Theor Appl Genet (1994) 88:722–726

N. Kishimoto · H. Higo · K. Abe · S. Arai · A. Saito
K. Higo

Identification of the duplicated segments in rice chromosomes 1 and 5 by linkage analysis of cDNA markers of known functions

Received: 21 October 1993 / Accepted: 21 December 1993

Abstract We mapped two loci for ADP-ribosylation factor homologues (ARF1, ARF2) and two loci for cysteine proteinase inhibitors (oryzacystatin-I and -II: OCI, OCII) by linkage analysis of restriction fragment length polymorphism loci in rice (*Oryza sativa* L.) genomic DNAs using their cDNAs as probes. *Oc-1* and *Arf-2* were found to be closely located to each other on chromosome 1, while *Oc-2* and *Arf-1, both found* on chromosome 5, were also located close to each other. The map distances are about 2 cM in both pairs. In each chromosome, the *Arf* locus was located about 27 cM from that of the aldolase gene (*Ald-2* in chromosome 1 and *Ald-1* in chromosome 5). These three genes are in the same order, *Ald-Arf-Oc*, but in opposite orientations relative to the distal ends of the linkage group. The presence of two sets of three linked genes on chromosomes 1 and 5 strongly suggests a structural similarity of the blocks of the two chromosomes, which probably reflects duplication of the segment. A recent investigation by other workers has shown that these rice blocks correspond to two regions in maize chromosomes 8 and 6, that have previously been shown to share many duplicated nucleotide sequences. It is therefore very likely that the duplication of the region occurred before the divergence of rice and maize during the evolution of the subfamilies of the grasses (Gramineae). In view of a recently discovered possible structural similarity between the small GTP-binding protein superfamily, which includes *Arf* and *ras* proteins, and the cystatin family, the close linkage of *Oc* and *Arf* loci found in the present study suggests a possible cluster of genes related to the small GTP-binding proteins.

Key words RFLP-mapping · *Oryza sativa* · Duplication
Oryzacystatin genes · ADP-ribosylation factor genes

Communicated by K. Tsunewaki

N. Kishimoto · H. Higo · A. Saito[1] · K. Higo (✉)
Department of Molecular Biology,
National Institute of Agrobiological Resources,
1-2 Kannondai 2-chome, Tsukuba, Ibaraki 305, Japan

K. Abe · S. Arai
Faculty of Agriculture, University of Tokyo, Yayoi, Bunkyo-ku,
Tokyo 113, Japan

Present address:
[1] Department of Crop Improvement, Kyushu National Agricultural Experimental Station, Nishigoushi, Kikuchi, Kumamoto 861-11, Japan

Introduction

Rice (*Oryza sativa* L.) is one of the most important crop plants in the world, and over the past 70 years a genetic map has been constructed that includes about 170 loci (for a review, see Kinoshita 1990). Rice has 12 chromosomes per genome, but their linkage relationships to other related cereals such as wheat and maize are not clear. The evolutionary and genetical relationships among chromosomes of cereal plants have important implications for both fundamental biology and agricultural research (Moore et al. 1993; Bennetzen and Freeling 1993).

Recent advances in the mapping of DNA markers by restriction fragment length polymorphism (RFLP) in higher plants has provided us with the means to examine and compare the fine genetical structure of chromosomes. For example, the interspecific comparison of the high density molecular linkage maps of the tomato and potato genomes, obtained by using mainly single-copy DNA markers, showed a conservation of marker order but differences in chromosome and total map length (Gebhardt et al. 1991; Tanksley et al. 1992b). DNA markers have been also used to examine the genetic structure of sorghum chromosomes (Hulbert et al. 1990) and also to compare fine genetic maps of homoeologous chromosomes between wheat, rye and barley (Devos et al. 1993 and references cited therein). On the other hand, by using multiple but low-copy number DNA markers, Helentjaris et al. (1988) identified the genomic locations of duplicate nucleotide sequences in maize. Genetical structural similarity among chromosomes in one plant species would suggest an evolutionary history of the chromosomes, i.e., a homoeologous relationship among the chromosomes.

