Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 5661–5666, May 1998
Genetics

Screening for overlapping bacterial artificial chromosome clones by PCR analysis with an arbitrary primer

(bacterial artificial chromosome library/contig/physical mapping)

JICHEN XU*, DAICHANG YANG*, JESSICA DOMINGO*, JUNJIAN NI*, AND NING HUANG*†‡

*Plant Breeding, Genetics and Biochemistry, International Rice Research Institute, P.O. Box 933, 1099 Manila, Philippines; and †Applied Phytologics, Inc., 4110 North Freeway Boulevard, Sacramento, CA 95834

Communicated by Diter von Wettstein, Washington State University, Pullman, WA, March 2, 1998 (received for review November 15, 1997)

ABSTRACT **In this article, we used PCR analysis with arbitrary primers (AP-PCR) to screen for overlapping bacterial artificial chromosome (BAC) clones and assembly of contigs. A rice BAC library with three genome equivalents was used to prepare pooled BAC DNA. Twenty-two arbitrary primers were used to survey the pooled BAC DNAs and individual BAC DNAs. Each primer identified 1–10 loci, and the average was 4.4 loci. There were 1–5 overlapping clones in each locus, and the average was 2.5 clones. A total of 245 BAC clones were identified as overlapping by AP-PCR and the identities were confirmed by DNA–DNA hybridization. The 245 BAC clones were then assembled into 80 contigs and 17 single-clone loci. The results indicated that PCR analysis with arbitrary primers is a powerful tool in screening for overlapping BAC clones with high accuracy and efficiency. The use of AP-PCR analysis should speed up the construction of physical maps of the plant and animal genomes, as well as the rice genome.**

Intensive efforts are being made to construct high-resolution physical maps of human, animal, and plant genomes through the generation of ordered overlapping DNA fragments. Because of their capacity for housing large fragments of exogenous DNA, yeast artificial chromosomes (YAC) (1) and bacterial artificial chromosomes (BAC) (2) are used to generate contigs of large genomes. Several approaches are used to identify overlapping clones. DNA–DNA hybridization with mapped DNA markers, combined with chromosomal walking, is widely used to identify overlapping YAC clones. A YAC contig map of the human genome has been constructed by using this technique (3). In plants, physical maps of overlapping YAC clones of chromosomes 2 (4) and 4 (5) of *Arabidopsis* have been constructed by colony hybridization of YAC libraries. With the use of 1,383 DNA markers, the rice physical map of YAC covers about 50% of the rice genome (6). To obtain more complete coverage of the rice genome, a higher-density marker map is being constructed (7). Obviously, the major limitation of this approach to generate a physical map is the requirement of a high-density DNA marker map of the target genomes. Furthermore, repetitive DNA sequences present in eukaryote genomes and chimeric YAC clones create enormous difficulties in chromosome walking (6, 8).

Another approach to assemble overlapping clones is based on DNA fingerprinting of random clones (9, 10). The restriction fragments of each DNA clone show typical banding patterns when separated in high-resolution gels and/or probed with DNA probes. Overlapping clones are identified by shared restriction fragments. By use of DNA fingerprinting of cosmid clones, a high-quality physical map of the nematode *Caenorhabditis elegans* genome was constructed (11). Though the DNA fingerprinting approach is effective in mapping a small genome, its application to a large genome such as the human genome can be difficult (3).

Sequence tag site (STS) content mapping is used for the identification of YAC clones (12). Early application of the method resulted in the assembly of YAC contigs of human chromosome 21q (13) and regions of the human Y chromosome (14). Recently, Hudson *et al*. (15) used more than 15,000 STS to generate a physical map covering 94% of the human genome. The method greatly simplified the procedure in identifying YAC clones. Furthermore, the data produced by STS can be easily stored in a database of the STS sequences (16). The major limitation of this approach is the requirement of a large number of sequenced STS that are based on DNA sequence and extensive specific primer synthesis.

Herein we describe an approach to identify overlapping clones. The approach takes advantage of the principles of STS content mapping but avoids the need of sequence specific primers. We used individual arbitrary primers to perform PCR (AP-PCR) in a way similar to STS content mapping. The primers used randomly but specifically amplify a few loci. For each locus, overlapping clones are identified as in STS content mapping. For each primer, several loci can be mapped simultaneously thus speeding the construction of a physical map of target genomes. Our approach will be very useful to generate physical maps of animal and plant genomes with high efficiency.

MATERIALS AND METHODS

BAC Library. A BAC library of a rice variety, IR64, was constructed at the Genome Mapping Laboratory at International Rice Research Institute (17). A total of 18,432 clones corresponding to 3.28 rice genome equivalents were maintained and grown on 48 microtiter plates. Each plate is an array of 384 wells with 16 rows and 24 columns. The insert sizes of the BAC clones range from 37 to 364 kb, with an average of 107 kb.

BAC DNA Pools. Two levels of BAC DNA pools (primary and secondary pools) were prepared for analysis. The primary pools were based on the entire BAC library and were prepared with a three-dimensional pooling scheme. The BAC clones were grown to saturation before pooling. The first dimension was microtiter plate pool. Bacteria of each plate were pooled and BAC DNA was isolated to form the plate pools. Each of the 48 plate pools thus contained DNA from 384 BAC clones. The second dimension consisted of row pools. Bacteria from the same row of the 48 plates were pooled to produce row pools. Each of the 16 row pools contained DNA from 1,152 BAC clones (24 × 48). The third dimension was column pools that consisted of DNA isolated from bacteria of the same column over the 48 plates of the BAC library.

