Theor Appl Genet (1998) 97: 356–369

A. E. Van Deynze · M. E. Sorrells · W. D. Park
N. M. Ayres · H. Fu · S. W. Cartinhour
E. Paul · S. R. McCouch

Anchor probes for comparative mapping of grass genera

Received: 27 June 1997 / Accepted: 17 July 1997

Abstract Comparative mapping of cDNA clones provides an important foundation for examining structural conservation among the chromosomes of diverse genera and for establishing hypotheses about the relationship between gene structure and function in a wide range of organisms. In this study, "anchor probes" were selected from cDNA libraries developed from rice, oat, and barley that were informative for comparative mapping in the grass family. One thousand eight hundred probes were screened on garden blots containing DNA of rice, maize, sorghum, sugarcane, wheat, barley, and oat, and 152 of them were selected as "anchors" because (1) they hybridized to the majority of target grass species based on Southern analysis, (2) they appeared to be low or single copy in rice, and (3) they helped provide reasonably good genome coverage in all species. Probes were screened for polymorphism on mapping parents, and polymorphic markers were mapped onto existing species-specific linkage maps of rice, oat, maize, and wheat. In wheat, both polymorphic and monomorphic markers could be assigned to chromosomes or chromosome arms based on hybridization to nullitetrasomic and ditelosomic stocks. Linkage among anchored loci allowed the identification of homoeologous regions of these distantly related genomes. Anchor probes were sequenced from both ends, providing an average of 260 bp in each direction, and sequences were deposited in GenBank. BLAST was used to compare the sequences with each other and with a non-redundant protein sequence database maintained at the European Molecular Biology Laboratory (EMBL). Of the anchor probes identified in this study 78% showed significant similarity to protein sequences for known genes with BLASTX scores exceeding 100.

Key words Grass family · Anchor probes · cDNA libraries · Comparative mapping

Communicated by F. Salamini

A. E. Van Deynze[1] · M. E. Sorrels · E. Paul
S. R. McCouch (✉)
Department of Plant Breeding and Biometry, Cornell University, Ithaca, NY 14853, USA
Fax: +1 607-255-6683
E-mail: srm4@cornell.edu

W. D. Park · N. M. Ayres · S. W. Cartinhour
Department of Biochemistry and Biophysics, Texas A&M University, College Park, TX 77843-2128, and Crop Biotechnology Center, Texas A&M University, College Station, TX 77843-2123, USA

H. Fu[2]
Department of Biochemistry and Biophysics, Texas A&M University, College Park, TX 77843-2128, USA

Present addresses:
[1] Calgene, 1920 Fifth St, Davis, CA 95616, USA
[2] Department of Horticulture, University of Wisconsin, 1575 Linden Dr. Madison, WI 53706-1590, USA

Introduction

The grass family consists of five principal subfamilies, and about 10,000 diverse species (Chapman 1996), including some of our most important agricultural food crops, such as rice, wheat, barley, oat, maize, sorghum, millet, and sugarcane. Grasses are believed to have originated in the late Cretaceous period more than 66 million years ago (Chapman 1996), and they now inhabit almost every land habitat known in both temperate and tropical regions. Despite the wide diversity of evolutionary modifications that have allowed this family to radiate throughout the world, there is a surprising consistency in the general plant body observable among members of the Gramineae. Thus, it is not surprising that the underlying genetic similarities, well documented by Vavilov (1940), have begun to find explanation in the increasing volume of molecular information currently being generated.

Comparative mapping of DNA probes provides an important foundation for combining genetic information from related species. The prerequisites for comparative mapping are a genetic linkage map for each species and a common set of DNA probes that can be used to align the maps. Restriction Fragment length polymorphism (RFLP) maps presently exist for rice (*Oryza sativa* L.) (Causse et al. 1994; Kurata et al. 1994a; Harushima et al. 1998), hexaploid wheat (*Triticum aestivum* L. em. Thell) (Chao et al. 1989; Liu and Tsunewaki 1991; Devos et al. 1992, 1993; Xie et al. 1993; Nelson et al. 1995a,b,c; Van Deynze et al. 1995a), barley (*Hordeum* sp.) (Heun et al. 1991; Kleinhofs et al. 1993; Graner et al. 1991, 1994), oat (*Avena* sp.) (O'Donoughue et al. 1992, 1995; Rayapati et al. 1995; Van Deynze et al. 1995b), maize (*Zea mays* L.) (Burr and Burr 1991; Gardiner et al. 1993), sorghum (*Sorghum* sp.) (Chittenden et al. 1994), and sugarcane (*Saccharum* sp.) (da Silva et al. 1993).

Comparative maps have been constructed for several grass genera. For species where linkage relationships have been conserved throughout the genome, comparative maps are known as consensus maps. These maps merge information about closely related species and are useful for cross-referencing genetic information from more distantly related species. A consensus map has been developed for Triticeae species, based on a common set of markers mapped onto the respective linkage groups of *T. aestivum*, *T. tauschii*, and *Hordeum* species (Nelson et al. 1995a,b; Van Deynze et al. 1995a,c).

Comparative maps in crop relatives within the Panicoideae subfamily, such as maize, sorghum, and sugarcane, and within the Pooideae subfamily, such as wheat, barley, and rye, have been developed based on cloned genes, random cDNA, and genomic DNA markers (Whitkus et al. 1992; Melake Berhan et al. 1993; Grivet et al. 1994; Naranjo et al. 1987; Devos et al. 1993). Substantial conservation of large segments of linkage groups was consistently observed within a subfamily. Within the Panicoideae, segments of most sorghum linkage groups showed homoeology to two independent segments in the maize genome (Whitkus et al. 1992; Melake Berhan et al. 1993), and the degree of conservation observed between sorghum and sugarcane was greater than between sorghum and maize (Grivet et al. 1994). In the Pooideae, only a few well-defined rearrangements distinguish the rye and wheat genomes (Naranjo et al. 1987; Devos et al. 1993).

Conservation of linkage relationships has also been demonstrated among more distantly related crops. Comparative maps of rice, maize, and wheat were developed by Ahn and Tanksley (1993), Wilson et al. (1998) and Ahn et al. (1993); maps of rice, maize, Triticeae species, and oat were published by Van Deynze et al. (1995a,b,c); maps of rice and wheat were presented by Kurata et al. (1994b), of rice and foxtail millet by Devos et al. (1998) and the relationship between wheat group 7 chromosomes and maize chromosome 9 was analyzed by Devos et al. (1994). Based on the information provided by these and other studies, Moore et al. (1995) proposed a conceptual framework for collating genetic information derived from six major grass species based on the alignment of 19 linkage segments from rice. This scheme has been revised based on higher resolution maps developed by Wilson et al. (Cornell University, personal communication).

Using data derived from global comparative mapping endeavors, several groups have reported the identification of putative orthologous loci in a range of grass genera. Lin et al. (1995) reported that quantitative trait loci (QTLs) with major effects on height and flowering in sorghum have counterparts in homeologous segments of the rice, wheat, barley, and maize genomes. Pereira and Lee (1995) reported the identification of genomic regions affecting plant height in sorghum and maize. Paterson et al. (1995) proposed that QTLs affecting seed size, non-shattering of grain, and photoperiod sensitivity are likely to be orthologous in sorghum, rice, and maize based on correspondence of map position across these genera. Harrington et al. (1996) mapped starch branching enzyme III in rice and used comparative maps to predict the location of this enzyme in maize, oat, and wheat. The maps that provide the basis for combining the genetic information available in these crops are based on probes which hybridize across the cereals.

The objective of the current study was to identify, map and end-sequence a set of cDNA probes that could be used by a wide range of researchers to detect anchor loci for comparative genome analysis among many members of the grass family.

Materials and methods

Selection of probes

Random cDNA clones from oat (CDO), rice (RZ), and barley (BCD) libraries were selected based on the reliability and intensity of hybridization signal when used as probes in Southern analysis of rice, wheat, barley, oat, maize, sorghum, and sugarcane DNA. Anchor probes were selected from surveys of over 1,800 clones based on the following criteria: (1) good hybridization signal in all or in most of these species, (2) apparently low or single copy in rice, (3) map position previously established in rice and maize, and/or in wheat, (4) uniform genome coverage in rice. Clones that improved genome coverage in all target crops were added to the set of Anchor Probes as they were identified. Probes were originally chosen based on the rice RFLP map because rice currently has one of the best characterized grass genomes (Causse et al. 1994; Kurata et al. 1994a; Harushima et al. 1998). It is a diploid with a large proportion of single-copy DNA (approximately 85% of the DNA behaves as single copy at high stringency, 0.5 × SSC at 65°C) (McCouch et al. 1988), and few patterns in the distribution of duplicated loci are evident (Causse et al. 1994). The BCD (barley cDNA), CDO (oat cDNA), WG (wheat genomic), and RZ (rice cDNA) probe libraries used in this study were described by Heun et al. (1991) and Causse et al. (1994), respectively.

RFLP and data analyses

Inserts from each of the probes were amplified using the polymerase chain reaction (PCR) with M13 primers and were hybridized (stringency of final wash = 0.5 × SSC at 65°C) to membranes (Hybond N+, Amersham) containing a single lane of rice (*Oryza sativa*, cv 'IR36'), oat (*Avena sativa*, cv 'Ogle'), barley (*Hordeum vulgare*, cv 'SE16'), wheat (*Triticum aestivum*, cv 'Chinese Spring'), sugarcane (*Saccharum spontaneum*, cv 'SE 208'), sorghum (*Sorghum bicolor*, cv 'BTx406'), and maize (*Zea mays*, cv 'CO159') DNA. DNA from each of the grasses was digested with *Eco*RI, and the amount used for Southern blotting was adjusted empirically to achieve uniform signal intensity with native probes for each of the DNA samples (approximately 8 ug/lane for rice, 15 ug/lane for oat, 12 ug/lane for barley, 15 ug/lane for wheat, 12 ug/lane for sugarcane, 8 ug/lane for sorghum, and 12 ug/lane for maize).

Loci detected as restriction fragment length polymorphisms with this set of anchor probes were placed on existing linkage maps for rice (Causse et al. 1944), maize (Ahn and Tanksley 1993), diploid oat (O'Donoughue et al. 1992; Van Deynze et al. 1995b), and wheat (Nelson et al. 1995a,b,c; Van Deynze et al. 1995a,c) using plant populations maintained at Cornell University. Polymorphic and non-polymorphic markers were assigned to wheat chromosomes or chromosome arms based on hybridization to nullitetrasomic (Sears 1966) and ditelosomic stocks (Sears and Sears 1978).

