3	Barley	
8	23.5	R1985	<i>pkc2</i>	Protein kinase C homolog	Rice	D24464
8	27.9	R2382	<i>pat</i>	Patatin T5	Potato	D24690
8	42.5	C929	<i>reg2</i>	14-3-3 protein	Barley	D22692
8	53.9	R1394A	<i>nod</i>	Nodulation protein	<i>Rhizobium leguminosarum</i>	D24124

Table 1 continued.

Chr.	Position	Marker	Gene	Protein name	Organism	DBJ ID No.
8	100.5	R2285	<i>gdh</i>	Glucose dehydrogenase (pyrroloquinoline-quinone)	<i>Acinetobacter calcoaceticus</i>	D24633
8	109.1	C922B	<i>gbp</i>	GTP-binding regulatory protein beta chain	<i>Chlamydomonas reinhardtii</i>	D22667
8	111.7	C277	<i>rpa</i>	Acidic ribosomal protein 4	Fruit fly	D15212
9	0.8	C711	<i>pab</i>	Polyadenylate-binding protein	Human	D15488
9	46.7	C397	<i>sco1</i>	SCO1 protein	Yeast	D22575
9	74.6	R1562	<i>hsp4</i>	Heat shock protein 82	Rice	D24234
9	75.1	C846	<i>pkc1</i>	Protein kinase C homolog	Rice	D15569
9	78.7	R3312	<i>gco3</i>	Beta-glucosidase B	<i>Bacillus polymyxa</i>	D28326
9	88.4	C985	<i>hsp2</i>	Heat shock protein 82	Rice	D22707
9	97	C506	<i>hmg3</i>	High mobility group protein	Maize	D22603
9	97.3	C632	<i>urt1</i>	UTP:glucose-1-phosphate uridylyltransferase	Potato	D15437
10	2.3	C701	<i>adh2</i>	Alcohol dehydrogenase	Human	D15481
10	11.7	C913A	<i>eno1</i>	Enolase	Tomato	D28210
10	17.6	C489	<i>atp2</i>	H ⁺ -transporting ATP synthase gamma chain	<i>Rhodospirillum rubrum</i>	D15343
10	42.7	R2604	<i>gcw4</i>	Glycine-rich cell wall structural protein	Rice	D24186
10	42.7	R2252	<i>hyp4</i>	Hypothetical protein YCL59C	Yeast	D24612
10	43.5	C677	<i>gcw2</i>	Glycine-rich cell wall structural protein	Rice	D13464
11	9.2	C950	<i>tum</i>	Tumor protein	<i>Arabidopsis thaliana</i>	D22697
11	65	R120	<i>ahc</i>	Adenosyl homocysteinase	Rat	D23773
11	65.8	C3	<i>sec2</i>	Sec23 protein	Yeast	D22492
11	91	R1572	<i>adh2</i>	Alcohol dehydrogenase	Rice	D24243
11	91.3	C496	<i>adh1</i>	Alcohol dehydrogenase	Maize	D15347
11	91.3	R682	<i>adh2</i>	Alcohol dehydrogenase	Maize	D23967
11	114	R3202	<i>cbp</i>	Calcium binding protein	Mouse	D25111
12	1.4	R2292	<i>rab5</i>	GTP-binding protein rab5	Dog	D28317
12	14.5	C1069	<i>hyp1</i>	Hypothetical protein	Maize	D15675
12	72.6	R3375	<i>cla</i>	Clathrin-associated protein 17	Rat	D25151
12	83	R2672	<i>elf4</i>	Elongation factor selB	<i>Escherichia coli</i>	D24864
12	87.1	C1336	<i>ald1</i>	Fructose-biphosphate aldolase	Rice	D28223
			15 mapped <i>pox</i>	Peroxidase	Horseradish and turnip	
			3 mapped <i>his2a</i>	Histone H2A	Mainly wheat and maize	
			4 mapped <i>his2b</i>	Histone H2B	Mainly wheat and maize	
			4 mapped <i>his3</i>	Histone H3	Mainly wheat and maize	
			5 mapped <i>his4</i>	Histone H4	Mainly wheat and maize	
			24 mapped <i>rpl</i>	Ribosomal protein large subunit	Mainly rat	
			15 mapped <i>rps</i>	Ribosomal protein small subunit	Mainly rat	