Abbreviations: AP-PCR, PCR with arbitrary primer; BAC, bacterial artificial chromosome; YAC, yeast artificial chromosome; STS, sequence tag site.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AF048983 and AF048984).

‡To whom reprint requests should be addressed. e-mail: Huang@APINC.COM.

Each of the 24 column pools contained DNA from 768 BAC clones (16 × 48). A total of 88 primary DNA pools were thus prepared.

The secondary pools were based on the individual microtiter plate of the BAC library. Two dimension (row and column) poolings were made for each of the 48 plates. For each plate, bacterial cells from each row (24 clones) or each column (16 clones) were separately pooled for DNA isolation. Each plate provided in this way 40 secondary DNA pools (16 row pools and 24 column pools). For the entire BAC library of the 48 microtiter plates, a total of 1,920 secondary pools were obtained.

Isolation of BAC DNAs. The BAC clones were picked from the library and inoculated into 2 ml of LB medium containing chloramphenicol (12.5 μg/ml) and incubated at 37°C overnight. Minipreparation of BAC DNAs was by the method of Yang *et al.* (17).

Arbitrary Primers and AP-PCR Analysis. Twenty-two arbitrary primers (Table 1) were obtained from Operon Technology (Alameda, CA). AP-PCR was conducted in a Perkin–Elmer model GeneAmp PCR System 9600. A reaction mixture of 20 μl contained 50 mM KCl, 10 mM Tris·HCl (pH 8.3), 2.5 mM $MgCl_2$, 0.01% gelatin, all four dNTPs (each at 0.1 mM), 20 ng of each primer, 30 ng of isolated BAC DNAs, and 1 unit of *Taq* polymerase. Amplification started with 2 min at 94°C, followed by 40 cycles of 1 min at 94°C (DNA denaturation), 1 min at 36°C (primer annealing), and 2 min at 72°C (primer extension). The reaction was terminated after a 2-min final primer extension at 72°C. PCR products were separated in an 1.5% agarose gel containing ethidium bromide (0.5 μg/ml). Separated PCR products were visualized under UV light and photographed to examine the banding patterns.

DNA Fingerprinting of BAC Clones. DNA of putative overlapping BAC clones was digested with *Hin*dIII and the fragments were separated in an 0.8% agarose gel. The separated DNA fragments were transferred to a Hybond-N^+ nylon membrane. For each group of putative overlapping BAC clones, the one producing the largest number of *Hin*dIII fragments was labeled as probe. Hybridization was carried out as described by Yang *et al.* (17). After autoradiography, banding patterns of putative overlapping clones were compared pairwise.

Table 1. Arbitrary primers used to screen rice BAC library

Primer	Sequence (5′ to 3′)	No. of contigs	No. of clones	Clones/ contig
A10	GTGATCGCAG	5	15	3
AA11	ACCCGACCTG	3	8	2.67
AA13	GAGCGTCGCT	1	3	3
AA7	CTACGCTCAC	5	13	2.6
AG16	CCTGCGACAG	2	8	4
AJ13	CAGCCGTTCC	6	15	2.5
AJ5	CAGCGTTGCC	4	10	2.5
AK10	CAAGCGTCAC	2	4	2
AK16	CTGCGTGCTC	2	10	5
AL11	GTCACGTCCT	5	10	2
AL12	CCCAGGCTAC	4	7	1.75
AL8	GTCGCCCTCA	10	21	2.1
B19	ACCCCCGAAG	2	6	3
C14	TGCGTGCTTG	8	19	2.38
D6	ACCTGAACGG	2	3	1.5
E01	CCCAAGGTCC	9	19	2.11
G5	CTGAGACGGA	8	13	1.63
H14	ACCAGGTTGG	6	19	3.17
M2	ACAACGCCTC	4	10	2.5
N1	CTCACGTTGG	4	15	3.75
O5	CCCAGTCACT	2	6	3
Q7	CCCCGATGGT	3	11	4
Total	—	97	245	
Mean	—	4.4	11.1	2.5

DNA Cloning and Sequencing. A common 3.3-kb fragment from BAC clones, 6N9, 13M5, and 15I2, were subcloned into pBluescript KS+. DNA sequences of both ends of cloned fragments were obtained by using the cycle sequencing system (GIBCO/BRL/Life Technologies). The primers used for sequencing were the M13 forward primer (5′-GTAAAACGACGGCCAGT-3′) and the M13 reverse primer (5′-AAACAGCTATGACCATG-3′).