The same procedures for RFLP analyses (Heun et al. 1991; Causse et al. 1994) were used for each population. The individual linkage maps were developed by placing markers at a LOD threshold of 2.0 using MAPMAKER v3.0 (Lander et al. 1987) with the Kosambi mapping function (Kosambi 1944). The consensus map for wheat, *T. tauschii*, and *Hordeum* species (Van Deynze et al. 1995c), hereafter referred to as the Triticeae consensus map, was used as the basis for comparisons among these species and rice, maize, and oat.

Sequencing and sequence analysis

DNA sequencing was performed on an ABI 377 DNA sequencer (Applied Biosystems) using a dye terminator cycle sequencing ready reaction kit (Perkin Elmer) according to the manufacturer's instructions. An M13 universal forward or reverse primer was used unless otherwise indicated. Several strategies were employed to sequence probes with very long polyA tails. Most clones with long poly A tails were sequenced using an anchored degenerate primer [T30,(A,C,G)]. For clone CDO105, approximately 500 bp were removed from the 3′ end by *Exo*III deletion. The deleted clone was sequenced using the M13 forward primer. For clone CDO118, part of the 3′ end was deleted by *Pst*I digestion.

Sequences obtained for the 152 anchor clones were compared with each other using a local BLAST database and BLASTN (Gish et al. 1993; Altschul et al. 1997). This comparison was designed to detect inconsistencies between sequencing runs, overlaps between forward and reverse sequences, and similarities between clones that might be due to the presence of multigene families. Most clones were sequenced twice and, when necessary, discrepencies were resolved by additional sequencing. Alignments generated by CLUSTALW (Thompson et al. 1994) were reconciled with electropherograms to produce the consensus sequences that were deposited at GenBank (Accession numbers AA231638–AA231938) and are also available in the RiceGenes database (http://probe.nalusda.gov). The BLAST network service provided at the EMBL was used to identify relationships between the anchor probes and protein sequences for known genes. BLASTX searches were run on January 23, 1998 using The Baylor College of Medicine "Search Launcher" analysis client and scores exceeding 100 were considered significant.

Results

Probes

A set of 152 anchor probes (21 BCDs, 63 CDOs, 67 RZs and 1 WG) was chosen based on the criteria outlined in the previous section. The intensity of hybridization signal and copy number for each probe in each reference species were determined based on an evaluation of screening filters (Fig. 1). This data was complemented by the results obtained on mapping filters and is summarized in Table 1. The proportion of probes in the current anchor set producing detectable hybridization signals for rice and maize is biased upward because many of these clones had been previously mapped on these species.

Oat cDNAs produced strong hybridization signals more consistently in all seven species tested than clones from any other library (Fig. 2). All 63 oat cDNAs hybridized to oat and rice, 95–97% hybridized to maize and barley, 85–88% to sorghum and wheat, and 75% to sugarcane. All barley cDNAs hybridized to barley, wheat, oat, and maize, 90% to rice, 82% to sorghum, and only 58% to sugarcane. All 67 rice cDNAs hybridized to rice, 90% to sorghum, 82% to sugarcane, 78% to maize, and 48–55% to wheat, barley, and oat genomic DNA.

Genome coverage

Anchor probes meeting the criteria outlined in the "selection of probes" section are well distributed throughout the rice linkage map except for distal portions of chromosomes 2, 7, and 12. Markers in these regions (Causse et al. 1994) consisted of rice genomic and cDNA clones that did not hybridize well to other species. Three probes in the anchor set, BCD454, CDO127, and CDO507, detected duplicate loci in rice. The anchor probes also provide fairly uniform genome coverage for maize and oat. In Triticeae, distal portions of chromosomes have fewer markers, and there is

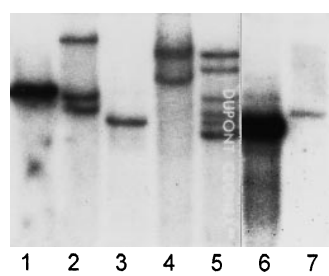

Fig. 1 Autoradiogram of CDO412 hybridized to rice ('IR36'), oat ('Ogle'), barley ('SE16'), wheat ('Chinese Spring'), sugarcane ('SES208'), sorghum ('BTx406'), and maize ('CO159') digested with *Eco*RI in *lanes 1–7*, respectively

Table 1 Information about anchor clones, including GenBank accession number, hybridization signal on garden blots, and chromosomal position in five genera

	Clone	ID	GenBank acc. no. F	GenBank acc. no. R	Insert size (kb)	Garden blots[a]							Chromosomal position					Hexaploid wheat arms	Wheat linkage map
						R	B	W	O	M	So	S	Rice	Maize	Diploid oat	Hexaploid oat	Barley		
1	BCD	98	AA231652	AA231667	1.3	G	G	G	G	G	G	G	8	1	B,D	–	5LS	1S,7S	1AS,7BS
2	BCD	134	AA231668	AA231669	0.85	G	G	G	G	G	G	G	1	3,8	C	–	–	3L	3DL
3	BCD	135	AA231670	AA231671	0.75	G	G	G	G	G	G	G	4	2,10	B	7,9	–	2L	2BL
4	BCD	147	AA231905	AA231672	0.95	F	G	G	F	G	G	F	8	10	D	–	–	3L	3BL
5	BCD	207	AA231835	AA231673	1.05	G	G	G	G	G	G	G	10	1,5	–	–	–	1L	–
6	BCD	348	AA231674	AA231675	0.75	F	G	G	G	G	G	ns	4	2,10	B	–	–	2S,4S,5,6S	2AS
7	BCD	349	AA231904	AA231676	1.5	G	G	G	G	–	F	ns	2	7	G	–	–	7S	7BS
8	BCD	386	AA231836	AA231678	1.1	G	G	G	G	G	G	F	10	1	A	–	–	1L	1BL
9	BCD	450	AA231679	AA231680	1.15	G	G	G	G	G	G	ns	3	1,2,5	E,F,G	–	–	5L	5BDL
10	BCD	454	AA231837	AA231681	0.95	G	G	G	G	G	G	G	3,5	6,8	A	–	–	1L	–
11	BCD	738	AA231872	AA231880	0.8	G	G	G	G	G	G	ns	5	6,8	–	–	–	1L	–
12	BCD	808	AA231685	AA231686	0.8	G	G	G	G	G	G	G	11	1,2,7	E	24,28	–	1L,4L	1AL
13	BCD	828	AA231838	AA231684	0.7	G	G	G	G	G	G	G	1	3,8	C	–	3L	3L	3AL
14	BCD	855	AA231839	AA231677	0.8	G	G	G	G	G	F	G	7	2,5,7,9		–	–	2S	2AS
15	BCD	880	AA231682	AA231683	0.8	G	G	G	G	–	G	G	2	–	G	33	–	6	6DL
16	BCD	926	AA231934	AA231935	1.4	G	G	G	G	G	G	F	9	2,7	E	–	–	5L	5AL
17	BCD	1072	AA231687	AA231688	1.2	G	G	G	G	G	G	F	5	1,5,10	A	–	–	1S	1AS
18	BCD	1087	AA231689	AA231690	0.7	G	G	G	G	G	F	F	9	2,7	E	–	–	5L	–
19	BCD	1261	AA231906	AA231691	0.85	G	G	G	G	–	G	G	5	–	A	1,3	–	1	1BDL
20	BCD	1421	AA231873	AA231857	0.8	G	G	G	G	–	G	G	5,7	–	E	–	–	4,5	5DL
21	BCD	1823	AA231692	AA231693	1.4	–	G	G	G	–	G	G		–	B	3,15	–	3	3AS
22	CDO	17	AA231858	AA231638	0.9	G	F	F	G	F	G	G	6	9	D	–	–	7	7AS
23	CDO	20	AA231642	AA231846	1.2	G	G	G	G	G	G	G	3	1,9	F	5	–	4L	4BL,5AL
24	CDO	36	AA231907	AA231751	1.2	G	F	G	G	G	G	G	4	2,10	C,E	–	1S	2L	2BDL
25	CDO	38	AA231754	AA231755	1.4	G	G	G	G	G	G	G	7	1,7	F,G	–	–	2S,4L	–
26	CDO	59	AA231859	AA231709	1.1	F	F	G	F	G	G	F	7	2,4,7	G	–	–	2S	–
27	CDO	78	AA231860	AA231717	1.5	G	G	G	G	G	G	G	6	9	D	–	–	7L	–
28	CDO	87	AA231720	AA231848	1.8	G	G	F	G	G	G	G	3	5	E	–	–	5L	–
29	CDO	89	AA231721	AA231722	1.2	G	G	G	G	G	G	G	5	6	D	–	–	1L	1ADL
30	CDO	98	AA231728	AA231729	2.3	G	G	G	G	G	G	G	10	1,5	A	–	–	1L	1ABL
31	CDO	105	AA231840	AA231700	2.3	G	G	G	G	G	G	G	5	3	G	–	3L,5L	1L	3BL
32	CDO	118	AA231932	AA231845	1.7	G	G	G	G	G	G	G	1	3	C	–	–	3L	3AL
33	CDO	122	AA231653	AA231694	1.2	G	G	G	G	G	G	G	3	1,5	E	28	–	4S	–
34	CDO	127	AA231695	AA231696	1.2	G	G	G	G	–	G	G	11,12	4,10	F	–	–	1S,4L,5L	1BS
35	CDO	202	AA231639	AA231908	1	G	G	G	G	G	G	G	5	2,5,6,8	A	–	–	1L	–
36	CDO	204	AA231640	AA231641	1.4	G	G	F	G	F	F	F	2,6	–	–	–	–	–	6AL
37	CDO	241	AA231909	AA231643	1.1	G	F	G	G	G	G	G	4	3	E	–	–	6L	–
38	CDO	244	AA231644	AA231645	1.5	G	G	G	G	G	F	G	4	2,4,10	B	–	–	–	–
39	CDO	328	AA231646	AA231647	1.5	G	ns	G	G	G	G	F	1	8	C	–	–	3S	3BS
40	CDO	344	AA231910	AA231648	1.25	G	G	G	G	G	G	G	12	1,3,4	F	30	–	5S	–
41	CDO	345	AA231649	AA231650	1.4	G	G	F	F	G	G	F	1	3,5	G	–	–	3L	–
42	CDO	365	AA231651	AA231882	1.3	G	G	G	G	G	G	G	11	4	B	3	–	6S	–
43	CDO	385	AA231752	AA231753	1.2	G	G	G	G	G	G	G	7	2	B	–	–	2S,5L	–
44	CDO	393	AA231847	AA231911	1.1	G	G	G	G	G	G	G	5	6	A	1,3,5	5L	1L	1AL
45	CDO	395	AA231756	AA231757	1.15	G	G	G	G	G	F	G	1,2	4,5	B,A	3,32	3S	2L,3S	3AS