We have embarked on the construction of a restriction fragment length polymorphism map of rice (*Oryza sativa* L.) (Saito et al. 1991) and recently also started mapping cloned cDNAs onto the RFLP linkage map (Kishimoto et al. 1994). In this paper, we report the results of cDNA mapping which led us to identify probable duplicated segments in rice chromosomes 1 and 5 and also to discuss a possible clustering of genes related to small GTP-binding proteins.

Materials and methods

Source of the probe

Probe for rice ADP-ribosylation factor (ARF) gene homologues

A cDNA of the rice ARF homologue was obtained during an attempt to clone the cDNA of the phospholipase A_2 homologue. Briefly, a cDNA library using λgt22A (GIBCO BRL, Md., USA) as vector was constructed using poly(A)$^+$RNA extracted from rice (*Oryza sativa* L., japonica type) callus culture derived from embryos. The DNA extracted from the library was used as a template in the polymerase chain reaction (PCR) to screen for the phospholipase A2 homologue using an oligonucleotide corresponding to the consensus amino acid sequence found in the mammalian non-pancreatic phospholipase A_2 family and another oligonucleotide that hybridizes to the poly(A) tail. Amplified DNA fragments were purified by agarose gel electrophoresis, and DNA recovered from the gel was inserted into pCR1000 vector (Invitrogen, Calif., USA). The partial nucleotide sequences of several clones were determined and compared with the sequences in the EMBL database. A region (113 bp) of a sequence (about 460 bp) of the clone pHHP1-1 insert was found to be similar (about 80% identity) to a part of that of ARF cDNA isolated from *Arabidopsis thaliana* (Regad et al. 1993). The cDNA sequence of the putative rice ARF homologue was also found to be similar to those of other ARF cDNAs isolated from animals or yeasts. The probe used in this study was prepared by amplifying the insert by PCR using the DNA of clone pHHP1-1 as template and two short oligonucleotides corresponding to the flanking regions of the pCR1000 vector as the primers. The DNA fragments to be used as probes were radiolabeled using a random oligonucleotide labeling kit (Amersham, UK). The complete sequence of the rice ARF cDNA will be published elsewhere.

Probe for rice cysteine proteinase inhibitor (oryzacystatin) genes

The clones λOC26 and λnOC1, which contain cDNAs of oryzacystatin-I and -II respectively, were isolated from cDNA libraries constructed using poly(A)$^+$RNA extracted from immature rice seeds harvested 2 weeks after flowering (Abe et al. 1987; Kondo et al. 1990). A comparison of these two nucleotide sequences showed 59% identity (Kondo et al. 1990). A detailed description of the characterization of cDNAs encompassing the entire coding sequence of oryzacystatin has been given elsewhere (Abe et al. 1987; Kondo et al. 1990).

Plant analysis

Plants were analyzed as described before (Kishimoto et al. 1994). Briefly, total DNA was extracted from the leaves of cvs 'Kasalath' (indica type) and 'FL134' (japonica type) of *Oryza sativa*, digested with either *Bam*HI, *Bgl*II, *Eco*RV, or *Hin*dIII, subjected to electrophoresis, blotted to a nylon membrane, and hybridized to the probes described above, all according to the method of Sambrook et al. (1989). All of the filters were washed in 2×SSC at 42 °C, and then subjected to autoradiography at −80 °C.

The MAPL software program (Ukai et al. 1990, 1991) was used for the genetic analysis. The recombination values between loci detected with the above-mentioned probes and the marker loci previously mapped in a rice RFLP map (Saito et al. 1991) were estimated by a maximum likelihood method (Allard 1956) in 144 individual F_2 plants derived from the cross 'Kasalath'×'FL134'. Within a linkage group, the putative order of the loci was determined by applying the metric multidimensional scaling method. Other details have been described by Saito et al. (1991).