RESULTS

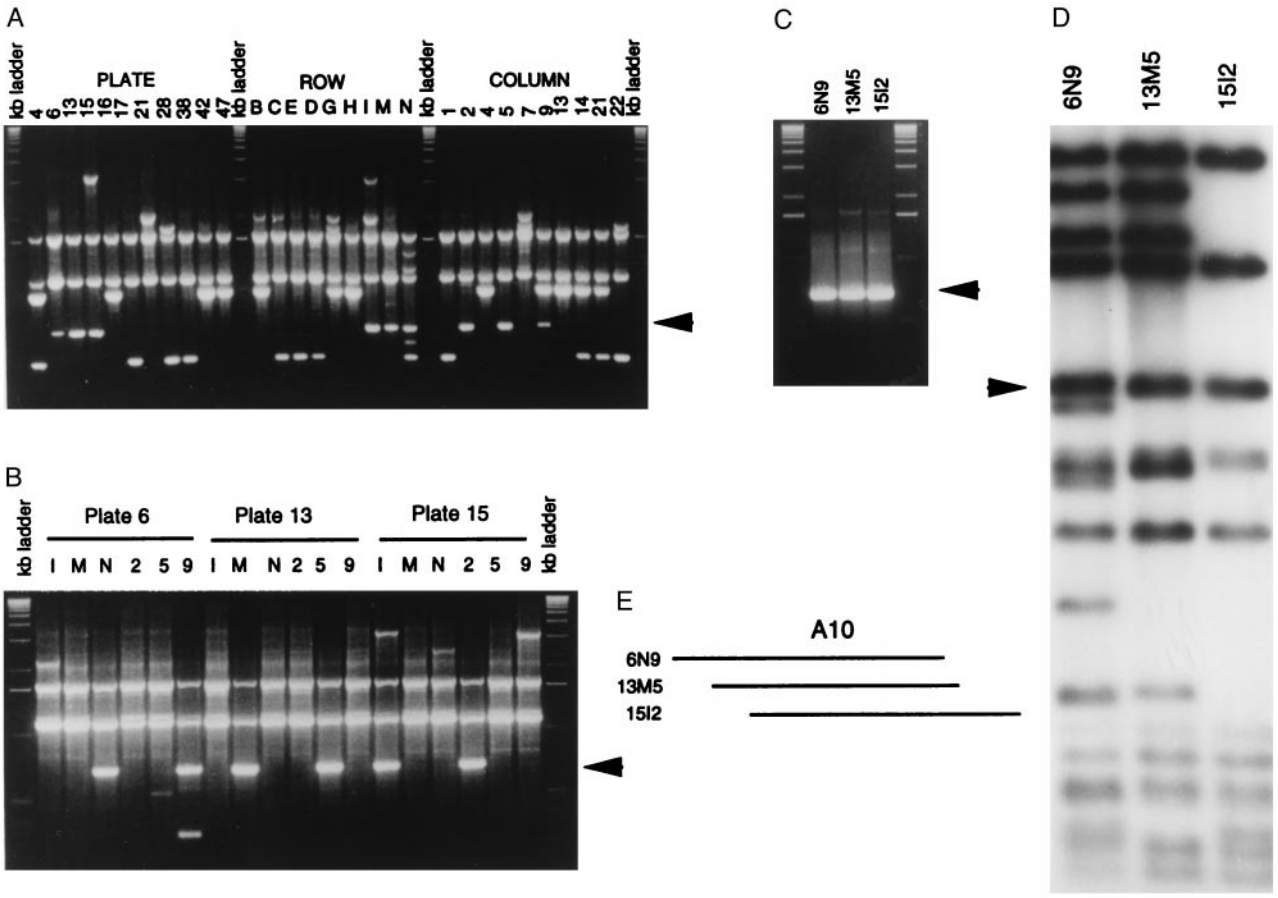
Identifying Overlapping BAC Clones via AP-PCR. To identify overlapping BAC clones via AP-PCR, the 88 primary DNA pools were first surveyed with arbitrary primer A10 (Fig. 1*A*). A10 was capable of amplifying multiple discrete PCR bands (DNA segments) from total genomic DNA isolated from IR64 (data not shown). Each band is assumed to represent a locus in the rice genome. Because the BAC library is about three genome equivalents, it is expected that on the average, each band is present three times in the BAC library or that three BAC clones contain that locus. If the three BAC clones were distributed over different plates, rows, and columns of the library, the survey of primary BAC pools would produce a DNA banding pattern as shown in Fig. 1*A*. For example, the 650-bp fragment is present in three plates (plates 6, 13, 15), in three rows (rows I, M, and N), and three columns (columns 2, 5, and 9). Similarly, the 500-bp band amplified by A10 distributed over four plates (plates 4, 17, 28, and 38), four rows (rows C, D, E, and N), and four columns (columns 1, 14, 21, and 22).

From the banding pattern, we can infer that BAC clones in plates 6, 13, and 15; in rows I, M, and N; and columns 2, 5, and 9 would carry the 650-bp fragment and were overlapping. To determine which three BAC clones containing the 650-bp fragment, relevant secondary BAC pools generated from each plate were amplified with A10 (Fig. 1*B*). From plate 6, BAC clones in row N and column 9 contained the 650-bp fragment; thus, 6N9 is the target clone. Similarly, 13M5 from plate 13 and 15I2 from plate 15 were putative overlapping clones, containing the 650-bp fragment. To further confirm that the three clones contain the DNA fragment, DNA from these three clones was isolated and amplified with A10 (Fig. 1*C*). All three BAC clones produced the expected 650-bp fragment, confirming that the three clones are overlapping BAC clones.

To examine the overlap of the three BAC clones 6N9, 13M5, and 15I2, these clones were fingerprinted by using *Hin*dIII digestion and probed with clone 6N9 (Fig. 1*D*). Many DNA fragments from the three clones were shared both in size and sequence homology. Thus the three BAC clones 6N9, 13M5, and 15I2 were determined to be overlapping and the degree of overlap is shown in Fig. 1*E* according to the results from DNA fingerprinting of the clones.