Table 1 Continued

	Clone	ID	GenBank acc. no. F	GenBank acc. no. R	Insert size (kb)	Garden blots[a]							Chromosomal position					Hexaploid wheat arms	Wheat linkage map
						R	B	W	O	M	So	S	Rice	Maize	Diploid oat	Hexaploid oat	Barley		
46	CDO	400	AA231758	AA231759	0.9	G	G	G	G	G	G	G	2	4,5	E	14	–	1L,5L	–
47	CDO	405	AA231912	AA231760	1.8	G	G	G	G	G	G	G	7	7	–	8	–	2S	2ABDS
48	CDO	407	AA231761	AA231762	0.9	G	G	G	G	G	G	G	7	7	D	–	–	3S	3DS
49	CDO	412	AA231763	AA231701	1.5	G	G	G	G	G	G	G	9	2,7	E	23,25	–	5	5ABDL
50	CDO	455	AA231871	AA231702	1.1	G	G	G	G	G	F	F	1	3,8	G		–	3L	–
51	CDO	456	AA231913	AA231914	1.4	G	G	G	G	G	G	G	4	2,5,10	B	9	–	2S	2ADS,U
52	CDO	457	AA231915	AA231862	1.5	G	G	G	G	G	F	F	3	1,5	E	24,28	7L	5L	5AL
53	CDO	459	AA231916	AA231917	1	G	G	G	G	G	F	F	12	3	F	–	–	–	–
54	CDO	460	AA231870	AA231918	1.3	G	G	G	G	–	–	G	1	8	C	13,15	7LS	3S	3ADS
55	CDO	470	AA231703	AA231887	1.1	G	G	F	G	G	G	G	1	8	F	–	–	–	–
56	CDO	475	AA231764	AA231884	1.9	G	G	G	G	G	F	G	6	1,4,5,9	–	–	1S	4,7	4AL,7AS
57	CDO	497	AA231919	AA231920	1.3	G	G	F	G	G	G	F	7	4	G	–	6L	6L	–
58	CDO	507	AA231844	AA231704	1.8	G	G	G	G	G	G	G	2,5	4,5	G	–	–	6L	6BL
59	CDO	516	AA231828	AA231921	1.6	G	G	F	G	G	G	G	2	5	D,G	–	–	6	–
60	CDO	520	AA231922	AA231923	1.6	G	G	F	G	G	G	F	11	4	A,E	–	–	5L	–
61	CDO	524	AA231938	AA231710	1	G	G	G	G	G	G	G	2	2	B	–	–	6S	6BS
62	CDO	534	AA231712	AA231705	1	G	G	G	G	G	G	F	2,11	4,5	B	–	–	1S,3LS,6S	6DS
63	CDO	545	AA231869	AA231713	2.4	G	G	G	G	F	G	G	6	6	D	17	1S	–	4AL,7AS
64	CDO	580	AA231888	AA231714	1.2	G	G	G	G	G	G	F	5	6,8	A	4	–	1S	1AS
65	CDO	590	AA231889	AA231711	1.3	G	G	G	G	G	G	G	9	–	A	3,24	–	5L	–
66	CDO	595	AA231924	AA231925	1.4	G	G	G	F	G	G	F	8	1	D	6	1S	7S	7AS
67	CDO	686	AA231937	AA231936	1.4	G	G	G	G	G	F	G	2	5	D	–	–	7L	7BL
68	CDO	718	AA231715	na[c]	1.1	G	G	G	G	G	G	G	2	5	A	4,6	–	3L	3BL
69	CDO	783	AA231868	AA231716	1.4	G	G	G	G	ns	G	G	4	2,10	–	–	–	2S,5,6L	–
70	CDO	795	AA231718	AA231719	1.5	G	G	G	G	G	G	G	3	1,5	E	22,24	4L	4S	4BS
71	CDO	920	AA231890	AA231842	1.1	G	G	G	G	G	F	F	1	3,8	C	–	–	2L,3L,7L	–
72	CDO	938	AA231725	AA231849	1.1	G	G	G	G	G	G	G	3	1,2,9	F	30	–	4L	4BL
73	CDO	941	AA231723	AA231724	0.85	G	G	F	G	G	F	ns	4	2	B,C	–	–	2L	–
74	CDO	962	AA231726	AA231727	1.2	G	G	G	G	G	G	G	1	3,8	D	7,17	–	7L	7AL
75	CDO	1081	AA231699	AA231841	1.1	G	G	G	G	G	G	G	3	1,8	F	5	–	4L	4DL
76	CDO	1160	AA231697	AA231698	1.3	G	G	G	G	G	G	F	5	3,6,8	C	–	–	1L,3L	1AL
77	CDO	1173	AA231926	AA231927	1.5	G	G	G	G	G	G	ns	5	5,8	A	4,5	–	1S	1BS
78	CDO	1328	AA231834	AA231830	1.1	G	G	G	G	G	G	G	4	2,10	B	10	–	2L	–
79	CDO	1338	AA231928	AA231929	1.2	G	G	G	G	F	G	G	11	3,10	E	–	–	4S,5S	5BS
80	CDO	1380	AA231886	AA231833	1.2	G	G	G	G	G	G	G	2	4,5	G	36	–	6L	–
81	CDO	1387	AA231881	AA231831	0.8	G	G	G	G	G	G	G	3	1,9	F	–	–	4L	4AS
82	CDO	1395	AA231832	AA231829	2	G	G	G	G	G	G	G	3	1,4,5,9	F	–	–	4L,7S	7AS
83	CDO	1417	AA231706	AA231707	1.8	G	G	G	G	G	G	G	4	2,10	B	–	–	2L	–
84	CDO	1508	AA231930	AA231708	0.9	G	G	G	G	F	G	G	9	–	E	23,27	–	–	5DL
85	RZ	2	AA231745	AA231863	1.55	G	G	G	G	G	G	G	6	5,6,9	D	–	–	5L	7DS
86	RZ	14	AA231891	AA231778	1.35	G	G	G	G	G	G	F	1	3	–	–	–	–	–
87	RZ	53	AA231931	AA231852	0.95	G	ns	ns	ns	G	G	ns	4	4,5	–	–	–	–	–
88	RZ	69	AA231805	AA231806	1.1	G	G	G	G	G	G	G	4	10	C	11	–	2S	2BS
89	RZ	87	AA231654	AA231821	1	G	F	F	F	G	F	F	2	5	B	–	–	–	–
90	RZ	141	AA231883	AA231730	1.2	G	F	F	F	G	G	ns	11	3,10	E	–	–	–	–

Table 1 Continued

	Clone	ID	GenBank acc. no. F	GenBank acc. no. R	Insert size (kb)	Garden blots[a]							Chromosomal position					Hexaploid wheat arms	Wheat linkage map
						R	B	W	O	M	So	S	Rice	Maize	Diploid oat	Hexaploid oat	Barley		
91	RZ	143	AA231892	AA231731	0.7	G	F	ns	F	G	G	G	8	4,6	–	–	–	–	–
92	RZ	144	AA231732	AA231777	1.7	G	F	F	F	G	G	G	6	6,9	–	–	–	–	–
93	RZ	166	AA231733	AA231734	1.5	G	F	F	F	G	G	G	2	5	–	–	–	1S	1BS
94	RZ	206	AA231735	AA231736	0.8	G	ns	F	ns	G	F	F	9	2,6,8	–	–	–	1L,5L	–
95	RZ	242	AA231737	AA231893	0.85	G	–	G	G	G	F	F	6	6	D	–	–	7L	–
96	RZ	244	AA231738	AA231739	0.9	G	G	G	G	G	G	G	5	3,8	A,C	–	–	–	1AS
97	RZ	251	AA231894	AA231853	1.6	G	F	G	F	G	G	G	3	–	F	–	–	4L	4BL
98	RZ	261	AA231655	AA231740	1.2	G	G	G	G	G	G	G	12	10	F	–	–	–	–
99	RZ	273	AA231741	AA231742	1.8	G	G	G	G	G	G	G	2	4,5	G	–	–	6L	6BL
100	RZ	296	AA231743	AA231744	3	G	ns	ns	G	G	G	G	5	1,3	–	–	–	–	–
101	RZ	323	AA231861	AA231854	1.75	G	ns	F	F	G	G	G	8	1,8	–	–	–	–	–
102	RZ	329	AA231746	AA231747	1.7	G	F	F	G	G	G	G	3	4,9	–	–	–	–	–
103	RZ	382	AA231866	AA231748	1.77	G	G	G	F	G	G	G	1	8	–	–	–	–	–
104	RZ	387	AA231749	AA231750	1.1	G	F	F	F	F	F	ns	7	2,10	B	–	–	–	–
105	RZ	390	AA231765	AA231766	1.1	G	F	G	G	G	G	G	5	3,8	A	–	–	–	–
106	RZ	395	AA231885	AA231767	1.6	G	G	G	G	G	G	G	7	7	E	–	–	2S	2AS,5ADL
107	RZ	397	AA231768	AA231769	0.5	G	F	F	F	F	F	F	12	–	D	–	–	–	–
108	RZ	400	AA231895	AA231896	0.9	G	G	G	G	F	G	G	10	–	F	–	–	–	–
109	RZ	403	AA231867	AA231770	0.7	G	–	–	–	F	F	G	3	1	–	–	–	–	–
110	RZ	404	AA231874	AA231771	0.85	G	ns	ns	ns	G	G	G	9	–	E	–	–	–	–
111	RZ	413	AA231779	AA231780	1.6	G	G	G	G	G	G	G	1	7	–	–	–	–	–
112	RZ	421	AA231781	AA231772	1.15	G	G	G	F	G	G	G	10	1	–	–	–	–	–
113	RZ	444	AA231782	AA231783	1.1	G	G	G	G	G	G	G	1	3,8	C	–	–	1L,3L,4S	2BS,2DL
114	RZ	446	AA231784	AA231875	0.8	G	F	F	F	G	G	G	2	4,5	G	–	–	–	–
115	RZ	455	AA231785	AA231786	1	G	G	G	G	G	G	G	5	6,8	–	–	–	–	–
116	RZ	474	AA231897	AA231787	0.8	G	ns	ns	ns	ns	F	ns	3	–	–	–	–	–	–
117	RZ	476	AA231773	AA231898	1.3	G	F	G	G	G	G	G	2	4,5,9	B	–	–	6S/7L	7BL
118	RZ	500	AA231788	AA231789	0.8	G	G	G	G	G	G	G	10	1	A	–	–	1L	–
119	RZ	508	AA231790	AA231899	0.87	G	G	G	G	G	G	G	6	5	D	–	–	7L	7BL
120	RZ	509	Same as R[b]	AA231791	0.6	G	F	F	F	F	–	F	7	2,7	–	–	–	–	–
121	RZ	516	AA231792	AA231855	0.69	G	F	F	F	F	G	G	6	9	D	–	–	–	–
122	RZ	525	AA231843	AA231876	0.55	G	ns	ns	ns	ns	ns	F	11	–	–	–	–	–	–
123	RZ	537	AA231656	AA231850	1.5	G	F	F	G	F	F	F	11	–	–	–	–	–	–
124	RZ	538	AA231793	AA231794	0.9	G	G	G	G	G	G	G	1	3,8	–	–	–	3L	–
125	RZ	543	AA231900	AA231901	0.85	G	G	G	G	G	G	G	1	3,8	C	–	–	3S	–
126	RZ	567	AA231774	na	1.1	G	G	G	G	G	G	G	2	4,5	–	–	–	6L	–
127	RZ	574	AA231813	AA231814	1.5	G	G	G	G	G	G	G	3	1	A	–	–	4L	4AS
128	RZ	583	AA231657	AA231658	0.8	G	G	G	G	G	G	G	10	5	–	–	–	1L	–
129	RZ	588	AA231815	AA231816	1.1	G	ns	F	ns	G	G	G	6	6,9	–	–	–	7	–
130	RZ	590	AA231865	AA231856	0.9	G	ns	F	ns	G	G	G	4	2,10	–	–	–	–	–
131	RZ	596	AA231659	AA231817	1.15	G	G	F	G	G	G	G	9	7	E	–	–	5	–
132	RZ	599	AA231864	AA231775	0.9	G	F	F	F	F	G	G	2	5	–	–	–	–	–
133	RZ	612	AA231776	AA231795	0.9	G	F	F	ns	G	G	G	6	9	D	–	–	–	–
134	RZ	614	AA231660	AA231796	1	G	ns	ns	ns	G	G	G	3	–	E	–	–	–	–
135	RZ	630	AA231797	AA231798	1.1	G	ns	ns	F	G	G	G	3	1,3,4,5	–	–	–	–	–

