

Construction of a first-generation physical map of the rice genome

I. Ashikawa, Y. Umehara, H. Tanoue, Z. Wang, S. Saji, T. Shimokawa, K. Yoshino, J. Wu, B. Antonio, K. Koike, W. Van Houten, A. Idonuma, M. Emoto, N. Kurata, and T. Sasaki

We aim to construct a physical map of rice consisting of overlapping yeast artificial chromosome (YAC) clones that span all of the rice chromosomes. Using 1,383 restriction fragment length polymorphism markers and sequence-tagged sites established on our rice linkage map, we have started to screen our rice YAC library, which contains 6,932 rice YAC clones and represents six genome equivalents. As of August 1995, we have used 1,126 markers for the screening. Of these, 1,004 markers hybridized with a total of 4,645 YAC clones in the library. Out of these YACs, 2,369 were independent. We matched the locations of the isolated YACs to corresponding positions on the rice linkage map, detected overlaps between the YACs, and assembled all YACs to contigs. About half of chromosome 6 has now been covered with YAC contigs. We expect that such a combined genetic/physical map will provide a basis for analyzing the genome structure and for map-based cloning of agronomically important genes.

A complete physical map, which consists of ordered recombinant genomic clones spanning an entire genome, is essential for investigating the genome structure of any organism. A physical map has high resolution at the level of deoxyribonucleic acid (DNA) length and directly provides the DNA clones located at distinct positions of the genome for further study. In crop plants, such a map is indispensable for isolating and characterizing genes for agronomically important traits based on their known map position by chromosome walking and/or chromosome landing with DNA clones.

With the advent of yeast artificial chromosome (YAC) vectors capable of cloning DNA fragments of 100 to more than 1000 kb, it has become feasible to assemble DNA clones to construct a physical map in higher eukaryotes. For the human genome, physical maps with ordered YAC clones have been constructed for the long arm of chromosome 21 (Chumakov et al 1992) and chromosome Y (Foote et al 1992). In

Arabidopsis, one-third of the genome was reported to have been covered with YACs (Hwang et al 1991), and these YACs have already been used for map-based cloning of genes (Arondel et al 1992, Bent et al 1994).

We started to use YACs as a cloning vector for the construction of a physical map in rice. We constructed a rice YAC library (Umehara et al 1995) and subsequently screened it for more than 1,100 markers that had been located on a rice linkage map (Kurata et al 1994). By using these markers in the screening, we aim to integrate the physical map with the established molecular genetic map. Such a combined genetic/physical map will serve as a tool in a map-based cloning strategy by facilitating the progress from the genetic locus to the cloned gene.

Constructing and characterizing the rice yac library

High molecular weight genomic DNA was prepared from agarose-embedded protoplasts from cultured suspension cells of rice (*Oryza sativa* L. cv Nipponbare). The DNA was digested with *NotI* or partially digested with *EcoRI*. The digested DNA was size-fractionated, ligated to pYAC4 (for DNA digested with *EcoRI*) and pYAC55 (for DNA digested with *NotI*), and used to transform *S. cerevisiae* AB1380. A total of 6,932 clones with an average insert size of 350 kb was obtained. This library represents approximately six genome equivalents. About 40% of the YACs were estimated to be chimeric. Most of the chimeric YACs were longer than 500 kb; YACs shorter than 400 kb were colinear. The transformants were stocked in 96-well microtiter plates at -80 °C.

In screening this library, 89% of the restriction fragment length polymorphism (RFLP) markers used identified at least one YAC clone. This suggests that YACs in the library distribute fairly evenly over the entire genome. All our results indicate that this YAC library is well-suited for physical map construction and map-based cloning.