Results

Polymorphism

The probe for the ARF homologue genes consisted of a 0.46-kbp (approximately) cDNA insert of the clone pHHP1-1 encoding the carboxyl-terminal 31 amino acid residues of rice ARF homologue flanked by the 3′-noncoding sequence, which is expressed in calli derived from embryos. The autoradiogram of Southern blots of DNAs from the 'Kasalath' and 'FL134' revealed that *Hin*dIII digestion produced a polymorphic major band (2.8 kbp in 'Kasalath' and 5.9 kbp in 'FL134') and a minor band (8.4 kbp in 'Kasalath' and 10.5 kbp in 'FL134') (Fig. 1). The genomic DNAs of 144 individual F_2 plants of the cross 'Kasalath'×'FL134' were therefore digested with *Hin*dIII, subjected to electrophoresis, blotted to a nylon membrane, and hybridized to the probes as described above. Autoradiogram data of the F_2 population were used for the linkage analysis. F_2 genotypes were classified according to the size of the bands observed: type 1 ('Kasalath' type band), type 2 ('FL134' type band), and type 3 (both bands) (Fig. 1).

The probe for the oryzacystatin-I gene was a 0.7-kbp cDNA insert of the clone λOC26 encoding the entire coding sequence flanked by short noncoding sequences at the 5′ and 3′ ends. The autoradiogram of Southern blots of DNAs from the 'Kasalath' and 'FL134' cultivars revealed that the *Eco*RV digestion produced polymorphic bands (11.5 kbp in 'Kasalath' and 6.3 kbp in 'FL134') (Fig. 2).

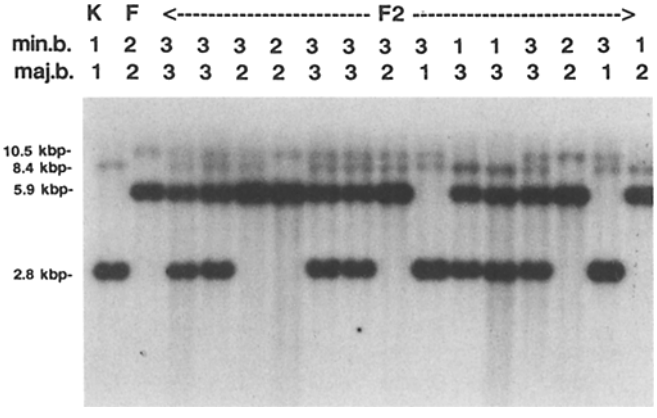

Fig. 1 Autoradiogram of *Hin*dIII-digested genomic DNA of two parental lines (*K* 'Kasalath', *F* 'FL134') and several representative F_2 plants hybridized with radiolabeled probes for rice ADP-ribosylation factor homologue genes. F_2 genotypes were classified according to the size of the bands observed: type 1 ('Kasalath' type band), type 2 ('FL134'-type band), and type 3 (both bands)

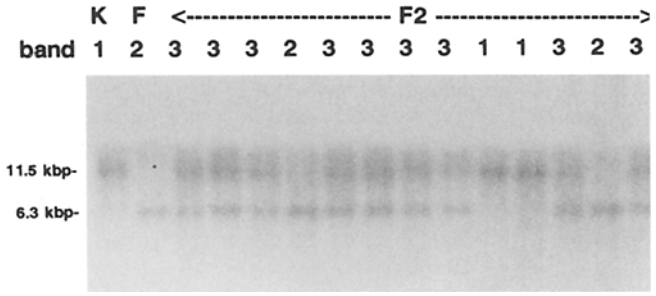

Fig. 2 Autoradiogram of *Eco*RV-digested genomic DNA of two parental lines (*K* 'Kasalath', *F* 'FL134') and several representative F_2 plants hybridized with radiolabeled probes for the oryzacystatin-I gene