these isozymes by mapping cDNA clones derived from callus and root cDNA libraries. Thus, such genes as *got*, *adh*, and *pox*, which have been assigned in the conventional linkage map by segregation analysis of gene products, could be accurately mapped with their exact locations in the chromosome. In addition, a number of genes, which code for structural proteins such as actin, tubulin and ubiquitin, genes associated with the glycolytic pathway, genes related to the cell cycle, as well as heat shock proteins,

were also mapped. Some of these genes, however, did not necessarily correspond to a specific gene sequence but rather to one of the highly conserved multiple copies in the genome and were mapped in several loci in one or more chromosomes.

Several multigene families such as ribosomal proteins and histones, which have been identified from the large-scale cDNA analysis, have also been mapped. Twenty-four genes of the large subunit ribosomal protein and 15 genes of the small subunit ribosomal protein were found to be widely distributed in the rice genome. We have also identified and mapped the genes for histone proteins, namely, H1, H2A, H2B, H3, and H4 proteins. In human and other animals, these five types of genes formed clusters or repeated tandem units. In rice, however, they were found to be widely distributed in several chromosomes.

Thus, construction of a detailed genetic map using expressed gene sequences may provide a vast amount of information on the structural and functional organization of the rice genome. This could be very useful in identifying a gene of interest as well as in the subsequent stage of manipulation and isolation.

Genomic DNA markers as sequence-tagged sites

The chromosomal distribution of genomic clones classified as random genomic clones (G-number), *NotI* linking clones (L-number), YAC-end clones (Y-number), and TELs were also determined (Fig. 1). One hundred and thirty-seven randomly selected genomic clones were evenly distributed on the map. Most of these genomic clones have been sequenced and registered at DDBJ. Thus, these clones can be referred to as STSs on the map. The YAC-end clones and *NotI* linking clones were used for mapping to determine the nature of these sequences, which was necessary for physical map construction. However, mapping of 33 YAC-end clones (Y-number) and 90 *NotI* linking clones (L-number) did not show any specific features in terms of distribution and chromosomal localization of these clones. Among the mapped YAC-end clones were those containing both ends of the DNA fragment in YAC. These clones were mapped at close proximity to each other so that the physical distance corresponding to the genetic distance in cM can be calculated.

The map positions of TELs isolated using cassette ligation-mediated PCR were also determined (Ashikawa et al 1994). Two of these clones have been located on opposite ends of chromosome 11 so that this chromosome could be completely saturated with DNA markers. Subtelomeric clones have also been mapped on one end of chromosome 12 as well as chromosome 5.

RAPD markers were used to fill such regions on the map with very few markers. More than 150 RAPD were detected between Nipponbare and Kasalath using 1,400 combinations of arbitrarily designed 10-nucleotide primers (Monna et al 1994). One hundred and forty-seven RAPD markers represented by P-number and T-number on the map were mapped on the 12 chromosomes of rice. The T-number markers correspond to RAPD markers, which were converted to STS. More importantly, regions in some chromosomes that cannot be linked by DNA markers had been successfully connected by RAPD markers. The distal regions of chromosomes 1, 6, and 8 were extended by RAPD markers P61, P73, and P122, respectively. These suggest that

RAPD markers can be very useful to fill gaps or to extend the linkage map of each chromosome.