Table 1 Continued

	Clone	ID	GenBank acc. no. F	GenBank acc. no. R	Insert size (kb)	Garden blots[a]							Chromosomal position					Hexaploid wheat arms	Wheat linkage map
						R	B	W	O	M	So	S	Rice	Maize	Diploid oat	Hexaploid oat	Barley		
136	RZ	670	AA231799	AA231800	0.6	G	G	G	G	–	G	G	12	–	–	–	–	–	–
137	RZ	672	AA231801	AA231933	1.65	G	G	G	G	G	G	F	3	1,9	U	–	–	–	4BL
138	RZ	682	AA231902	AA231802	1	G	F	F	F	G	G	F	6	9	D	–	–	7L	7AL
139	RZ	698	AA231803	AA231804	0.7	G	F	G	F	G	G	G	9	1,2,6	U	–	–	5	–
140	RZ	740	AA231807	AA231818	1.15	G	G	G	G	G	G	G	4	–	–	–	–	2L	–
141	RZ	753	AA231903	AA231808	1.55	G	ns	ns	ns	G	G	G	7	7	–	–	–	–	–
142	RZ	783	AA231661	AA231809	0.7	G	G	G	F	G	G	G	1	–	–	–	–	–	–
143	RZ	797	AA231810	AA231811	1.6	G	ns	ns	F	F	G	ns	11	–	–	–	–	–	–
144	RZ	816	AA231819	AA231820	0.8	G	F	F	F	F	F	F	12	–	–	–	–	–	–
145	RZ	836	AA231662	AA231851	1.2	G	F	F	F	F	G	F	1	–	–	–	–	–	–
146	RZ	892	AA231822	AA231823	1.6	G	F	F	F	F	G	G	10	–	A	–	–	1S	–
147	RZ	900	AA231824	AA231825	1	G	–	–	–	G	G	G	11	4	–	–	–	–	–
148	RZ	912	AA231663	AA231664	1.5	G	–	–	–	G	G	G	3	1,5	–	–	–	–	–
149	RZ	952	AA231812	AA231877	1.9	G	F	G	F	G	G	G	8	1	D	–	–	–	–
150	RZ	953	AA231826	AA231878	1.4	G	ns	ns	ns	ns	G	G	6	–	–	–	–	–	–
151	RZ	995	AA231827	AA231879	0.95	G	G	G	G	G	G	G	1	3,8	–	–	–	–	6BS
152	WG	110	AA231665	AA231666	1	G	G	G	G	–	–	–	1	–	G	3,34	3L	3L	3DL,4BL

[a] R, Rice; B, barley; W, wheat; O, oat; M, maize; So, sorghum; S, sugarcane
[b] Clone RZ509 was sequenced in a single pass (Acc. no. AA231791)
[c] na, Sequence not available
[d] G, Good signal; F, fair signal; ns, no signal on Southern hybridization

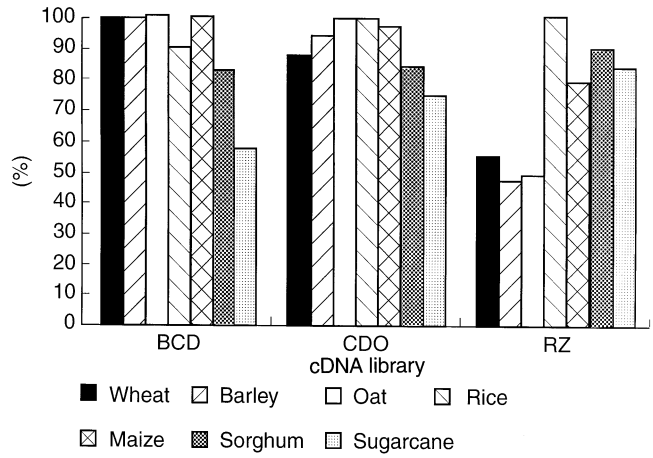

Fig. 2 Proportion of clones providing good hybridization signal (stringency = 0.5 × SSC at 65°C) for selected barley (*BCD*), oat (*CDO*), and rice (*RZ*) cDNA clones

clustering of marker loci near centromeres on six of the seven chromosomes. Although fewer anchor probes were mapped in wheat due to limited polymorphism, most fragments having good hybridization signal (polymorphic and nonpolymorphic) were assigned to wheat chromosomes or chromosome arms with nullitetrasomic and/or ditelosomic stocks (Table 1).

Identification of homoeologous segments across grass genera

Single species maps portray static images of single genome configurations of the Anchor Probe loci. To identify and visualize conserved linkage blocks (homeologous segments) in multiple genera simultaneously, we have developed an interactive display in the RiceGenes database (http: //probe.nalusda.gov) (Paul et al. 1998). As illustrated in Fig. 3, this display allows the user to clearly observe the alignment of homoeologous segments from multiple grass genera in relation to rice chromosomes. In the left portion of the figure, maize chromosome 1 is shown in terms of segments shared between maize and rice. Rice chromosome 3 is shown on the right, with vertical lines indicating conserved segments with other grass species (R3-OF indicates a region shared between rice chromosome 3 and oat chromosome F, R3-M9 a region shared between rice 3 and maize 9, and so on). The display also encourages the user to contemplate internal duplications within a single genome, as evidenced by the regions "R3-M1" and "R3-M5" on the lower part of rice chromosome 3, whose markers appear on both maize chromosome 1 and maize chromosome 5. Upon clicking on one of the thin vertical bands indicating a conserved segment, the individual markers on the rice map that define the syntenic region in the related genus are highlighted (see R3-M1-3). By double clicking on the vertical band, additional information, including notes about genes or QTLs occurring in the region, can be readily accessed (i.e. the Chrom_Block window in the foreground in Fig. 3). Alternatively, clicking on the R3-M1-1 label moves the user directly to the Cornell University (CU) maize map (Ahn and Tanksley, 1993), where synteny with rice can be visualized in relation to maize-based information. This dynamic display provides a highly informative look at the evolutionary relationships among the grasses.

Gene identification

Table 2 summarizes the results of an analysis in which probe sequences were compared using BLASTX (Gish et al. 1993; Altschul et al. 1997) to the protein sequence database maintained at the EMBL. Out of 152 anchor probes, 119 (78%) showed significant similarity to previously characterized genes or putative genes in other organisms.

Discussion

Identification of a set of anchor probes

The availability of a set of probes that can be used to simultaneously "anchor" loci on species-specific maps offers a point of departure for the development of increasingly comprehensive comparative maps in a wide range of species and genera. The use of a common set of anchor probes for mapping of grass family members allows the immediate transfer of genetic information across species and genera of this important plant family and has profound implications for practical applications in plant improvement. Probes that are most informative for comparative mapping in grasses will (1) hybridize to the majority of target species, (2) be single or low copy in a reference species (such as rice), and (3) provide good genome coverage in all species. The use of these probes in linkage analysis and sequence-based studies of more distantly related organisms will provide an informative link across greater evolutionary distances.