Methods for assembling yac contigs on the rice linkage map

Several different methods have been developed and used to construct YAC contig maps (Cohen et al 1993). All of these techniques aim to establish overlaps between YACs that allow the reconstitution of the original genome order of the cloned DNA fragments. Of these techniques, we have mainly employed a strategy of screening the library with our RFLP and sequence-tagged site (STS) markers. We have already constructed a high-density genetic RFLP map of rice (Kurata et al 1994) where 1,236 RFLP markers and 147 STSs were established, on average, at every 300 kb over the entire rice genome. With the average insert size of a YAC being 350 kb, picking up YACs with these mapped DNA markers could assemble the YACs into contigs located at particular positions on the chromosomes. This strategy has been successfully applied to cover the long arm of human chromosome 21 (Chumakov et al 1992) and chromosome Y (Foote et al 1992) with YACs.

To screen the library for the RFLP markers, we made high-density replica filters on which 6,932 DNAs from each YAC clone were gridded. YACs corresponding to

each RFLP marker in the genetic RFLP map were identified by colony hybridization using a set of these high-density filters. YACs recognized as positive by this hybridization were further characterized by Southern hybridization, and this confirmed whether the candidate YACs were derived from the marker loci.

An example of this screening is presented (Fig. 1). The high-density replica blots were hybridized with the probe R1553. Four YACs (indicated by the arrowheads) appeared to be positive in this colony hybridization. DNAs were isolated from these positive YACs, digested with endonucleases, electrophoresed, and hybridized with

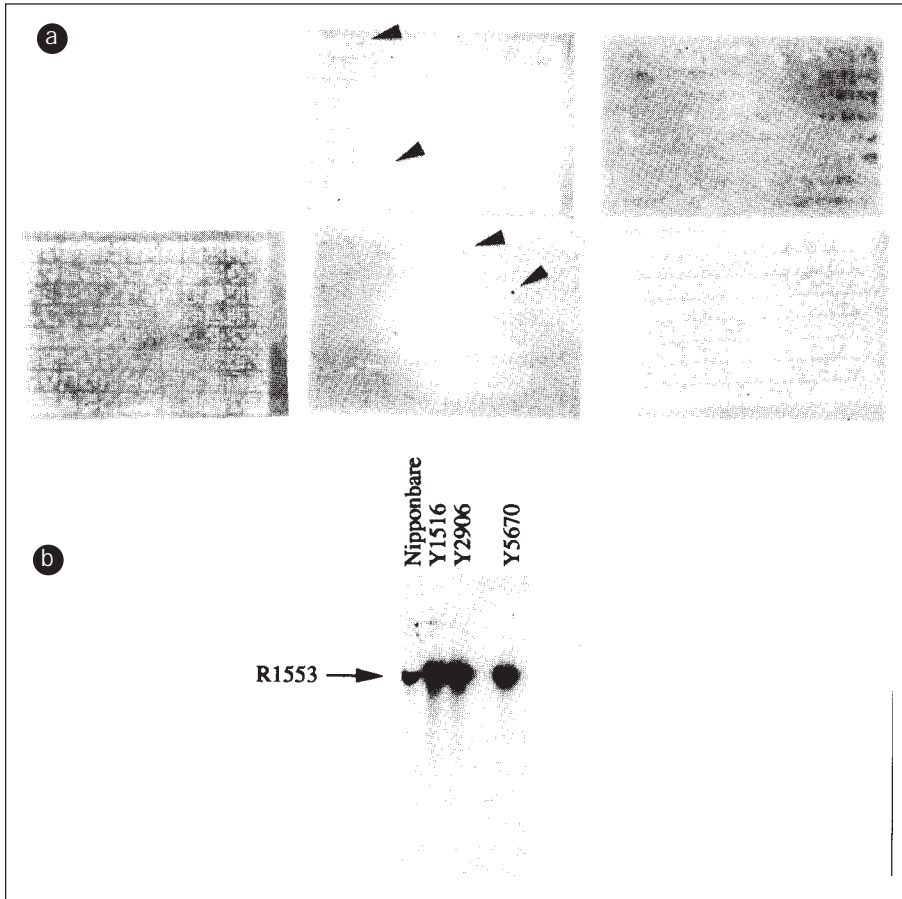


Fig. 1. A result of YAC screening by colony hybridization and Southern blot analysis. a) An autoradiogram image of a set of YAC high-density replica filters hybridized with R1553 probe. Colony hybridization was performed with the ECL system (Amersham). b) Southern blot analysis of candidate YAC DNAs as well as Nipponbare DNA with R1553 probe. The arrow points to the band mapped on the R1553 locus of chromosome 5. From the hybridization pattern, we assigned YACs Y1516, Y2906, and Y5670 to the locus R1553.