To further confirm that the three BAC clones 6N9, 13M5, and 15I2 were derived from the same locus in the rice genome and, thus, were indeed overlapping, we obtained partial DNA sequences from a 3.3-kb fragment shared by the three clones (Fig. 1*F*). If the three clones are truly overlapping, then the sequences from the fragment of each clone should be identical. Comparison of a total of the 367-bp DNA sequences obtained from both ends of the cloned fragments showed that they were identical, providing definitive evidence that BAC clones 6N9, 13M5, and 15I2 are overlapping, and thus AP-PCR can be used to identify overlapping BAC clones.

Identification of Overlapping BAC Clones with 22 Arbitrary Primers. When surveying the primary BAC DNA pools with 22 arbitrary primers, 5 to 15 bands (loci) were observed for each primer (see Fig. 1*A*). The size of the bands ranged from 100 bp to 4,000 bp, with the majority of the bands between 500 bp and 2,500 bp. The number of DNA pools generating a specific band (locus) varied from one to many. All 22 primers produced 1 or 2 bands across all 88 primary DNA pools and, sometimes, in secondary DNA pools as well. Presumably, the variation is due to



F

```
                10         20         30         40         50
15I2-F  AAGCTTTCTC AAGTTTGGCT TGGCCGTCTT AACATAGCCC ACACACCCAA
13M5-F  AAGCTTTCTC AAGTTTGGCT TGGCCGTCTT AACATAGCCC ACACACCCAA
6N9-F   AAGCTTTCTC AAGTTTGGCT TGGCCGTCTT AACATAGCCC ACACACCCAA
                60         70         80         90        100
15I2-F  ACACACGCAG GTGTTCGACA GTGGGTCGGC GACCATGCCA CGTTCATACG
13M5-F  ACACACGCAG GTGTTCGACA GTGGGTCGGC GACCATGCCA CGTTCATACG
6N9-F   ACACACGCAG GTGTTCGACA GTGGGTCGGC GACCATGCCA CGTTCATACG
               110        120        130        140        150
15I2-F  CGTCACCCAT CAAGACTCTT CGTTAGGTGA TCTGTTCAGG ACGTAGACGC
13M5-F  CGTCACCCAT CAAGACTCTT CGTTAGGTGA TCTGTTCAGG ACGTAGACGC
6N9-F   CGTCACCCAT CAAGACTCTT CGTTAGGTGA TCTGTTCAGG ACGTAGACGC
               160        170        180
15I2-F  TGCACACACT TCCAGAATAT GCGCATGTTG CGT
13M5-F  TGCACACACT TCCAGAATAT GCGCATGTTG CGT
6N9-F   TGCACACACT TCCAGAATAT GCGCATGTTG CGT
```

```
                10         20         30         40         50
6N9-R   ATGAGCTTGA TTTCACGAGC ATGAAACGTA ACACATTCCA ATGGAGGCTA
15I2-R  ATGAGCTTGA TTTCACGAGC ATGAAACGTA ACACATTCCA ATGGAGGCTA
13M5-R  ATGAGCTTGA TTTCACGAGC ATGAAACGTA ACACATTCCA ATGGAGGCTA
                60         70         80         90        100
6N9-R   AGTAAAACTA TCTTATGAAA TATGATACTT TAATGCATCT GCTCGCGGAG
15I2-R  AGTAAAACTA TCTTATGAAA TATGATACTT TAATGCATCT GCTCGCGGAG
13M5-R  AGTAAAACTA TCTTATGAAA TATGATACTT TAATGCATCT GCTCGCGGAG
               110        120        130        140        150
6N9-R   CTCATGATGC AATATGCATC TGACGTCAGA TGTCTCGGAC GACGCGGAGC
15I2-R  CTCATGATGC AATATGCATC TGACGTCAGA TGTCTCGGAC GACGCGGAGC
13M5-R  CTCATGATGC AATATGCATC TGACGTCAGA TGTCTCGGAC GACGCGGAGC
               160        170        180
6N9-R   ATGTGAGTGG CACACCGATG ACTTTACAGA TCAC
15I2-R  ATGTGAGTGG CACACCGATG ACTTTACAGA TCAC
13M5-R  ATGTGAGTGG CACACCGATG ACTTTACAGA TCAC
```

FIG. 1. Screening for overlapping BAC clones by PCR with arbitrary primer A10. (*A*) AP-PCR products amplified from primary BAC DNA pools with primer A10. Only a portion of the pools from the three dimensions (plate, row, and column) are shown. The arrowhead indicates the 650-bp band (locus) analyzed in *B*. The kb ladder was used as molecular size markers (GIBCO/Life Technology). (*B*) Screening of selected secondary BAC DNA pools with primer A10. The arrowhead indicates the same 650-bp band shown in *A*. (*C*) AP-PCR on the three putative overlapping clones with primer A10. The arrowhead indicates the same 650-bp band. (*D*) DNA fingerprinting of the three overlapping BAC clones with 6N9 as probe. The arrowhead indicates the 3.3-kb band that was cloned and partially sequenced from both ends of cloned fragments. (*E*) Contig of the three overlapping clones based on DNA fingerprints in *D*. (*F*) DNA sequences from both ends of the 3.3-kb fragment of the three overlapping BAC clones.

the limited number of BAC clones, which in turn limits the number of the DNA pools, carrying the same DNA fragment. The variation also could be due to the presence of repetitive DNA sequences dispersed in the rice genome or bacterial DNA contamination during BAC DNA isolation. Both can be amplified by arbitrary primers if there is sufficient sequence homology between primer and the repetitive DNA sequence or bacterial DNA contamination. Because the BAC library covers three genome equivalents, the chance of a single-copy DNA fragment in the rice genome being cloned in more than six BAC clones is less than 1%, we arbitrarily decided that only specific bands generated from six or fewer BAC DNA pools would be analyzed with secondary DNA pools. By this criterion, 1–10 bands (loci), with an average of 4.4 bands (loci), were scored with individual primers (Table 1). Primer AL8 produced a maximum number of 10 loci that could be scored (Table 1).