This set of anchor probes provides good coverage of the linkage maps of rice, maize, and oat. While linkage data are not available for 84 of the 152 probes for Triticeae species, mainly due to the low polymorphism in cultivated wheat, this is compensated by the use of aneuploid stocks that provide arm locations for nearly all hybridizing fragments (Table 1). For example, loci detected with CDO122 map to the same relative positions in homoeologous segments of rice chromosome 3, maize 1 and 5, and oat E (Fig. 3). If the positions of orthologous loci detected with this probe were conserved relative to Triticeae species, the distal

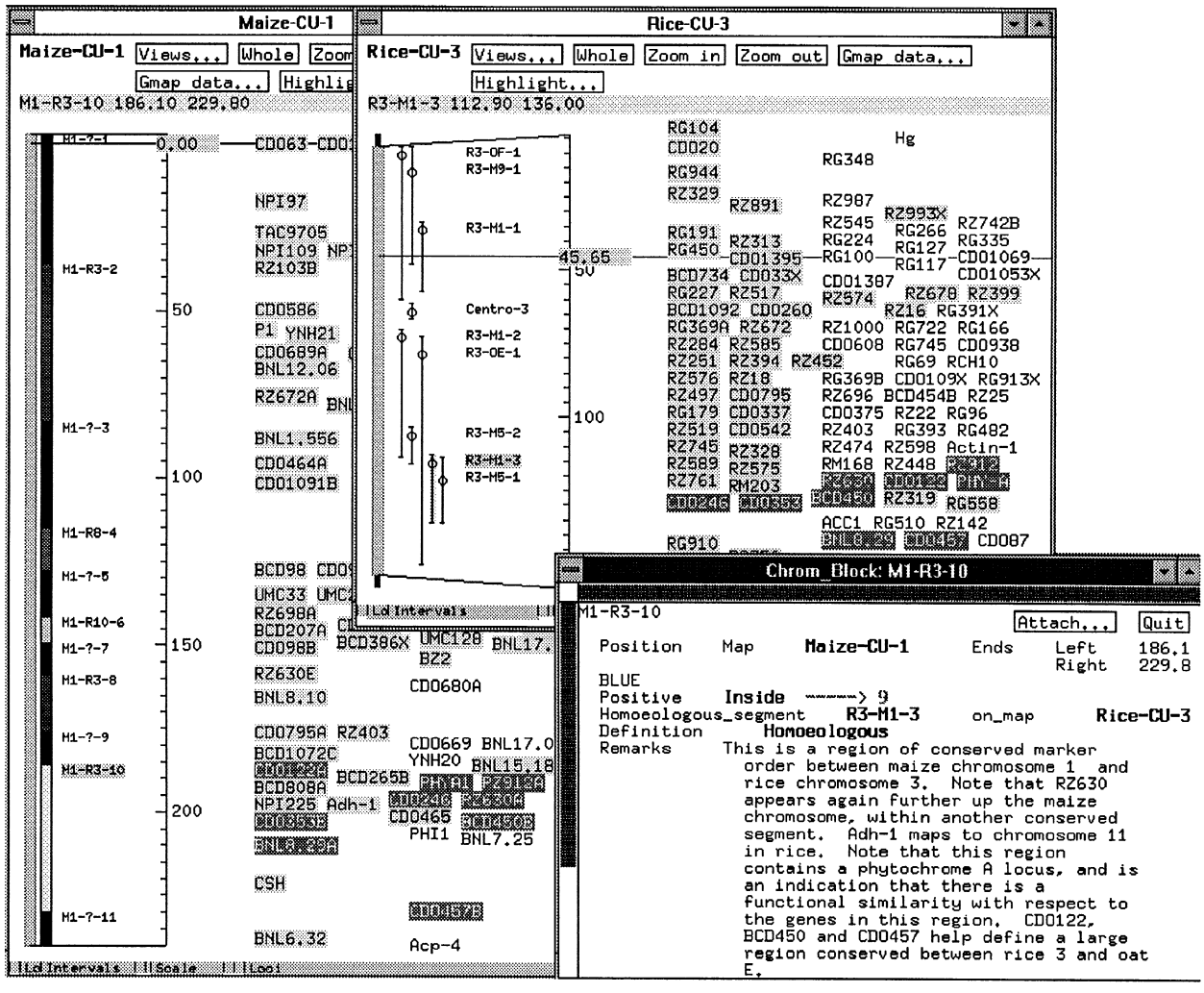

Fig. 3 Screen shot of comparative map display in the RiceGenes database (Paul et al. 1998). Displays are hot-linked to each other and accessible by clicking on segments of interest. *Left window* shows maize chromosome 1, color-coded to indicate regions of homoeology with rice. "M1-R3-10" region is *highlighted* and when clicked, opens the *lower right window*, displaying additional information about the conserved region, including positions of noteworthy genes/QTLs that may be homologous in grasses. Clicking on Rice-CU-3 brings up the *right window* which depicts rice chromosome 3 with homoeologous segments from maize and oat overlayed as *thin vertical lines* to the *right* of the chromosome bar

portion of Triticeae chromosome 4S would be better represented (Van Deynze et al. 1995c). The positions of these markers must be confirmed by mapping in additional populations of Triticeae species or by using the deletion stocks developed by Werner et al. (1992). Additional probes mapping to poorly represented regions in the Triticeae should be screened to meet the criteria and added to this set of anchor probes in the future.

Sequence data for these anchor probes serve several roles. First, these cDNAs can be tentatively associated with protein products and putative functions based on structural similarity with genes from a wide array of organisms (Table 2). This information can be used to help clarify the relationship between candidate genes and traits of agronomic importance (governed by single genes and quantitative trait loci, or QTLs) using positional information about these genes/QTLs in several species.

Second, the sequence information generated in this study provided immediate points of integration between the Japanese rice map (Kurata et al. 1994a) and the four maps presented in this study; 8 cDNAs from the libraries developed at Cornell were homologous to cDNAs on the high-density rice map developed independently in Japan. The 8 Japanese clones were isolated from a rice callus library, while 5 of the Cornell counterparts were derived from an etiolated leaf library of rice and 3 from an etiolated leaf library of oat (*Avena*). In addition, 6 homologues were identified among rice and barley or rice and oat cDNAs derived from etiolated leaf tissue libraries developed at Cornell. These clones had been previously located on species-specific maps, and clarification of their sequence identity converted them into the basis for comparisons of allelic divergence among the grass genera.

Table 2 BLASTX similarities for anchor probes. The 119 anchor probes with BLAST scores greater than 100 are listed along with their GenBank accession numbers, BLASTX score, P-value with number of "High-Scoring Segment Pairs" in parentheses, match accession number and description. The search was conducted on Jan. 23, 1998 using the default BLASTX parameters. "F" and "R" indicate forward and reverse sequences, respectively; "ALL" indicates that the entire clone was sequenced in a single pass. "GB" = GenBank; "SP" = SwissProt; "PIR" = Protein Identification Resource

Clone[a]	Accession	Score	P-value	Match[b]	Description
BCD98.F	AA231652	175	5.5e-12 (1)	SP:O04503	Possible gene in BAC F21M12.20 [*Arabidopsis thaliana*]
BCD134.F	AA231668	131	4.2e-07 (1)	SP:Q20445	Possible gene in cosmid F46B6.3 [*Caenorhabditis elegans*]
BCD147.R	AA231672	240	2.6e-19 (1)	SP:P91917	Possible gene in cosmid W08E3.3 [*Caenorhabditis elegans*]
BCD207.F	AA231835	101	0.00043 (1)	SP:Q99962	Protein containing SH3 domain, SH3GL2 [*Homo sapiens*]
BCD454.R	AA231681	314	2.8e-26 (1)	SP:P11155	Pyruvate, orthophosphate dikinase [*Zea mays*]
BCD738.F	AA231872	205	5.9e-20 (2)	SP:P12782	Phosphoglycerate kinase [*Triticum aestivum*]
BCD738.R	AA231880	310	1.0e-36 (2)	SP:P12782	Phosphoglycerate kinase [*Triticum aestivum*]
BCD808.F	AA231685	477	9.5e-45 (1)	SP:Q96459	Alpha tubulin [*Hordeum vulgare*]
BCD808.R	AA231686	164	9.7e-11 (1)	SP:Q96459	Alpha tubulin [*Hordeum vulgare*]
BCD828.R	AA231684	278	6.2e-23 (1)	SP:Q41534	ATP synthase beta subunit [*T. aestivum*]
BCD880.F	AA231682	327	1.2e-28 (1)	SP:P48642	Blutathione reductase, cytosolic [*Oryza sativa*]
BCD880.R	AA231683	157	6.7e-10 (1)	SP:P48642	Blutathione reductase, cytosolic [*Oryza sativa*]
BCD855.R	AA231677	170	7.1e-11 (1)	SP:Z97340_23	Similar to extensin class 1 [*Arabidopsis thaliana*]
BCD926.F	AA231934	434	3.5e-40 (1)	PIR:S17200	Protein kinase [*Oryza sativa*]
BCD1072.F	AA231687	104	0.00055 (1)	SP:O04056	Heat shock protein 70 precursor [*Citrullus lanatus*]
BCD1072.R	AA231688	181	3.8e-28 (2)	SP:Q08080	Chloroplast stroma 70 kd heat shock-related protein [*Spinacia oleracea*]
CDO17.R	AA231638	108	0.00025 (1)	SP:Q21453	Possible gene in cosmid M01F1.4 [*Caenorhabditis elegans*]
CDO20.F	AA231642	167	6.9e-12 (1)	SP:Q93499	Possible gene in cosmid F15C11.2 [*Caenorhabditis elegans*]
CDO20.R	AA231846	129	2.5e-07 (1)	SP:Z98262_1	Possible gene in cosmid F15C11.2 [*Caenorhabditis elegans*]
CDO36.R	AA231751	134	1.1e-07 (1)	SP:AF006493_1	Acyl-CoA binding protein ACBP/ECHM [*Cyprinus carpio*]
CDO38.R	AA231755	253	1.8e-20 (1)	SP:P52914	Nucleoside triphosphatase [*Pisum sativum*]
CDO59.F	AA231859	145	1.5e-09 (1)	SP:P56330	Translation initiation factor [*Zea mays*]
CDO89.R	AA231722	308	2.2e-26 (1)	SP:O08616	Beta-alanine-pyruvate aminotransferase [*Rattus norvegicus*]
CDO98.R	AA231729	146	2.7e-08 (1)	SP:Q99189	mRNA transport regulator Mtr10p [*Saccharomyces cerevisiae*]
CDO105.R	AA231700	281	2.8e-22 (1)	SP:Z97336_6	TMV resistance protein homolog [*Arabidopsis thaliana*]
CDO118.R	AA231845	126	1.9e-06 (1)	SP:P53632	Topoisomerase 1-related protein [*Saccharomyces cerevisiae*]
CDO122.R	AA231694	275	2.6e-30 (2)	SP:Q43644	76 kd mitochondrial complex I subunit precursor [*Solanum tuberosum*]
CDO127.R	AA231696	191	1.3e-13 (1)	SP:O06134	Pyruvate kinase (Pyk) [*Mycobacterium tuberculosis*]
CDO202.F	AA231639	407	2.5e-37 (1)	SP:O04619	Similar to mitochondrial carrier family [*Arabidopsis thaliana*]
CDO202.R	AA231908	230	1.7e-18 (1)	SP:O04619	Similar to mitochondrial carrier family [*Arabidopsis thaliana*]
CDO244.R	AA231645	139	3.5e-09 (2)	SP:U82481_1	KI domain interacting kinase 1 [*Zea mays*]
CDO344.F	AA231910	135	2.0e-07 (1)	SP:Y11336_1	RGA1 protein [*Arabidopsis thaliana*]
CDO344.R	AA231648	137	1.2e-07 (1)	SP:Y11336_1	RGA1 protein [*Arabidopsis thaliana*]
CDO365.R	AA231882	275	3.3e-23 (1)	SP:Z99107_117	YerN gene [*Bacillus subtilis*]
CDO385.F	AA231752	100	0.00065 (1)	SP:Q19437	Similar to beta-ureidopropionase [*Caenorhabditis elegans*]
CDO385.R	AA231753	148	4.0e-09 (1)	SP:Q03248	Beta-alanine synthase [*Rattus norvegicus*]
CDO393.R	AA231911	245	9.0e-19 (1)	SP:P33194	Possible DNA-repair protein [*Cercopithecus aethiops*]
CDO395.F	AA231756	118	1.4e-06 (1)	SP:Q05737	GTP-binding protein yptm2 gene product [*Zea mays*]
CDO395.R	AA231757	198	3.5e-15 (1)	SP:P40392	Ras-related GTP binding protein [*Oryza sativa*]
CDO400.R	AA231759	154	1.6e-10 (1)	SP:P27924	Huntingtin interacting protein [*Homo sapiens*]
CDO405.R	AA231760	331	2.9e-27 (1)	SP:AC003952_5	Hypothetical protein [*Arabidopsis thaliana*]
CDO407.R	AA231762	157	6.0e-10 (1)	GB:U10400	Gene product Ysc84p [*Saccharomyces cerevisiae*]
CDO412.R	AA231701	292	1.2e-24 (1)	SP:Z99707_20	Hypothetical 55.1 KD protein [*Arabidopsis thaliana*]
CDO456.R	AA231914	206	2.1e-15 (1)	SP:Z97336_19	Hypothetical 47.3 kd protein [*Arabidopsis thaliana*]
CDO457.F	AA231915	133	2.6e-07 (1)	SP:Z99708_19	Myo-inositol transport protein homolog [*Arabidopsis thaliana*]
CDO470.F	AA231703	126	1.5e-06 (1)	SP:Q12874	Splicesomal protein (SAP 61) [*Homo sapiens*]
CDO507.R	AA231704	227	6.6e-18 (1)	SP:P31691	ATP/ADP translocator [*Oryza sativa*]
CDO524.R	AA231710	191	2.7e-13 (1)	SP:Q04585	Probable membrane protein YDR109c [*Saccharomyces cerevisiae*]
CDO534.F	AA231712	121	4.6e-06 (1)	SP:X84226_1	Citrate synthase [*Nicotiana tabacum*]
CDO534.R	AA231705	481	3.6e-45 (1)	SP:P49298	Citrate synthase precursor [*Citrus maxima*]
CDO580.F	AA231888	193	5.6e-14 (1)	SP:AF038605_4	Similar to acyl-CoA dehydrogenase [*Caenorhabditis elegans*]
CDO580.R	AA231714	276	2.8e-23 (1)	SP:P26440	Isovaleryl-coA dehydrogenase (IVD) [*Homo sapiens*]
CDO590.R	AA231711	235	9.6e-19 (1)	SP:P36875	Phosphoprotein phosphatase 2 A 65 kDa regulatory subunit [*Pisum sativum*]
CDO686.R	AA231936	315	4.5e-51 (2)	SP:P52903	E1 alpha subunit of pyruvate dehydrogenase precursor [*Solanum tuberosum*]
CDO718.F	AA231715	108	9.8e-05 (1)	SP:P78784	Protein similar to yeast Lph16p [*Saccharomyces cerevisiae*]
CDO783.R	AA231716	125	6.5e-12 (2)	SP:AC002329_4	Possible gene in BAC F5J6 from chromosome IV [*Arabidopsis thaliana*]
CDO795.R	AA231719	113	3.6e-06 (1)	SP:P74262	Hypothetical 13.0 KD protein [*Synechocystis sp.*]
CDO920.R	AA231842	159	1.4e-11 (2)	SP:P92030	egl [*Drosophila melanogaster*]