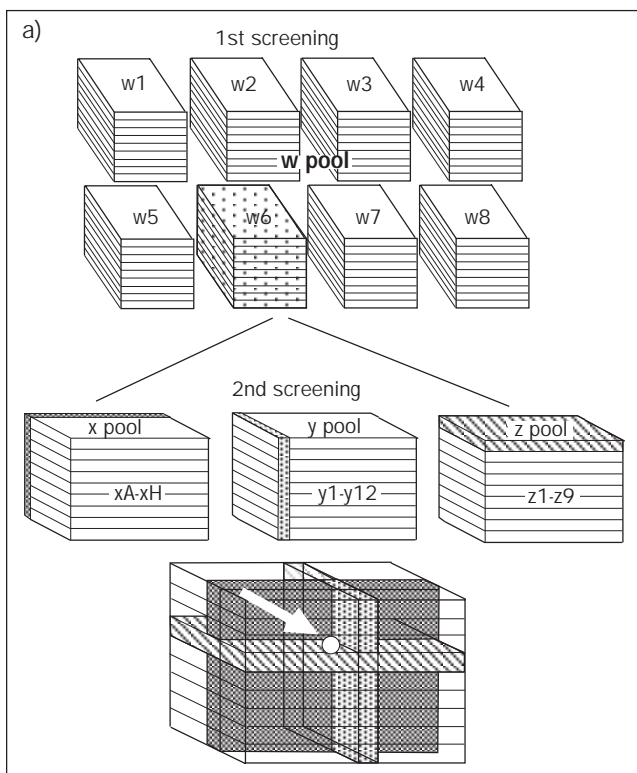
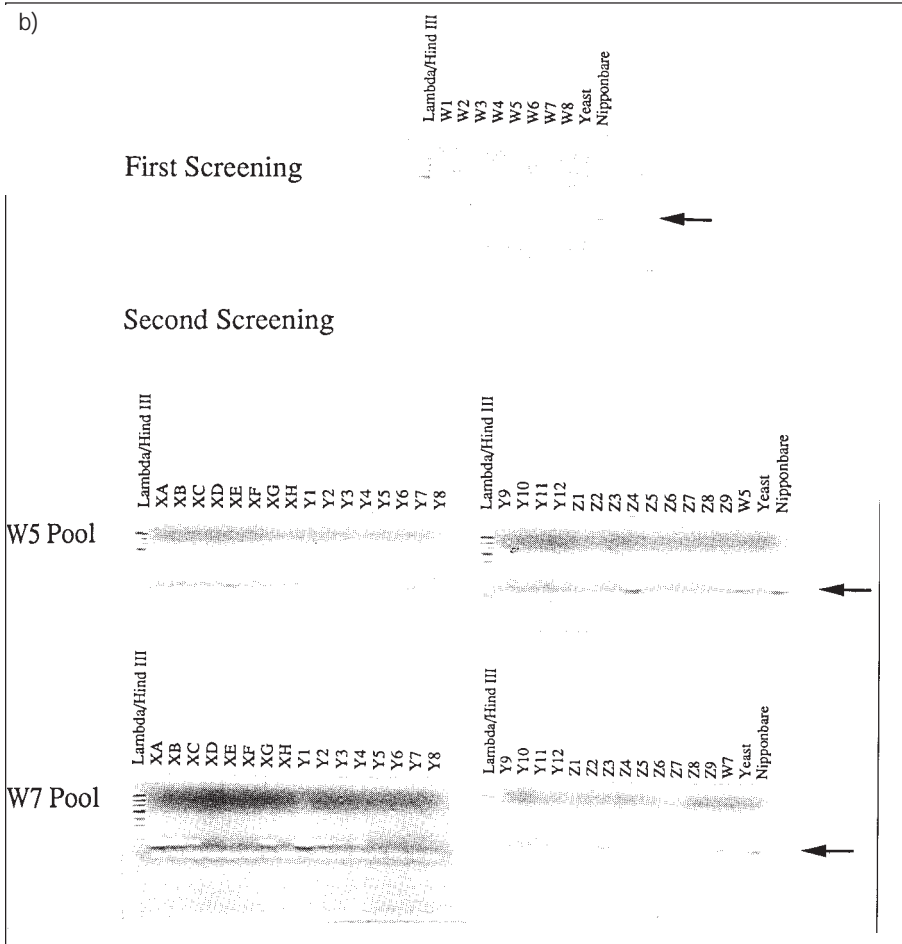