Syntenry with the wheat genome

To clarify the relationships of the rice genome with other crops, 60 wheat genomic DNA fragments (W-number) have been mapped on our high-density linkage map in collaboration with the Cambridge Laboratory, John Innes Centre, UK. The results showed that most of these markers have the same linkage order in wheat and rice (Kurata et al 1994a). Furthermore, it has been clarified that rice chromosome 1 corresponds to wheat group 3, rice chromosome 2 to wheat 6, rice 3 to wheat 4, rice 4 and 7 to wheat 2, rice 5 to wheat 1, rice 6 to wheat 7, and rice 9 to wheat 5. This suggests conservation of genome structure between rice and wheat, which are from different Gramineae tribes and differ in both chromosome number and genome size. We are also pursuing reciprocal mapping of DNA probes with other crops such as barley and maize. Eventually, we hope to clarify the extent of syntenry and linkage conservation among cereal crops.

Conserved linkage order in chromosomes 11 and 12

Although most of the clones used as probes showed a single-copy band on genomic Southern hybridization, some DNA probes had two or more bands and were located in duplicate or triplicate loci. Seventy-nine probes (6.1% of the total mapped DNA probes) were mapped on more than one locus. Duplicate segments were particularly observed between chromosomes 11 and 12 (Nagamura et al 1995, Fig. 2). Thirteen of the 33 mapped DNA markers at the distal regions of these chromosomes, including a TEL (TEL2), were mapped as duplicate loci. These duplicated segments occupy 10 and 11.8 cM in chromosomes 11 and 12, respectively. The other 20 markers in these regions also showed two or more main bands, but only one band was polymorphic, which was mapped in either chromosome 11 or 12. This suggests that RFLP mapping can also be an effective method to clarify chromosomal rearrangements as well as conservation of gene order accompanied by the evolution of a species.

Toward a saturated linkage map and more

At present, we are mapping additional markers in our RFLP linkage map to create a tighter linkage. In addition to callus and root, we are also using cDNA clones from green shoot, etiolated shoot, and developing seed cDNA libraries. As of Mar 1995, we have mapped an additional 521 DNA markers so that our map now has 1,904 DNA markers and a length of 1,556 cM. The average interval between markers is about 0.8 cM. However, there are still several regions in some chromosomes with very few markers as well as long stretches without any markers. Thus, it is necessary to screen for more markers to fill these gaps or to analyze the exact nature of such regions in the chromosomes.

Ultimately, we would like to establish a map with about 2,000 DNA markers at very close intervals necessary for physical map construction and gene tagging. Selection and ordering of YAC clones covering the entire genome to construct a detailed physical