Generally, the same number of lanes in plate pools showing a specific band were observed in row pools and in column pools as is the case with the 650-bp band in Fig. 1*A*. Sometimes, the number of bands in plate pools was different from that in row or column pools. The 1,000-bp band in Fig. 1*A* is an example. There were four bands in plate pools, three bands in row pools, and five bands in column pools. This could result from two clones containing the same fragment being in the same plate, row, or column. Alternatively, this might be due to nonspecific amplification of BAC DNA with the arbitrary primer. These two possibilities can be resolved by AP-PCR with secondary BAC DNA pools.

Putative overlapping clones identified by screening the primary BAC DNA pools with the arbitrary primers were then grouped

Table 2. Overlapping BAC clones identified by 22 arbitrary primers

Primer	Fragment size, bp	Overlapping BAC clones					Primers	Fragment size, bp	Overlapping clones				
A10	500	04C01	17D21	28N14	38E22		B19	1,200	31B11	28J15	12L22		
	650	06N09	I5I02	13M05				3,000	07I13	21H16	30K06		
	1,000	04B21	16L04	42H13	47H14		C14	200	11J11	04J11	35B14		
	1,700	28G15						450	45N09				
	2,000	46L07	21G24	48C20				550	23G09	36106	37C07		
AA11	600	11C22	09G13					800	22G19	26J24	26F24		
	1,200	18O05	44G06	45K22	29B12			900	04N02	20O22			
	1,600	12K15	19A05					950	15H02	24H08			
AA13	1,700	15O11	45N15	48J20				1,600	10B18	22E06			
AA7	550	13L10	39H22					2,900	33O05	47L09	31N11		
	650	09J16	40J21				D6	1,200	26M02				
	800	34D02	38N03	29M19	24N10			2,200	04P01	46K09			
	1,000	25M15	36N23				E1	500	06L11	12B20			
	1,800	46K04	25E13	25B05				600	26O1	43C21			
AG16	800	23106	09B24	08G13				650	27C07				
	1,500	06E11	29K05	45D14	45D15	06J23		750	16N09	17C11	31NI0		
AJ13	600	18D13	13N19					800	19J08	20M19	43K17		
	1,000	25P07	07P19	01P12				1,100	26H04				
	1,050	15J05	07N19					1,200	31P10				
	1,100	15N17	23O22	24K03	29H22			1,600	07118	48D09			
	2,000	18O23	23M21					1,900	04F06	44F17	29P05	20P05	
	2,200	47K05	12C21				G5	700	25L12	14P05			
A15	1,200	11J21	43N01					1,000	44P24				
	1,500	25J20	43N10					1,200	17D02	40F22			
	2,000	24P05	37P14	41B8	42M8			1,300	02K19	25A11			
	600	03M23	30F11					1,800	48B18				
AK10	200	17E05	17O06					1,900	04H10	37J09			
	1,100	44L20	17J18					1,950	37J10				
AK16	200	30F09	09F05	10O18	17C12	24I08		2,200	08B03	09P21			
	900	08B13	08K10	21B16	39K17	17L23	H14	900	39C05	06F24	20M03	37B06	30K07
AL11	500	I0F18						1,000	05L06	46B17			
	600	18112						1,100	25M12	33D18	37M18	38B21	
	650	28E10	20A03	40N01				1,200	11E04	27C7			
	1,500	04D23	16H01	44K03	48C02			1,700	14N17	21J13			
	2,000	43C21						1,800	16N23	33L06	27A06	07H19	
AL12	500	11H08	26M05				M2	700	09F21				
	650	20KI1						1,100	12A03	24P05	42M08	45D21	34G07
	800	41J06	29J10					1,600	08L17	08N16			
	1,100	20L12	29O02					2,000	23F02	23F22			
AL8	300	27K14	20J07				N1	800	33H19	36G14	14M14	23B01	19E12
	450	13L19	47D08					1000	09H16	32115			
	600	46H23						1,300	25G0S	02N06	14L05	02N24	
	750	01E08	42J12					3,000	10L11	23P16	27N24	47A05	
	750	11P21					O5	1,100	04E18	42A15	46C18	38N17	
	900	02E13						1400	4K21	20F2			
	1,000	02K04	36L06	15E13			Q7	700	33K20	09D06	11P15		
	1,100	09C20	10K04	16G23				800	07O03	09A06	09D04	28P16	46I01
	1,300	08H05	08I05	13M21				2,900	01L08	01M07	40109		
	1,400	27L14	17E04	15B13									

according to specifically amplified bands. The same primer was used to survey selected secondary BAC DNA pools. Generally, much clearer banding patterns were observed in AP-PCR with secondary DNA pools presumably because of more abundant target DNA fragments (only 16 or 24 BAC clones were present in a secondary row or column pool). In most of the cases, only 6 of the 40 secondary DNA pools of each plate needed to be screened (Fig. 1 *A* and *B*). The survey of the secondary BAC DNA pools identified specific overlapping BAC clones from the putative set of BAC clones identified by surveying primary DNA pools. Thus, a total of 261 BAC clones for 97 loci were identified by AP-PCR analysis with 22 arbitrary primers.