Fig. 2. a) Strategy for screening YACs with rice STSs by PCR amplification. DNAs for all YACs in every 7 or 8 plates (96-well microtiter) were pooled to create a W pool (with 672 or 768 clones in each W pool), and DNAs for YACs in 12 columns (YA-YH: 56 or 64 clones in each Y pool) and in 8 rows (X1-X12: 84 or 96 clones in each X pool) in every 7 or 8 plates were pooled to create a Y pool and an X pool, respectively. DNAs in all YACs in each plate were pooled to make a Z pool (96 clones). Two rounds of PCR, the first for the W pools and the second for the X, Y, and Z pools, identified coordinates of individual YACs containing a given STS. **b)** A result of YAC screening with a rice STS using the strategy in Fig. 2a. In this screening, we used a set of primers for the locus P130. These figures represent results of PCR amplification using the pooled DNAs as templates and the set of primers. The amplified products are indicated by the arrows. The results of the first screening indicate that the sequence corresponding to the locus P130 was contained in YACs in the W5 and W7 pools. From the results of the subsequent second screenings, we identified a YAC in the W5, XE, Y7, and Z4 well and another one in the W7, XA, Y1, and Z3 wells to be positive.

Figure 2 continued



R1553. Out of four candidate YACs, three (Y1516, Y2906, and Y5670) contained the band corresponding to the locus R1553, which assigned this YAC to locus R1553.

Polymerase chain reaction (PCR) amplification was used to identify YACs with rice STSs. In this process, it is not feasible to check amplification for each YAC DNA in the library. We instead created several pools of total YAC DNAs (Fig. 2a). Two rounds of PCR using the pooled DNAs as templates identified coordinates of individual YACs containing a given STS. DNA was then prepared from a single colony and analyzed by PCR to confirm the presence of the STS used to select the clone. An example of YAC identification with the P130 STS marker is given (Fig. 2b). Two YAC clones (Y3955 and Y5601) were identified in this screening.

Results of the screening were stored in a data base that runs on the relational data base system of 4th DIMENSION software. We use this data base to assist in managing and analyzing the YAC screening data on all chromosomes.

Constructing ordered yac clone libraries on all 12 chromosomes

Current status of assembly of rice YACs on all chromosomes

As of Aug 1995, we had screened the library for 1,126 markers (Table 1). Of these, 1,004 markers identified at least one YAC clone. On average, 4.1 YACs were isolated per probe; this number seems smaller than that expected from the size of this library (six haploid equivalents). There may be several regions for which the DNAs are difficult to clone into YACs because of the nature of their physical structure.

In this screening, the 1,004 markers identified 4,645 YACs, of which 2,369 YACs were independent. Through Southern hybridization of the positive YACs with the markers, 1,892 YACs were located at their specific positions on the chromosomes. The remaining 477 (2369-1892) YACs have not yet been located at particular positions because of no RFLP for the detected hybridization bands in those YACs. Among 2,369 YACs, 262 were assigned to more than two separate positions on one or two chromosomes, and these are considered to be chimeric.