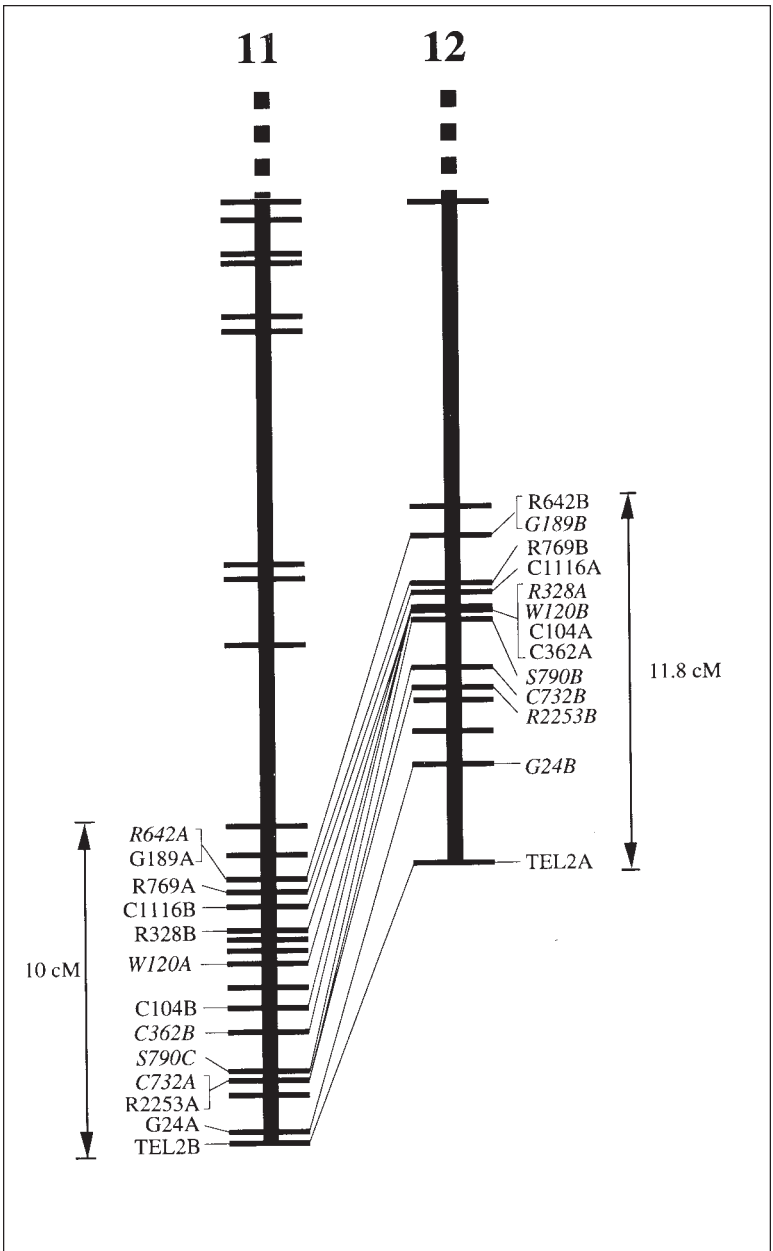


Fig. 2. The distal region of chromosomes 11 and 12 with highly conserved linkage of 13 DNA markers. Marker designations are described in Figure 1. Markers in italics were mapped after the publication of the linkage map in Kurata et al (1994b).

map of rice is in progress. Tagging of genes controlling phenotypical traits, which are important agronomically and for scientific studies, is also under way. We have already identified the chromosomal locations of such genes as *Xa1* (bacterial blight resistance gene) and *Se1* (photoperiod sensitivity gene). Isolation of these genes is expected to progress efficiently through positional map-based cloning with tagged DNA markers by using physically arrayed YAC or cosmid clones.

Thus, a high-density linkage map of rice will have far-reaching applications in understanding genome organization, function, and evolution. More importantly, it is expected to have enormous impact on the more practical aspect of plant genetic manipulation, that is, for marker-aided selection in breeding programs as well as for map-based cloning of agronomically important genes.