The putative overlapping BAC clones were further tested by AP-PCR with BAC DNA isolated from individual BAC clones. The AP-PCR results were backed up by DNA fingerprinting needed to generate contig maps. Of the 261 putative overlapping BAC clones identified with the secondary BAC DNA pools, 245 (94%) BAC clones proved to be overlapped. They were in 97 loci of the rice genome (Table 1). Seventeen of these loci were represented by only a single clone. The other 228 clones were grouped into 80 overlapping groups. On the average, each arbitrary primer identified 11.1 BAC clones that formed 4.4 contigs, with 2.5 BAC clones in each contig.

Linking Overlapping BAC Clones Identified by Two Primers. To link overlapping BAC clones identified by two arbitrary primers, there should be at least a common BAC clone present in the two contigs. Examination of the results in Table 2 showed three such cases (Table 3). Arbitrary primers AL11 and E1 identified BAC clone 43C21. Similarly, overlapping BAC clones identified by arbitrary primers E1 and H14 can be linked through

Table 3. Overlapping BAC clones identified by two arbitrary primers

Primer	Fragment size, bp	BAC clones identified
AL11	2,000	43C21
E1	600	43C21, 26O1
AJ5	2,000	24P5, 42M8, 37P14, 41B8
M2	1,100	24P5, 42M8, 12A3, 45D21, 34G7
E1	650	27C7
H14	1,200	27C7, 11E4

common clone 27C7. The most significant case would be the overlapping clones identified by arbitrary primers AJ5 and M2. In the analysis of the positive overlapping clones, two clones (24P5 and 42M8) were found repeatedly present in two groups. The first group contains four clones (24P5, 42M8, 37P14, and 41B8) that have the same fragment of 2,000 bp amplified by primer AJ5. The second group contains five clones (24P5, 42P8, 12A3, 45D21, and 34G7) that have the same fragment of 1,100 bp amplified by primer M2 (Fig. 2*A*). Southern blot hybridization of these clones with 24P5 as probe proved them to be overlapping (Fig. 2*B*), and the overlaps could be schematically diagrammed as shown in Fig. 2*C*. The common clones identified by two different primers linked the two contigs together. It is believed that an increase in the number of arbitrary primers used will increase the frequency of linkages to generate larger contigs.

DISCUSSION

Generating high-resolution physical maps is the central effort of eukaryote genome research. With sufficient resources, physical maps of the human genome have been established (15) via STS content mapping. However, such resources are hardly available for other eukaryote genomes. Thus, efficient approaches are needed to speed up genome mapping. In this report, we present an approach to generate physical maps of eukaryote genomes. This approach used arbitrary primers rather than sequence-specific STS to identify overlapping BAC clones.

There are several distinct advantages in our approach. First, our approach is rapid. AP-PCR is performed basically as STS content mapping. Because multiple loci can be mapped simultaneously, our approach should be faster than STS content mapping. In this report, we used only 22 primers and were able to map 97 loci (Table 1) instead of the 22 loci that could have been mapped by STS content mapping. Because our library is three genome equivalents and the average insert size is 107 kb (17), the 245 BAC clones identified in the 97 loci would cover about 10 million base pairs or about 2.3% of the rice genome. Thus, to rapidly generate a physical map covering most of the rice genome, we need to screen the BAC library with about 1,000 primers that will generate about 4,400 loci with an average size of 100 kb. The total coverage will be 440 million base pairs, which is the estimated genome size of rice (18).

Second, our approach is convenient and low cost. Our approach requires no information of DNA sequences for primer design, therefore, reducing the cost in obtaining the DNA sequence. There are large numbers of arbitrary primers available. Furthermore, arbitrary primers can be used in pairwise combinations, which will increase the available number of primers. The large scale use of arbitrary primers in gene mapping has demonstrated that 95% of arbitrary primers work well with DNA amplification (for two examples, see refs. 19 and 20). Because of the nature of screening for multiple loci, a considerable reduced number of PCRs are needed to identify an individual locus. In our approach, the analysis with each primer involved three steps: (*i*) identify putative overlapping clones from primary DNA pools (88 reactions), (*ii*) further determine the overlapping clones with secondary DNA pools (72 reactions, if four loci each primer and three overlapping clones per locus are assumed), and (*iii*) indi-