Table 2 Continued

Clone[a]	Accession	Score	P-value	Match[b]	Description
CDO938.R	AA231849	166	3.0e-23 (2)	SP:U87163_1	eIF-2 beta subunit [*Triticum aestivum*]
CDO1081.F	AA231699	112	3.1e-05 (1)	SP:AC002342_6	Putative polygalacturonase [*Arabidopsis thaliana*]
CDO1081.R	AA231841	154	5.0e-10 (1)	SP:P27644	Polygalacturonase [*Agrobacterium tumefaciens*]
CDO1160.F	AA231697	114	3.7e-05 (1)	SP:Q42559	Ketol-acid reductoisomerase [*Arabidopsis thaliana*]
CDO1160.R	AA231698	350	1.2e-32 (2)	SP:Q01292	Acetohydroxy acid isomeroreductase [*Arabidopsis thaliana*]
CDO1328.R	AA231830	346	4.2e-33 (2)	SP:AF028842_1	DegP protease precursor [*Arabidopsis thaliana*]
CDO1338.R	AA231929	214	8.9e-16 (1)	GB:X94753	Hexosephosphate aminotransferase [*Candida albicans*]
CDO1387.R	AA231831	266	1.6e-21 (1)	SP:U94831_1	Multispanning membrane protein [*Homo sapiens*]
CDO1417.R	AA231707	126	2.9e-06 (1)	SP:P14680.	Yak1 kinase [*Saccharomyces cerevisiae*]
RZ2.R	AA231863	460	3.1e-41 (1)	SP:AB007510_1	PRP8 protein [*Homo sapiens*]
RZ69.R	AA231806	142	3.0e-09 (1)	GP:S78994	Copper amine oxidase C-terminalEC 1.4.3.6[*Lens culinaris*]
RZ87.R	AA231821	129	1.5e-06 (1)	SP:P53533	Heat shock protein ClpB [*Synechococcus sp.*]
RZ141.R	AA231730	438	1.3e-40 (1)	SP:O04091	Endomembrane protein EMP70 precusor isolog [*Arabidopsis thaliana*]
RZ143.F	AA231892	264	3.6e-22 (1)	SP:Q42977	Glyceraldehyde-3-phosphate dehydrogenase [*Oryza sativa*]
RZ166.F	AA231733	163	2.1e-10 (1)	SP:O04504	Similar to N. tabacum salt-inducible protein [*Arabidopsis thaliana*]
RZ206.F	AA231735	290	6.2e-25 (1)	SP:P46274	Voltage dependent anion channel (VDAC) [*Triticum aestivum*]
RZ206.R	AA231736	111	2.2e-05 (1)	SP:P46274	Voltage dependent anion channel (VDAC) [*Triticum aestivum*]
RZ244.F	AA231738	357	4.8e-32 (1)	SP:P31023	Dihydrolipoamide dehydrogenase [*Pisum sativum*]
RZ244.R	AA231739	227	2.8e-19 (2)	SP:Q41219	Ferric leghemoglobin reductase [*Glycine max*]
RZ261.R	AA231740	171	2.7e-11 (1)	SP:Q42947	Dehydroquinate dehydratase/shikimate dehydrogenase [*Nicotiana tabacum*]
RZ273.R	AA231742	297	1.1e-25 (1)	SP:P31691	ATP/ADP translocator [*Oryza sativa*]
RZ323.R	AA231854	168	2.1e-13 (2)	SP:AJ002414_1	HnRNP-like protein [*Arabidopsis thaliana*]
RZ329.R	AA231747	220	2.7e-16 (1)	SP:Z99708_49	Beta-galactosidase [*Arabidopsis thaliana*]
RZ387.F	AA231749	151	5.3e-14 (2)	SP:Q59083	UDP-glucose 4-epimerase [*Azospirillum brasilense*]
RZ390.F	AA231765	311	3.8e-27 (1)	SP:P49100	Cytochrome b5 [*Oryza sativa*]
RZ390.R	AA231766	174	3.9e-15 (2)	SP:P49100	Cytochrome b5 [*Oryza sativa*]
RZ397.R	AA231769	114	4.9e-06 (1)	SP:O04437	Glutathione-S-transferase [*Triticum aestivum*]
RZ400.F	AA231895	404	1.3e-47 (2)	SP:Q40210	Small GTP-binding protein, RAB5B [*Lotus japonicus*]
RZ403.F	AA231867	215	2.7e-16 (1)	GB:AC000104	Similar to aldo-keto reductase [*Arabidopsis thaliana*]
RZ404.F	AA231874	215	5.6e-17 (1)	SP:O04182	CaMB-channel protein [*Nicotiana tabacum*]
RZ421.R	AA231772	105	0.00051 (1)	SP:U78525	Translation initiation factor [*Homo sapiens*]
RZ455.R	AA231786	477	9.6e-45 (1)	SP:Q39433	Small G protein [*Beta vulgaris*]
RZ474.F	AA231897	276	2.6e-35(2)	SP:Q08267	Tuber-induction gene 3' region[*Solanum tuberosum*]
RZ474.R	AA231787	143	8.7e-09 (1)	SP:Q08267	Tuber-induction gene 3' region[*Solanum tuberosum*]
RZ476.F	AA231773	213	3.4e-16 (1)	SP:O04487	Similar to elongation factor 1-gamma [*Arabidopsis thaliana*]
RZ500.F	AA231788	131	7.5e-07 (1)	SP:Q96290	Sugar transporter [*Arabidopsis thaliana*]
RZ500.R	AA231789	131	7.3e-07 (1)	SP:Q96290	Sugar transporter [*Arabidopsis thaliana*]
RZ508.F	AA231790	445	2.4e-41 (1)	SP:P55309	Catalase [*Oryza sativa*]
RZ508.R	AA231899	149	7.1e-12 (2)	SP:AF035256_1	Catalase [*Glycine max*]
RZ509.ALL	AA231791	192	1.5e-14 (1)	SP:Q39439	Plasma membrane major intrinsic protein 1 [*Beta vulgaris*]
RZ516.R	AA231855	182	5.8e-13 (1)	SP:Q39892	Nucleosome assembly protein 1 [*Glycine max*]
RZ567.F	AA231774	273	4.0e-23 (1)	SP:Q42038	Kinesin light chain [*Arabidopsis thaliana*]
RZ574.R	AA231814	195	8.6e-14 (1)	SP:Z97340_20	Putative cell wall protein [*Arabidopsis thaliana*]
RZ583.F	AA231657	153	2.0e-10 (1)	SP:O00264	putative progesterone binding protein [*Homo sapiens*]
RZ588.R	AA231816	369	2.7e-33 (1)	SP:P23902	Beta-ketoacyl-ACP synthase I [*Hordeum vulgare*]
RZ614.F	AA231660	312	5.6e-30 (2)	SP:P27788	Ferredoxin III [*Zea mays*]
RZ630.F	AA231797	251	3.8e-20 (1)	SP:AF016305_1	ATP sulfurylase [*Zea mays*]
RZ630.R	AA231798	422	5.1e-50 (2)	SP:AF016305_1	ATP sulfurylase [*Zea mays*]
RZ672.F	AA231801	469	6.7e-44 (1)	SP:O04982	Cystathionine gamma-synthase [*Zea mays*]
RZ682.R	AA231802	224	5.9e-18 (1)	SP:P32055	Colanic acid biosynthesis protein [*E. coli*]
RZ698.F	AA231803	282	1.9e-29 (2)	SP:P92991	Phosphate/phosphoenolpyruvate translocator precursor; PPT [*Arabidopsis thaliana*]
RZ698.R	AA231804	369	2.7e-33 (1)	SP:P93642	Phosphate/phosphoenolpyruvate translocator precursor [*Zea mays*]
RZ740.F	AA231807	343	1.5e-30 (1)	SP:Y07766_1	S-adenosylmethionine decarboxylase [*Oryza sativa*]
RZ753.R	AA231808	540	2.0e-51 (1)	SP:P53683	Calcium-dependent protein kinase [*Oryza sativa*]
RZ836.F	AA231662	444	3.0e-41 (1)	SP:Q43605	Alpha-tubulin [*Oryza sativa*]
RZ836.R	AA231851	168	1.5e-14 (2)	SP:Q43605	Alpha-tubulin [*Oryza sativa*]
RZ900.F	AA231824	588	1.7e-56 (1)	SP:Z97335_3	Adenosylhomocysteinase [*Arabidopsis thaliana*]
RZ912.F	AA231663	464	3.4e-42 (1)	SP:P10931	Phytochrome (phy 18) [*Oryza sativa*]
RZ995.F	AA231827	170	3.3e-12 (1)	SP:Q57752	Hypothetical ferripyochelin binding protein [*Methanococcus jannaschii*]
WG110.F	AA231665	384	6.9e-35 (1)	SP:Z97335_12	Selenium-binding protein [*Arabidopsis thaliana*]
WG110.R	AA231666	139	3.2e-08 (2)	SP:Z97335_13	Selenium-binding protein [*Arabidopsis thaliana*]