Cited references

- Ahn S, Tanksley SD. 1993. Comparative linkage maps of the rice and maize genomes. *Proc. Natl. Acad. Sci. USA* 90:7980-7984.
- Ashikawa I, Kurata N, Nagamura Y, Minobe Y. 1994. Cloning and mapping of telomere-associated sequences from rice. *DNA Res.* 1:67-76.
- Bernatzky R, Tanksley SD. 1986. Towards a saturated linkage map in tomato based on isozymes and random cDNA sequences. *Genetics* 112:887-898.
- Chang C, Bowman JL, DeJohn AW, Lander ES, Meyerowitz EM. 1988. Restriction fragment length polymorphism linkage map for *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* 85:6856-6860.
- Da Silva AG, Sorrells ME, Burnquist WL, Tanksley SD. 1993. RFLP linkage map and genome analysis of *Saccharum spontaneum*. *Genome* 36:782-791.
- Kleinhofs A, Kilian A, Saghai Maroof MA, Biyashev RM, Hayes P, Chen FQ, Lapitan N, Fenwick A, Blake TK, Kanazin V, Ananiev E, Dahleen L, Kudrna D, Bollinger J, Knapp SJ, Liu B, Sorrells M, Heun M, Franckowiak JD, Hoffman D, Skadsen R, Steffenson BJ. 1993. A molecular, isozyme and morphological map of the barley (*Hordeum vulgare*) genome. *Theor. Appl. Genet.* 86:705-712.
- Kosambi DD. 1944. The estimation of map distances from recombination values. *Ann. Eugenet.* 12:172-175.
- Kurata N, Moore G, Nagamura Y, Foote T, Yano M, Minobe Y, Gale M. 1994a. Conservation of genome structure between rice and wheat. *Bio/Technology* 12:276-278.
- Kurata N, Nagamura Y, Yamamoto K, Harushima Y, Sue N, Wu J, Antonio BA, Shomura A, Shimizu T, Lin SY, Fukuda A, Shimano T, Kuboki Y, Toyama T, Miyamoto Y, Kirihara T, Hayasaka K, Miyao A, Monna L, Zhong HS, Tamura Y, Wang ZX, Momma T, Umehara Y, Yano M, Sasaki T, Minobe Y. 1994b. A 300-kilobase interval genetic map of rice including 883 expressed sequences. *Nat. Genet.* 8:365-372.
- 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.
- McCouch SR, Kochert G, Yu ZH, Wang ZY, Khush GS, Coffman WR, Tanksley SD 1988. Molecular mapping of rice chromosomes. *Theor. Appl. Genet.* 76:815-829.
- Monna L, Miyao A, Inoue T, Fukouka S, Yamazaki M, Zhong SH, Sasaki T, Minobe Y. 1994. Determination of RAPD markers in rice and their conversion into sequence-tagged sites (STSs) and STS-specific primers. *DNA Res.* 1:139-148.

- Murray MG, Thompson WF. 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 8:4321-4325.
- Nagamura Y, Inoue T, Antonio BA, Shimano T, Kajiya H, Shomura A, Lin SY, Kuboki Y, Harushima Y, Kurata N, Minobe Y, Yano M, Sasaki T. 1995. Conservation of duplicated segments between rice chromosomes 11 and 12. *Breeding Sci.* 45: 373-376.
- Rognli OA, Devos KM, Chinos CN, Harcourt RL, Atkinson MD, Gale MD. 1992. RFLP mapping of rye chromosome 7R reveals a highly translocated chromosome relative to wheat. *Genome* 35:1026-1031.
- Saito A, Yano M, Kishimoto N, Nakagahra M, Yoshimura A, Saito K, Kuhara S, Ukai Y, Kawase M, Nagamine T, Yoshimura S, Ideta O, Ohsawa R, Hayano Y, Iwata N, Sugiura M. 1991. Linkage map of restriction fragment length polymorphism loci in rice. *Jpn. J. Breed.* 41:665-670.
- Tanksley SD, Ganai MW, Prince JP, de Vicente MC, Bonierbale MW, Broun P, Fulton TM, Giovannoni JJ, Grandillo S, Martin GB, Messeguer R, Miller JC, Miller L, Paterson AH, Pineda O, Roder MS, Wing RA, Wu W, Young ND. 1992. High density molecular linkage maps of the tomato and potato genomes. *Genetics* 132:1141-1160.
- Wu KS, Glaszmann JC, Khush GS. 1988. Chromosomal locations of ten isozyme loci in rice (*Oryza sativa* L.) through trisomic analysis. *Biochem. Genet.* 26:303-320.

Notes

Authors' addresses: B.A. Antonio, A. Shomura, T. Shimano, Y. Kuboki, S.Y. Lin, T. Inoue, Y. Harushima, K. Yamamoto, Institute of Society for Techno-innovation of Agriculture, Forestry and Fisheries, 446-1 Ippaizuka, Kamiyokoba, Tsukuba, 305 Japan; H. Kajiya, Y. Nagamura, N. Kurata, M. Yano, Y. Minobe, and T. Sasaki, National Institute of Agrobiological Resources, 2-1-2 Kannondai, Tsukuba, 305 Japan.

Citation: [IRRI] International Rice Research Institute. 1996. Rice genetics III. Proceedings of the Third International Rice Genetics Symposium, 16-20 Oct 1995. Manila (Philippines): IRRI.