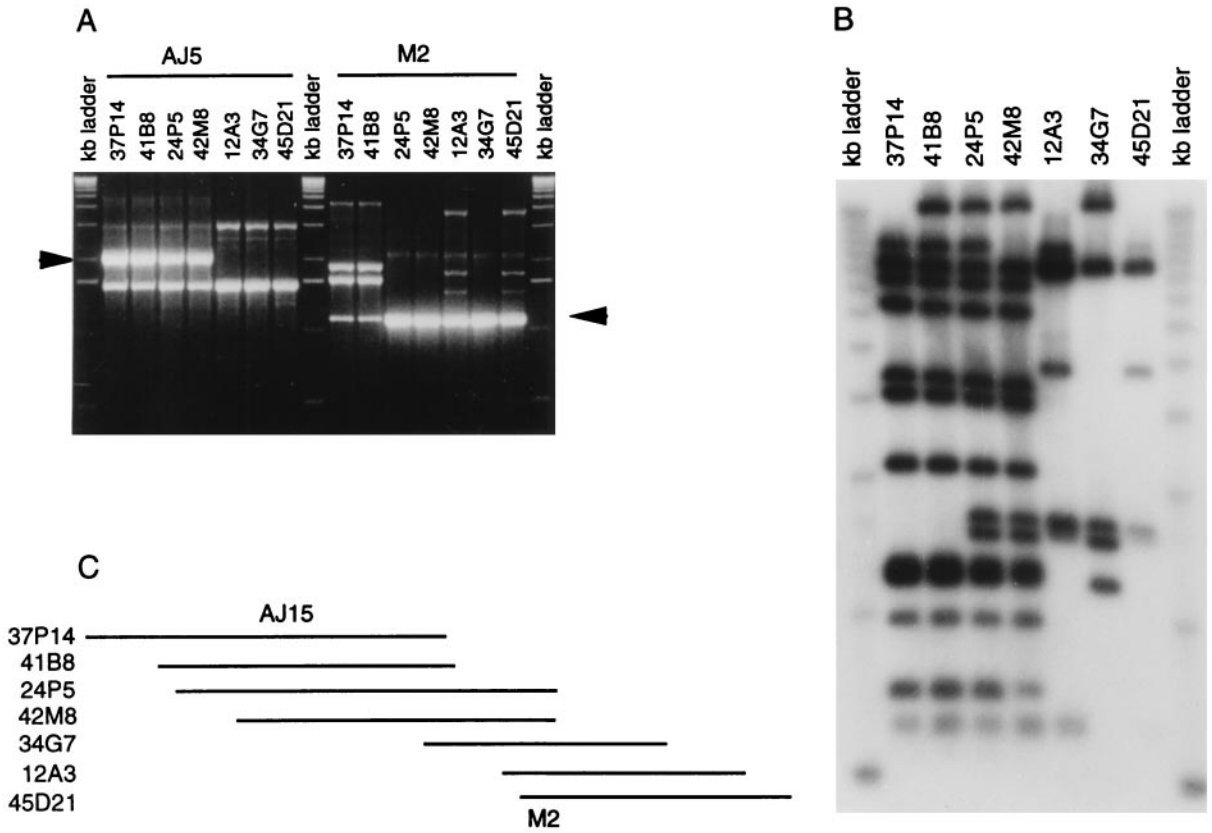


FIG. 2. Screening for overlapping clones by primers AJ5 and M2. (*A*) Banding patterns of individual clones amplified with primer AJ5 and M2. The kb ladder was used as molecular size markers. The arrowheads indicate the specific fragments amplified by each primer. (*B*) Hybridization patterns of individual clones probed with clone 24P5. (*C*) The seven BAC clones were assembled into a contig based on the DNA fingerprints in *B*.

vidual BAC clone confirmation (12 reactions). This gives only 43 PCRs per locus, which is about half the number needed with STS content mapping (72 PCRs) (3).

Third, our approach has broad coverage. Because of the large numbers of random primers used, the primers should amplify DNA loci covering the entire genome. Molecular mapping with DNA fragments amplified with arbitrary primers has shown that the random DNA fragments are distributed over the entire *Arabidopsis* (19) and rice genomes (P. Subudhi and N.H., unpublished data). STS primers constitute additional sources of arbitrary primers and have been used in AP-PCR analyses (21). They have the advantage of amplifying several loci in addition to the specific locus from which the sequence originates. Because STS primers also have been shown to cover entire genomes, broad coverage of animal or plant genomes by large sets of AP-PCR is assured.

Four, the primers can be used in all genomes, both animal and plants. In STS content mapping, the STS primers generated from one species can generally not be used in other species. Alu PCR primers can only be used for the human genome research. The arbitrary primers, on the other hand, are species-independent.

Most of the initial screening with AP-PCR results in the establishment of individual contigs for a given locus. With an increase of the number of primers used, some clones will be identified by two primers. These common clones serve as points to link small contigs together to generate larger contigs. With the use of only 22 primers, we have already seen three such cases (Table 3). So we believe that our approach would eventually, like STS content mapping, permit generation of contiguous physical maps of ordered overlapping DNA fragments.

AP-PCR has been successfully used in germ-plasm characterization (21, 22), generation of a genetic map (19), and genetic mapping of important genes (for an example, see ref. 23). Due to the requirement of a low annealing temperature in AP-PCR, AP-PCR is sensitive to many factors such as the concentration of template DNA, Mg^{2+}, and nucleotides. Among the many specific and reproducible bands, nonspecific (nonreproducible) bands were also observed in AP-PCR analysis. Such random bands give weak signals and can be seen in Fig. 1*A*. In some rare cases, the molecular weight of random bands can be the same as that of targeted bands. However, such random bands impose no problems in the identification of overlapping BAC clones because all clones identified as overlapped must go through three steps of AP-PCR tests that are backed up by DNA-DNA hybridization and fingerprinting. In the present study involving 97 loci and 22 primers, all overlapping clones identified in AP-PCR of individual BAC clones were confirmed by DNA hybridization and fingerprinting.