We aligned each YAC identified in the screening at a corresponding position on the linkage map, detected overlaps—if any—between neighboring YACs, and integrated these YACs into contigs on the genetic map.

YAC assembly on chromosome 6

The most progress in chromosome assembling of YACs has occurred on chromosome 6, where 257 unique YACs were identified by the 157 markers, with a total of 710 hits to YACs. The YAC locations were identified at particular positions on the chromosome, and they were assembled into 43 contigs. The minimal path of YACs in the contigs comprises 52 YAC clones. Based on the average size of these YACs (350 kb), we estimate the contigs now encompass about 18 Mb or about 50% of this chromosome.

Table 1. Number of isolated YAC clones on different rice chromosomes.^a

Chromosome	Probes	Clones ^b	Clones assigned ^c
1	230	814	341
2	113	446	236
3	162	711	285
4	71	369	212
5	104	595	346
6	157	710	257
7	107	480	193
8	16	71	45
9	21	21	27
10	31	106	85
11	85	257	108
12	47	120	59

^aData of Aug 1995. ^bNumber of YAC clones isolated (simple sums of the number of clones identified with each probe). ^cNumber of independent clones assigned to a particular position on a chromosome.

We are presently refining the preliminary YAC contig maps of regions 1 and 2 on chromosome 6 (Fig. 3a,b). Complete maps of these regions will be published elsewhere. In these regions, most of the markers identified 1 to 11 YACs, confirming that most DNAs of this chromosome were incorporated in our YAC library and could be screened. The largest contig spans more than 6 centiMorgans (cM) (from markers C1003B to C952), and other contigs cover 1-3 cM on the chromosome linkage map. However, in the region of the markers from C1368 to R2749 (indicated by the arrow in Figure 3a), we could not isolate any YACs in our screening. DNA in this region might comprise sequences that are difficult to clone in yeast AB1380.

Physical mapping of DNA markers using assembled YAC clones

The resolution of this physical map appears to exceed the accuracy of our high-resolution genetic map. Many markers, which were not resolvable on the genetic map and were mapped at the same position, could be resolved on the physical map. One example illustrates the YAC contigs in the *eg* (extra glume gene) region on chromosome 1 (Fig. 4). There is one position where six markers (R2018, C346, R2414, G2200, L1082, and R2657A) were located together with no resolution at the level of this linkage map. Close to this position, another clustering of the markers G1133, R2625, and L543 at one position was also observed. These markers were used for screening and each one isolated YACs; some isolated YACs were commonly identified by several independent markers, and other YAC were uniquely isolated by one marker. The YACs isolated were arranged on the linkage map so that their alignment was consistent with the screening results. These arrangements of multiple YAC clones could consequently show us the order of the markers on the linkage map (Fig. 4).

Deducing locations of multiple-copy genes on the chromosomes

Copies of a multiple-copy gene are thought to be distributed at multiple positions on chromosomes. A hybridization pattern of genomic DNA probed with such a gene copy consists of many or several hybridizing bands. However, it is usually fairly difficult to assign the location of each copy corresponding to one of the bands on the chromosomes through genetic mapping. Even in this case, a physical map could give us information on the location or distribution of each copy of a multiple gene on chromosomes.

An example of the possible assignment on rice chromosomes of copies derived from one multiple-copy gene is given (Fig. 5). Southern hybridization patterns of Nipponbare genomic DNA and DNAs from 11 YACs probed with cDNA clone C425 are shown. Although the hybridization pattern of Nipponbare DNA consists of more than eight distinct bands, only one copy corresponding to the band at 3.6 kb (indicated by the arrow) could be assigned to a position on chromosome 6, at locus C425A, through genetic mapping. The hybridization pattern reveals that only YAC Y3685 contains this copy but not that of other copies of the gene. This indicates that the other copies are not near this locus. The copies revealed at 15 kb, 6.5 kb, and 5 kb are contained together in Y1061, Y4089, Y5223, and Y6828, and should thus map closely to each other on the chromosome. On the other hand, the copy corresponding to the