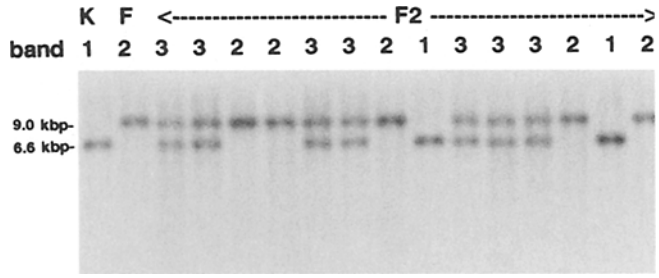

Fig. 3 Autoradiogram of *Hin*dIII-digested genomic DNA of two parental lines (*K* 'Kasalath', *F* 'FL134') and several representative F_2 plants hybridized with radiolabeled probes for the oryzacystatin-II gene

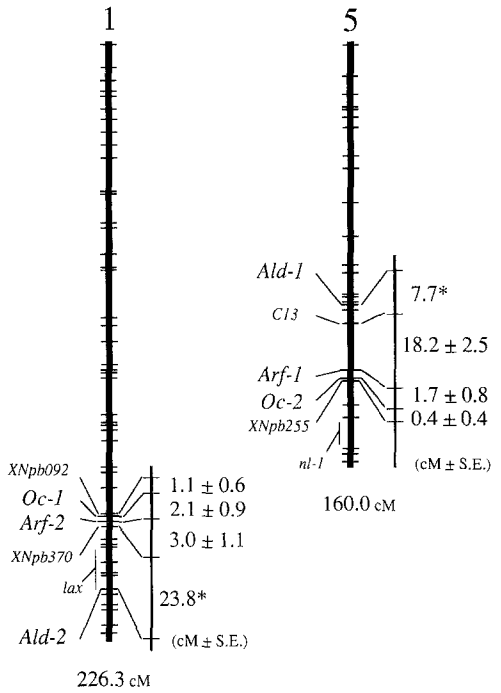

Fig. 4 Linkage map of rice chromosomes 1 and 5, and the location of the ARF genes (*Arf-1*, *Arf-2*) and oryzacystatin genes (*Oc-1* and *Oc-2*). Linkage relations and linear orders between these genes and RFLP markers are shown at the *right* and the *left* of the *thick bar*, respectively. The map distances for two aldolase loci (*) were taken from Saito et al. (1991). The total map distance of chromosomes (Saito et al. 1991) are shown at the *bottom* of each *bar*. Anonymous *horizontal bars* indicate other RFLP loci in the map. *lax* and *nl-1* are genetic markers that had been previously located in the map (Saito et al. 1991)

The *Eco*RV-digested genomic DNAs of 144 individual F_2 plants were therefore subjected to electrophoresis, blotted to a nylon membrane, and hybridized to the probes as described above. Autoradiogram data of the F_2 population were used for the linkage analysis.

The probe for the oryzacystatin-II gene was a 0.65-kbp cDNA insert of the clone λnOC1 encoding the entire coding sequence flanked by short noncoding sequences at the 5′ and 3′ ends. The autoradiogram of Southern blots of DNAs from the 'Kasalath' and 'FL134' cultivars revealed that the *Hin*dIII digestion produced polymorphic bands (6.6 kbp in 'Kasalath' and 9.0 kbp in 'FL134') (Fig. 3). The *Hin*dIII-digested genomic DNAs of 144 individual F_2 plants were therefore subjected to electrophoresis, blotted to a nylon membrane, and hybridized to the probes as described above. Autoradiogram data of the F_2 population were used for the linkage analysis.