[a] F and R indicate forward and reverse sequences, respectively
[b] GB, GenBank: SP, SwissProt; PIR, Protein Identification Resource

Third, the data provides the foundation for the design of sequence-based mapping tools which are amenable to rapid, large-scale analyses, for studies of allelic diversity across a range of taxa, and as the basis for understanding the evolution of gene families. Sequence information can be combined with positional information in grasses to distinguish between orthologous and paralogous family members, underscoring the value of having both sequence and mapping data when trying to interpret evolutionary relationships.

Finally, sequence information is the key to probe identity and provides an economical and reliable way of distributing probe information to the scientific community, thereby diminishing the reliance on the physical distribution of probes, which is costly, time-consuming, and prone to errors.

As might have been expected, a large proportion of the proteins of known function identified in Table 2 are indicative of heterotrophic activity, involved in the TVA cycle or the glycolitic pathway, while plastocyanin, RuBisco and proteins from the light-harvesting complex, which would be abundant in cDNA libraries from green plant tissue, are notably absent in all of these libraries derived from etiolated leaf tissue.

Rice as blueprint for comparative mapping in Gramineae

In establishing the anchor probe set, we gave priority to probes detecting single- or low-copy sequences and providing good genome coverage in rice. The majority of the probes selected also detected low-copy sequences in barley, wheat, oat, maize, sorghum, and sugarcane. The rice genome provides a good basis for comparative mapping in monocots because of its small, diploid genome (0.45 pg per haploid cell) (Arumuganathan and Earle, 1991) and its well-developed classical and molecular maps (Kinoshita 1995; Causse et al. 1994; Kurata et al. 1994a; Harushima et al. 1998).

Information is available on over 150 morphological mutants and 3,000 DNA markers. There is an average spacing of 1 DNA marker every 0.5 cM, and an average DNA/cM ratio of 250–300 kb/cM (Van Houten et al. 1996). Efforts to develop a physical map of the rice genome are underway (Kurata et al. 1997; Hong 1997; Zhang et al. 1997), and in combination with YAC (Umehara et al. 1994), BAC (Wang et al. 1995; Yang et al. 1997; Zhang et al. 1996), and cosmid libraries (Song et al. 1995) this physical map will facilitate high-resolution comparative mapping (Chen et al. 1997; Bennetzen et al. 1998) and positional gene cloning in grass species (Kilian et al. 1995; Song et al. 1995; Yoshimura et al. 1998).

The use of anchor probes for comparative mapping is an efficient way of establishing genetic relationships for comparisons among all the species and genera being studied (Ahn and Tanksley 1993; Ahn et al. 1993; Van Deynze et al. 1995a,b,c). Collaboration among research groups involved in the mapping of other organisms will contribute to extending the set of anchor probes. As the development of comparative maps is a dynamic process, we envision the expansion of this set of clones to include: (1) cDNA clones from a wider array of species and libraries, (2) cDNA clones of known gene function which are of agronomic importance, and (3) clones from poorly represented regions of specific genomes that give single-copy hybridization signal across a majority of species tested.

Certain anchor probes hybridize to more than one locus per genome, and some are very similar to members of complex multigene families such as tubulin and GTP-binding proteins. This potentially complicates map alignment between species as heterologous probes may sometimes fail to distinguish sequences which are genuinely orthologous (direct descendents of an ancestral gene) from those which are parologous (descendents of a duplicated ancestral gene) (Hillis 1994). However, the complications can be addressed by comparing multiple maps with a high density of common markers and by sequence analysis of putatively orthologous genes.

All materials and information relating to this work are available to other researchers. The anchor probes described in this report are available over the WWW at "http://kernel.cit.cornell.edu/rice/probe.html". The end sequences for all of the anchor probes reported here can be accessed in GenBank, RiceGenes, or GrainGenes. The data, images of screening filters and linkage maps developed using the anchor probes are available in RiceGenes or GrainGenes at the National Agricultural Library, accessible over the World Wide Web [http://probe.nalusda.gov]. The interactive comparative grass genome display is available in RiceGenes at the above address.

Note in press: The similarities shown in Table 2 represent relationships between anchor probe sequences and public sequence data analyzed on Jan. 23, 1998. The authors recommend that searches be repeated to extract the most recent information. In addition, the choice of gapped vs. ungapped BLAST, and whether or not non-redundant nucleotide or protein databases are queried, will influence the results.

Acknowledgements We wish to thank Steve Tanksley, Bill Wilson, Daniella Braga, Clare Nelson, Dave Matthews, and Ernest Retzel for critical reading of this manuscript; John Bennett and his staff at the International Rice Research Institute, Philippines, for making some unpublished sequence data available to help confirm clone identities; Sandra Harrington for DNA preps on all sequenced probes; Andrew Paterson for DNA for *Sorghum bicolor* used in survey filters, Marci Blinstrub and Mauricio la Rota for data modeling in the RiceGenes database, and Carole Morehouse for help formatting the manuscript. Funding for this project was provided by the USDA Plant Genome National Research Initiative (USDA NRI Grant No. 94-37300-0324; USDA NRI/CGP Grant No. 94-37310-0661; and subcontract numbers 92-G0161-Cornell of USDA NRI Grant No. 92-37300-7550), Copersucar Technology Center, the International Consortium for Sugarcane Biotechnology, Texas Advanced Technology Program, the Texas Agricultural Experiment Station, the Institute for Biosciences and Technology at

Texas A&M University, the Texas Rice Research Foundation, and the USDA-ARS Plant Genome Program.

References

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

Ahn S, Anderson JA, Sorrells ME, Tanksley SD (1993) Homoeologous relationships of rice, wheat and maize chromosomes. Mol Gen Genet 241: 483–490

Altuschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389–3402

Arumunagathan K, Earle ED (1991) Nuclear DNA content of some important plant species. Plant Mol Bio Rep 9: 208–218

Ashikawa I, Umehara Y, Wang ZX, Saji S, Shimokawa T, Sue N, Yoshino K, Kurata N (1994) Assembling rice YAC contigs. Rice Genome 3: 6–7

Bennetzen JL, San Miguel P, Chen M, Tikhonov A, Franki M, Avramova Z (1998) Grass Genomes. Proc Natl Acad Sci 95: 1975–1978

Burr B, Burr FA (1991) Recombinant inbreds for molecular mapping in maize. Trends Genet 7: 55–60

Causse M, Fulton TM, Cho YG, Ahn SN, Chunwongse J, Wu K, Xiao J, Yu Z, Ronald PC, Harrington SB, Second GA, McCouch SR, Tanksley SD (1994) Saturated molecular map of the rice genome based on an interspecific backcross population. Genetics 138: 1251–1274

Chao S, Sharp PJ, Worland AJ, Warham EJ, Koebner RMD, Gale MD (1989) RFLP-based genetic maps of wheat homoeologous group 7 chromosomes. Theor Appl Genet 78: 495–504

Chen M, SanMiguel P, de Oliveira AC, Woo SS, Zhang H, Wing RA, Bennetzen JL (1997) Microcolinearity in *sh*2-homologous regions of the maize, rice, and sorghum genomes. Proc Natl Acad Sci USA 94: 3431–3435

Chitttenden LM, Schertz KF, Lin YR, Wing RA, Paterson AH (1994) A detailed RFLP map of *Sorghum bicolor* × *S. Propinquum* suitable for high-density mapping suggests ancestral duplication of *Sorghum* chromosomes or chromosomal segments. Theor Appl Genet 87: 925–933

Da Silva JAG, Sorrells ME, Burnquist WL, Tanksley SD (1993) RFLP linkage map and genome analysis of *Saccharum spontaneum*. Genome 36: 782–791

Devos KM, Atkinson MD, Chinoy CN, Liu CJ, Gale MD (1992) RFLP-based genetic map of the homoeologous group 3 chromosomes of wheat and rye. Theor Appl Genet 83: 931–939

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

Devos KM, Chao S, Li QY, Simonetti MC, Gale M (1994) Relationship between chromosome 9 of maize and wheat homoeologous group 7 chromosomes. Genetics 138: 1287–1292