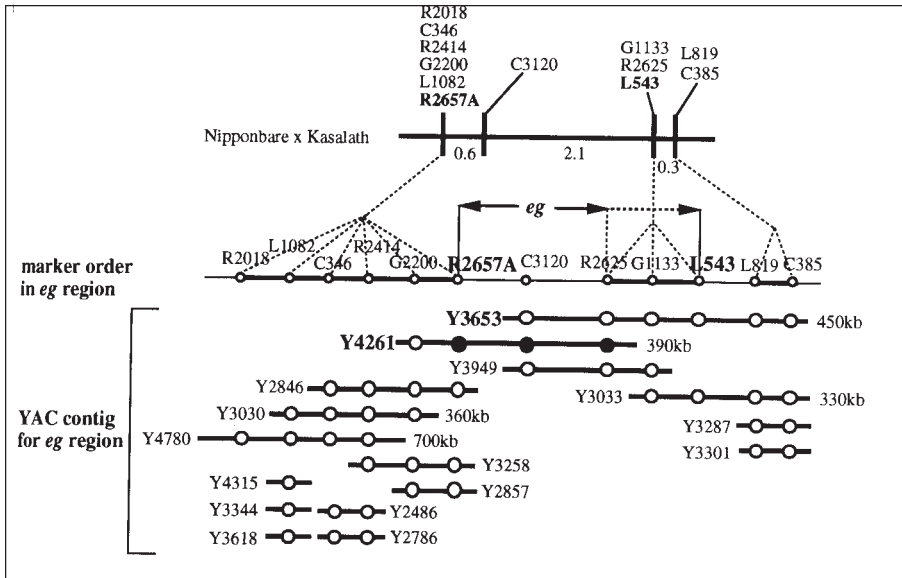


Fig. 4. Genetic map and YAC contig map in the *eg* (extra glume gene) region on rice chromosome 1.

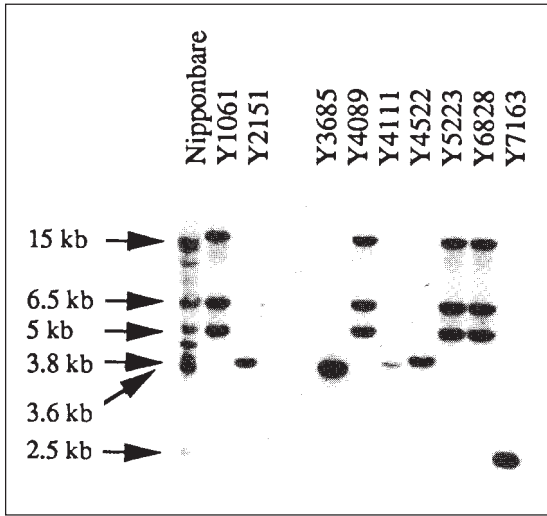


Fig. 5. Southern hybridization patterns of DNAs from 11 YACs together with Nipponbare DNA probed with C425 marker. The hybridizing bands at 3.6 kb, which were detected in Nipponbare DNA and Y3685, have been mapped to the locus C425A. Locations of copies corresponding to other bands have not yet been assigned on any rice chromosomes.

band appearing at 3.8 kb in Y2152, Y4111, and Y4522 and the copy at 2.5 kb in Y7163 should map to different positions on rice chromosomes.

Estimating physical distances on the map

Rate of recombination varies greatly from one region of a chromosome to another (Tanksley et al 1992). Now that we have integrated YAC contigs onto the linkage map, we can investigate this point quantitatively. An average physical distance per cM in the rice genome is calculated to be 273 kb (genome size [430 Mb]/total length of the linkage map [1575 cM]). In the case of chromosome 6, this value varied from 115 kb to more than 2 Mb in various regions, suggesting a very heterogeneous rate of recombination.