Location and order of the probe and gene loci

The map locations of the ARF homologue gene (locus name: *Arf-1*) based on the segregation analysis of the major band (5.9 kbp/2.8 kbp in Fig. 1) and the oryzacystatin-II gene (locus name: *Oc-2*) were determined to lie between two previously known RFLP markers on chromosome 5 (Saito et al. 1991), *C13* (a cDNA marker locus) and *XNpb255*. The linear order was *C13-Arf-1-Oc-2-XNpb255*, with map distances of 18.2±2.5 cM, 1.7±0.8 cM and 0.4±0.4 cM, respectively (Fig. 4). The locus for an aldolase gene (*Ald-1*) has been located 7.7 cM from *C13* (Saito et al. 1991).

The map locations of the second putative ARF homologue gene (locus name: *Arf-2*) based on the segregation analysis of the minor bands (10.5 kbp/8.4 kbp in Fig. 1) and the oryzacystatin-I gene (locus name: *Oc-1*) were determined to lie between two previously known genomic RFLP markers on chromosome 1 (Saito et al. 1991), *XNpb092* and *XNpb370*. The linear order was *XNpb092-Oc-1-Arf-2-XNpb370*, with map distances of 1.1±0.6 cM, 2.1±0.9 cM and 3.0±1.1 cM, respectively (Fig. 4). The locus for another aldolase gene (*Ald-2*) has been located 23.8 cM from *XNpb370* (Saito et al. 1991)

The linear orders of the three cDNA markers described above were therefore *Ald-2-Arf-2-Oc-1* in chromosome 1 with approximate distances of 27 cM and 2 cM, respectively, and *Ald-1-Arf-1-Oc-2* in chromosome 5 with 26 cM and 2 cM, respectively.

Discussion

Nature of marker genes of known function

The presence of ARF in plants was confirmed quite recently by a report on the cloning of an *Arabidopsis* homologue of mammalian and yeast ARF cDNA (Regad et al 1993). Mammalian and yeast ARFs are believed to be important in vesicular transport and essential for cell viability (Balch et al. 1992), and their biological roles have been the subject of numerous recent intensive studies. Two ARF genes have been identified in *Arabidopsis*, but their chromosomal locations have not been determined (Regad et al. 1993). In yeast (*Saccharomyces cerevisiae*), only two ARF genes are known, and these have been mapped within 28 cM of each other on chromosome IV (Stearns et al. 1990). In higher plant, ARFs are also expected to play very important roles because plant physiology involves various vesicular transport steps such as the intracellular transport of storage proteins from the Golgi body to vacuoles.

Oryzacystatin was the first proteinaceous cysteine-proteinase inhibitor of plant origin to be purified and characterized in detail, and its cDNA was cloned by Abe et al. (1987). Similar cDNAs have also recently been isolated from maize and other plants. To the best of our knowledge, the chromosomal locations of ARF and cysteine-proteinase inhibitor genes of a plant origin have never been reported before.

Aldolase (fructose-1,6-bisphosphate aldolase) is one of the key enzymes of the glycolytic pathway, catalyzing the breakdown of the fructose ring to two trioses. Higher plants have two forms of aldolase, cytoplasmic and plastidic types, and their cDNAs have been isolated from maize (Hake et al. 1985) and other plants.

Possible duplication of the chromosome segment

Genetical structural similarity among chromosomes within one plant genome is likely a reflection of a homoeologous relationship of the chromosomes. In maize, duplicated isozymic loci have been found in parallel linkages, and chromosome segment duplication has been suggested (Wendel et al. 1986 and references cited therein). Comparison of the fine genetical structure of chromosomes has been made possible by the mapping of molecular markers. Through the use of cloned maize DNAs as probes, Helentjaris et al. (1988) identified the genomic locations of many duplicate nucleotide sequences on RFLP maps. These were usually duplicated on separate chromosomes, with the loci sometimes being in the same linear order. The chromosome pair 2 and 7 and the group involving chromosomes 3, 6, and 8 contain the largest number of pairs of duplicate loci detected in this study. The presence of chromosomal segments having duplicated loci in a generally ordered arrangement strongly suggests that two segments share a common origin.