Devos KM, Wang ZM, Beales J, Sasaki T, Gale MD (1998) Comparative genetic maps of foxtail millet (*Setaria italica*) and rice (*Oryza sativa*). Theor Appl Genet 96: 63–68

Gardiner JM, Coe EH, Melia-Hancock S, Hoisington DA, Chao S (1993) Development of a core RFLP map using an immortalized F_2 population. Genetics 134: 917–930

Gish W, States DJ (1993) Identification of protein coding regions by database similarity search. Nat Genet 3: 266–72

Graner A, Jahoor A, Schondelmaier J, Siedler H, Pillen K, Fischbeck G, Wenzel G, Herrmann RG (1991) Construction of an RFLP map of barley. Theor Appl Genet 83: 250–256

Graner A, Bauer E, Kellermann A, Kirchner S, Muraya JK, Jahoor A, Wenzel G (1994) Progress of RFLP map construction in winter barley. Barley Genet Newsl 23: 53–61

Grivet L, D'Hont A, Dufour P, Hamon P, Roques D, Glaszmann JC (1994) Comparative mapping of sugar cane with other species within the Andropogoneae tribe. Heredity 73: 500–508

Harrington S, Bligh HFJ, Park W, Jones C, McCouch SR (1997) Linkage mapping of starch branching enzyme III in rice (*Oryza sativa* L.) and prediction of location of orthologous genes in other grasses. Theor Appl Genet 94: 564–568

Havukkala I, Ichimura H, Nagamura Y, Sasaki T (1995) Rice genome analysis by integration of sequencing and mapping data. J Biotechnol 41: 139–148

Heun M, Kennedy AE, Anderson JA, Lapitan NLV, Sorrells ME, Tanksley SD (1991) Construction of a restriction fragment length polymorphism map for barley (*Hordeum vulgare*). Genome 34: 437–447

Hillis DM (1994) Homology in molecular biology. In Hall BK (ed) Homology: the hierarchical basis of comparative biology. Academic Press, New York, pp 339–368

Hong G (1997) A rapid and accurate strategy for rice contig map construction by combination of fingerprinting and hybridization. Pl Mol Biol 35: 129–133

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

Kilian A, Kudrna DA, Kleinhofs A, Yano M, Kurata N, Steffenson B, Sasaki T (1995) Rice-barley synteny and its application to saturation mapping of the barley *Rpg*1 region. Nucleic Acids Res 23: 2729–2733

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugen 12: 172–175

Kurata N, Nagumara Y, Yamamoto K, Harushima Y, Sue N, Wu J, Antonio BA, Shomura A, Shimuzu T, Lin S-Y, Inoue T, Fukuda A, Shimano T, Kuboki Y, Toyama T, Miyamoto Y, Kirihara T, Hayasaka K, Miyao A, Monna L. Zhong HS, Tamura Y, Wang Z-X, Momma T, Umehara Y, Yano M, Sasaki T, Minobe Y (1994a) A 300-kilobase interval genetic map of rice including 883 expressed sequences. Nat Genet 8: 365–372

Kurata N, Moore G, Nagumara Y, Foote T, Yano M, Minobe Y, Gale M (1994b) Conservation of genome structure between rice and wheat. Bio/technology 12: 276–278

Kurata N, Umehara Y, Tanoue H, Sasaki T (1997) Physical mapping of the rice genome with YAC clones. Pl Mol Biol 35: 101–113

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

Lin YR, Schertz KF, Paterson AH (1995) Comparative analysis of QTLs affecting plant height and maturity across the Poaceae, in reference to an interspecific sorghum population. Genetics 141: 391–412

Liu YG, Tsunewaki K (1991) Restriction fragment length polymorphism analysis of wheat. II. Linkage maps of the RFLP sites in common wheat. Jpn J Genet 66: 617–633

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

Melake Berhan A, Hulbert SH, Butler LG, Bennetzen JL (1993) Structure and evolution of the genomes of *Sorghum bicolor* and *Zea mays* Theor Appl Genet 86: 598–604

Naranjo T, Roca P, Goicoechea PG, Giraldez R (1987) Arm homoeology of wheat and rye chromosomes. Genome 29: 873–882

Nelson JC, Van Deynze AE, Autrique E, Sorrells ME, Lu YH, Merlino M, Atkinson M, Leroy P (1995a) Molecular mapping of wheat. Homoeologous group 2. Genome 38: 517–524

Nelson JC, Van Deynze AE, Autrique E, Sorrells ME, Lu YH, Negre S, Bernard M. Leroy P (1995b) Molecular mapping of wheat. Homoeologous group 3. Genome 38 : 525–533

Nelson JC, Sorrells ME, Van Deynze AE, Lu YH, Atkinson M, Bernard M, Leroy P, Faris J, Anderson J (1995c) Molecular mapping of wheat. Major genes and rearrangements in homoeologous groups 4, 5, and 7. Genetics 141 : 721–731

O'Donoughue LS, Wang Z, Röder M, Kneen B, Leggett M, Sorrells ME, Tanksley SD (1992) An RFLP-based map of oat on a cross between two diploid taxa (*Avena atlantica* × *A. hirtula*). Genome 35 : 765–771

O'Donoughue LS, Kianian SF, Rayapati PJ, Penner GA, Sorrells ME, Tanksley SD, Phillips RL, Rines HW, Lee M, Fedak G, Molnar SJ, Hoffman D, Salas CA, Wu B, Autrique E, Van Deynze A (1995) A molecular map of cultivated oat. Genome 38 : 368–380

Paterson AH, Lin YR, Li Z, Schertz KF, Doebley JF, Pinson SRM, Liu SC, Stansel JW, Irvine JE (1995) Convergent domestication of cereal crops by independent mutations at corresponding genetic loci. Science 269 : 1714–1718

Paul E, Blinstrub M, Hart M, McCouch SR (Curators) (1998) RiceGenes database. Database accessible via World Wide Web, http: //probe.nalusda.gov

Rayapati PJ, Gregory JW, Lee M, Wise RP (1995) A linkage map of diploid oat *Avena* based on RFLP loci and a locus conferring resistance to *Puccinia coronata* var avenae. Theor Appl Genet 89 : 831

Sasaki T, Song J, Koga-Ban Y, Matsui E, Fang F, Higo H, Nagasaki H, Hori M, Miya M, Murayama-Kayano E, Takiguchi T, Takasuga A, Niki T, Ishimaru K, Ikeda H, Yamamoto Y, Mukai Y, Ohta I, Miyadera N, Havukkala I, Minobe Y (1994) Toward cataloguing all rice genes: large scale sequencing of randomly chosen rice cDNAs from a callus cDNA library. Plant J 6 : 615–624

Sears ER (1966) Nullisomic-tetrasomic combinations in hexaploid wheat In: Riley R, Lewis KR Chromosome manupulation and plant genetics. Oliver and Boyd, Edinburgh, pp 29–45

Sears ER, Sears LMS (1978) The telocentric chromosomes of common wheat. In: Ramanujam S (ed) Proc 5th Int Wheat Genet Symp. Kapoor Art Press, New Delhi, pp 389–407

Song W-Y, Wang G-L, Chen L, Kim H-S, Wang B, Holsten T, Zhai W-X, Zhu L-H, Fauquet C, Ronald P (1995) The rice disease resistance gene *Xa21* encodes a receptor kinase-like protein. Science 270 : 1804–1806

Thompson JD, Higgins DG, Gibson TJ (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res 22 : 4673–4680

Umehara Y, Inagaki A, Tanoue H, Yasukochi T, Nagamura Y, Saji S, Otsuki Y, Fujimura T, Kurata N, Minobe Y (1994) Construction and characterization of rice YAC libraries for physical mapping. Mol Breed 1 : 79–89

Van Deynze AE, Dubcovsky J, Gill KS, Nelson JC, Sorrells ME, Dvorak J, Gill BS, Lagudah ES, McCouch SR, Appels R (1995a) Molecular-genetic maps for chromosome 1 in Triticeae species and their relation to chromosomes in rice and oats. Genome 38 : 45–59

Van Deynze AE, Nelson JC, O'Donoughue LS, Ahn SN, Siripoonwiwat W, Harrington SE, Yglesias ES, Braga DP, McCouch SR, Sorrells ME (1995b) Comparative mapping in grasses. Oat relationships. Mol Gen Genet 249 : 349–356

Van Deynze AE, Nelson JC, Yglesias ES, Harrington SE, Braga DP, McCouch SR, Sorrells ME (1995c) Comparative mapping in grasses. Wheat relationships. Mol Gen Genet 248 : 744–754

Van Houten W, Kurata N, Umehara Y, Sasaki T, Minobe Y (1996) Generation of a YAC contig encompassing the extra glume gene, eg, in rice. Genome 39 : 1072–1077

Vavilov NI (1940) The new systematics of cultivated plants. In: Huxley J (ed) The new systematics. Clarendon Press, Oxford, pp 549–566

Wang GL, Holsten TE, Song WY, Wang HP, Ronald PC (1995) Construction of a rice bacterial artificial chromosome library and identification of clones linked to the *Xa-21* disease resistance locus. Plant J 5 : 525

Werner JE, Endo TR, Gill BS (1992) Toward a cytogenetically based physical map of the wheat genome. Proc Natl Acad Sci USA 89 : 11307–11311

Whitkus R, Doebley J, Lee M (1992) Comparative genome mapping of sorghum and maize. Genetics 132 : 119–1130

Wilson WA, Harrington SE, Woodman WL, Lee M, Sorrells ME, McCouch SR (1998) Inferring the genome structure of progenitor maize through comparative analysis of rice, maize, and the domesticated panicoids. Genetics (in press)

Xie DX, Devos KM, Moore G, Gale MD (1993) RFLP-based maps of the homoeologous group 5 chromosomes of bread wheat (*Triticum aestivum* L.). Theor Appl Genet 87 : 70–74

Yang D, Parco A, Nandi S, Subudhi P, Zhu Y, Wang G, Huang N (1997) Construction of a bacterial artificial chromosome (BAC) library and identification of overlapping BAC clones with chromosome 4-specific RFLP markers in rice. Theor Appl Genet 95 : 1147–1154

Yoshimura S, Yamanouchi U, Katayose Y, Toki S, Wang ZX, Kono I, Kurata N, Yano M, Iwata N, Sasaki T (1998) Expression of *Xal*, a bacterial blight-resistance gene in rice is induced by bacterial inoculation. Poc Natl Acad Sci USA 95 : 1663–1668

Zhang HB, Wing RA (1997) Physical mapping of the rice genome with BACs. Pl Mol Biol 35 : 115–127