If a DNA marker map is available, linkage between the physical map and genetic map can be easily established by probing the BAC library with the DNA marker. Alternatively, contigs identified by AP-PCR analysis can be anchored to chromosomes by the segregation of AP-PCR bands in a mapping population mapped with restriction fragment length polymorphism and STS markers (24). Once the linkage between genetic and physical maps is established, the genes mapped in the genetic map can be assigned to the physical map. A segment of ordered DNA fragments becomes immediately available and cloning genes by phenotype will be easy and fast.

Note. After submission of this article, a note was published on the identification of cosmid clones linked to avirulence genes of the fungus *Magnaporthe grisea* by random amplified polymorphic DNA-based screening of a genomic library (25).

We thank Prasanta Subudhi for technical assistance and the Rockefeller Foundation for financial support.

1. Burke, D. T., Carle, G. F. & Olson, M. (1987) *Science* **236,** 806–812.
2. Shizuya, H., Birren, B., Kim, U., Mancino, V., Slepak, T., Tachiiri, Y. & Simon, M. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 8794–8797.
3. Chumakov, I. M., Rigault, P., Le Gall, I., Bellanne-Chantelot, C., Billault, A., Guillou, S., Soularue, P., Guasconi, G., Poullier, E., Gros, I., *et al.* (1995) *Nature (London)* **377,** 175–297.
4. Zachgo, E. A., Wang, M. L., Dewdney, J., Bouchez, D., Camilleri, C., Belmonte, S., Huang, L., Dolan, M. & Goodman, H. M. (1996) *Genome Res.* **6,** 19–25.
5. Schmidt, R., West, J., Love, K., Lenehan, Z., Lister, C., Thompson, H., Bouchez, D. & Dean, C. (1995) *Science* **270,** 480–483.
6. Kurata, N., Umehara, Y., Tanoue, H. & Sasaki, T. (1997) *Plant Mol. Biol.* **35,** 101–113.
7. Harushima, Y., Yano, M., Shomura, P., Sato, M., Shimano, T., Kuboki, Y., Yamamoto, T., Lin, S. Y., Antonio, B. A., Parco, A. *et al.* (1998) *Genetics* **148,** 479–494.
8. Tanksley, S. D.; Ganal, M. W. & Martin, G. B. (1995) *Trends Genet.* **11,** 63–68.
9. Coulson, A., Sulston, J., Brenner. S. & Karn, J. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 7821–7825.
10. Olson, M. V., Dutchik, J. E., Graham, M. Y., Brodeur, G. M., Helms, C., Frank, M., MacCollin, M., Scheinman, R. & Frank, T. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 7826–7830.
11. Riles, L., Dutchik, J. E., Baktha, A., McCauley, B. K., Thayer, E. C., Leckie, M. P., Braden, V. V., Depke, J. E. & Olson, M. V. (1993) *Genetics* **134,** 81–150.
12. Green, C. D. & Olson, M. V. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 1213–1217.
13. Chumakov, I., Rigault, P., Guillou, S., Ougen, P., Billault, A., Guasconi, G., Gervy, P., Le Gall, I., Soularue, P., Grinas, L., *et al.* (1992) *Nature (London)* **359,** 380–387.
14. Foote, S., Vollrath, D., Hilton, A. & Page, D. C. (1992) *Science* **258,** 60–66.
15. Hudson, T. J., Stein, L. D., Gerety, S. S., Ma, J., Castle, A. B., Silva, J., Slonim, D. K., Baptista, R., Kruglyak, L., Xu, S. H., *et al.* (1995) *Science* **270,** 1945–1954.
16. Olson, M., Hood, L., Cantor, C. & Botstein, D., (1989) *Science* **245,** 1434–1435.
17. Yang, D. C., Parco, A., Nandi, S., Subudhi, P., Zhu, Y. G., Wang, G. L. & Huang, N. (1997) *Theor. Appl. Genet.* **95**. 1147–1154.
18. Arumuganathan, K. & Earle, D. E. (1991) *Plant Mol. Biol. Reporter* **9,** 208–218.
19. Reiter, R. S., Williams, J. G. K., Feldmann, K. A., Rafalski, J. A., Tingey, S. V. & Scolnik, P. A. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 1477–1481.
20. Subudhi, P. K., Borkakati, R. P., Virmani, S. S. & Huang, N. (1997) *Genome* **40,** 188–194.
21. Welsh, J. & McClelland, M. (1990) *Nucleic Acids Res.* **18,** 7213–7218.
22. Williams, J. G., Kubelik, A. R., Livak, K. J., Rafalski, J. A. & Tingey, S. V. (1990) *Nucleic Acids Res.* **18,** 6531–6535.
23. Zhang, G., Angeles, E. R., Abenes, M. L. P., Khush, G. S. & Huang, N. (1996) *Theor. Appl. Genet.* **93,** 65–70.
24. Huang, N., Parco, A., Mew, T., Magpantay, G., McCouch, S., Guiderdoni, E., Xu, J., Subudhi, P. K., Angeles, E. R. & Khush, G. S. (1997) *Mol. Breeding* **3,** 105–113.
25. Dioh, W., Tharreau, D. & Lebrun, M. H. (1997) *Nucleic Acids Res.* **25,** 5130–5131.