The chromosomal locations of two sets of rice genes for aldolase, the ARF protein, and oryzacystatin with almost identical map distances, although in opposite orientation, on two separate chromosomes (chromosomes 1 and 5) found in this study also suggest the genetical structural similarity of (at least) the block(s) of the two chromosomes. This may be attributable to an ancestral relationship of these two chromosomes, i.e., these two chromosomes are homoeologous. Alternatively, duplication of the segment of one chromosome was followed by its translocation to another.

The correlation of many molecular markers in our RFLP map (Tsukuba map: Saito et al. 1991; Kishimoto et al. 1992, 1993) and those in the Cornell map (McCouch et al. 1988; Tanksley et al. 1992a) are now in progress (Xiao et al. 1992). According to a recent study on comparative linkage maps of the rice and maize genomes (Ahn and Tanksley 1993), the putative duplicated regions on rice chromosomes 1 and 5 correspond to two regions in maize chromosomes 8 and 6, respectively. The genetic structural similarity of the regions of maize chromosomes 6 and 8 has been suggested before as described above (Helentjaris et al. 1988). It would therefore be reasonable to assume that duplication of this region occurred before the divergence of maize and rice. An attempt to map these and other markers on wheat and other higher plants is currently under way.

Possible cluster of genes related to small GTP-binding proteins

Arf genes constitute a "small GTP-binding protein gene" family (or sometimes called "ras super family") together with the *ras*, *rho*, and *rab/ypt* protein gene sub-families in animals and yeasts (Sewell and Kahn 1988). The ARF and *rab/ypt* proteins are both involved in vesicular transport, although their precise distinct functions are still not clear. Almost all of the plant genes for small GTP-binding proteins so far isolated belong to the *rab/ypt* type, as judged from their nucleotide sequence similarity (Terryn et al. 1993). On the other hand, cysteine proteinase inhibitor activity of the mammalian *ras* protein (Hiwasa et al. 1987) and, subsequently, a possible structural similarity of the *ras* protein to previously known cystatins (cysteine proteinase inhibitor) have been recently reported (Hiwasa et al. 1989). These observations suggest a possible evolutionary relationship between the *ras* protein and cystatin, and therefore between the "small GTP-binding protein family" and cystatin. If this is the case, the close linkage of genes for the ARF protein and oryzacystatin found in the present study might be a reflection of the duplication of a common "ancestral GTP-binding protein gene", and other genes for proteins of a similar nature might be found in this region.

References

Abe K, Emori Y, Kondo H, Suzuki K, Arai S (1987) Molecular cloning of a cysteine proteinase inhibitor of rice (oryzacystatin). J Biol Chem 262:16793–16797

Ahn S, Tanksley SD (1993) Comparative linkage maps of the rice and maize genomes. Proc Natl Acad Sci USA 90:7980–7984

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

Balch WE, Kahn RA, Schwaninger R (1992) ADP-ribosylation factor is required for vesicular trafficking between the endoplasmic reticulum and the *cis*-Golgi compartment. J Biol Chem 267: 13053–13061

Bennetzen JL, Freeling M (1993) Grasses as a single genetic system: genome composition, collinearity and compatibility. Trends Genet 9:259–261

Devos KM, Millan T, Gale MD (1993) Comparative RFLP maps of the homoeologous group-2 chromosomes of wheat, rye and barley. Theor Appl Genet 85:784–792

Gebhardt C, Ritter E, Barone A, Debener T, Walkemeier B, Schachtschabel U, Kaufmann H, Thompson RD, Bonierbale MW, Ganal MW, Tanksley SD, Salamini E (1991) RFLP maps of potato and their alignment with the homoeologous tomato genome. Theor Appl Genet 83:49–57

Hake S, Kelly PM, Tayler WC, Freeling M (1985) Coordinate induction of alcohol dehydrogenase 1, aldolase, and other anaerobic RNAs in maize. J Biol Chem 260:5050–5054

Helentjaris T, Weber D, Wright S (1988) Identification of the genomic locations of duplicate nucleotide sequences in maize by analysis of restriction fragment length polymorphisms. Genetics 118:353–363

Hiwasa T, Yokoyama S, Ha J-M, Noguchi S, Sakiyama S (1987) c-Ha-*ras* gene products are potent inhibitors of cathepsins B and L. FEBS Lett 211:23–26

Hiwasa T, Sawada T, Sakiyama S, Kominami E, Katunuma N (1989) Cysteine proteinase-inhibitory activity of *ras* gene product is not affected by mutations at GTPase-activating protein-binding sites. Biol Chem Hoppe-Seyler 370:1215–1220

Hulbert SH, Richter TE, Axtell JD, Bennetzen JL (1990) Genetic mapping and characterization of sorghum and related crops by means of maize DNA probes. Proc Natl Acad Sci USA 87: 4251–4255

Kinoshita T (1990) Report of the committee on gene symbolization, nomenclature and linkage groups. Rice Genet Newsl 7:16–50

Kishimoto N, Shimosaka E, Matsuura S, Saito A (1992) A current RFLP linkage map of rice: alignment of the molecular map with the classical map. Rice Genet Newsl 9:118–124

Kishimoto N, Foolad MR, Shimosaka E, Matsuura S, Saito A (1993) Alignment of molecular and classical linkage maps of rice, *Oryza sativa*. Plant Cell Rep 12:457–461

Kishimoto N, Higo K, Higo H, Saito A (1994) The coding sequence for rice seed catalase detects a locus different from that determined by isozyme analysis. Theor Appl Genet 87:625–626

Kondo H, Abe K, Nishimura I, Watanabe H, Emori Y, Arai S (1990) Two distinct cystatin species in rice seeds with different specificities against cysteine proteinases. J Biol Chem 265: 15832–15837

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

Moore G, Gale MD, Kurata N, Flavell RB (1993) Molecular analysis of small grain cereal genomes: current status and prospects. Bio/Technology 11:584–589

Regad F, Bardet C, Tremousaygue D, Moisan A, Lescure B, Axelos M (1993) cDNA cloning and expression of an *Arabidopsis* GTP-binding protein of the ARF family. FEBS Lett 316:133–136

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

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Sewell JL, Kahn RA (1988) Sequences of the bovine and yeast ADP-ribosylation factor and comparison to other GTP-binding proteins. Proc Natl Acad Sci USA 85:4620–4624

Stearns T, Kahn RA, Botstein D, Hoyt MA (1990) ADP ribosylation factor is an essential protein in *Saccharomyces cerevisiae* and is encoded by two genes. Mol Cell Biol 10:6690–6699

Tanksley S, Causse M, Fulton T, Ahn N, Wang Z, Wu K, Xiao J, Yu Z, Second G, McCouch S (1992a) A high density molecular map of the rice genome. Rice Genet Newsl 9:111–115

Tanksley SD, Ganal 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 (1992b) High density molecular linkage maps of the tomato and potato genomes. Genetics 132:1141–1160

Terryn N, Van Montagu M, Inze D (1993) GTP-binding proteins in plants. Plant Mol Biol 22:143–152

Ukai Y, Ohsawa R, Saito A (1990) Automatic determination of the order of RFLP loci in a linkage group by metric multidimensional scaling method. Jpn J Breed 40 [Suppl 2]:302–303

Ukai Y, Ohsawa R, Saito A (1991) MAPL: A package of microcomputer program for RFLP linkage mapping. Rice Genet Newsl 8:155–158

Wendel JF, Stuber CW, Edwards MD, Goodman MM (1986) Duplicated chromosome segments in maize (*Zea mays* L.): further evidence from hexokinase isozymes. Theor Appl Genet 72:178–185

Xiao J, Fulton T, McCouch S, Tanksley S, Kishimoto N, Ohsawa R, Ukai Y, Saito A (1992) Progress in integration of the molecular maps of rice. Rice Genet Newsl 9:124–128