Toward refining the physical map

After using more than 1,100 DNA markers for YAC assembly, many gaps between YAC contigs remain on the map. To fill these gaps and to construct a more complete map, we are introducing the following additional methods.

First, after we complete the screening for all 1,383 markers on the genetic RFLP map, another 600 markers—which have been additionally located on our map—will be used for further screening of YACs. Using these additional markers will allow us to isolate YACs that might fill the gaps in the current physical map. Second, after we have used all available DNA markers, we intend to determine the locations of the remaining YAC clones (those that have not been picked up by any DNA marker) by genetic mapping using YAC end-clones or subclones. These YAC subclones are expected to be located in the gap regions. In addition, chromosome walking using YAC end-clones located at the most distal positions of each contig will help fill in the gaps of the physical map. Finally, we are trying to apply a restriction fragment fingerprinting approach to YAC contig assembly, with which we could identify YAC overlaps among neighboring clones and obtain information on the alignment of YACs, including overlapping length in a contig.

Using these combined additional approaches will facilitate assembly of YAC contigs on all rice chromosomes. The constructed physical map will be most useful for studying the entire genome structure of rice. Such a physical map can be used to isolate a collection of overlapping cosmids or phage clones in the area of interest, thereby enabling detailed studies of any genetic locus.

Cited references

- Arondel V, Lemieux B, Hwang I, Gibson S, Goodman HM, Somerville CR. 1992. Map-based cloning of a gene controlling omega-3 fatty acid desaturation in *Arabidopsis*. *Science* 258:1353-1355.
- Bent AF, Kunkel BN, Dahlbeck D, Brown KL, Schmidt R, Giraudat J, Leung J, Staskawicz BJ. 1994. RPS2 of *Arabidopsis thaliana*: a leucine-rich repeat class of plant disease resistance genes. *Science* 265:1856-1860.

- Chumakov I, Rigault P, Guillou S, Ougen P, Billaut A, Guasconi G, Gervy P, LeGall I, Soularue P, Grinas L, Bougueleret L, Chantelot CB, Lacroix B, Barillot E, Gesnoux P, Pook S, Vaysseix G, Frelat G, Schmitz A, Sambucy JL, Bosch A, Estivill X, Weissenbach J, Vignal A, Riethman H, Cox D, Patterson D, Gardinar K, Hattori M, Sakaki Y, Ichikawa H, Ohki M, Paslier DL, Heilig R, Antonarakis S, Cohen D. 1992. Continuum of overlapping clones spanning the entire human chromosome 21q. *Nature* 359:380-387.
- Cohen D, Chumakov I, Weissenbach J. 1993. A first-generation physical map of the human genome. *Nature* 366:698-701.
- Foot S, Vollrath D, Hilton A, Page DC. 1992. The human Y chromosome: overlapping DNA clones spanning the euchromatic region. *Science* 258:60-66.
- Hwang I, Kohchi T, Hauge BM, Goodman HM, Schmidt R, Cnops G, Dean C, Gibson S, Iba K, Lemieux B, Arondel V, Danhoff L, Somerville. 1991. Identification and map position of YAC clones comprising one-third of the *Arabidopsis* genome. *Plant J.* 1:367-374.
- Kurata N, Nagamura Y, Yamamoto K, Harushima Y, Sue N, Wu J, Antonio BA, Shomura A, Shimizu 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. 1994. A 300-kilobase interval genetic map of rice including 883 expressed sequences. *Nat. Genet.* 8:365-372.
- Tanksley SD, Ganai MW, Prince J, de Vicente M, Bonierbale MW, Broun P, Fulton T, Giovannoni J, Grandillo S, Martin G, Messeguer R, Miller J, Miller L, Paterson A, Pineda O, Roder M, Wing R, Wu W, Young N. 1992. High-density molecular linkage map of the tomato and potato genomes: biological inferences and practical applications. *Genetics* 132:1141-1160.
- Umehara Y, Inagaki A, Tanoue H, Yasukochi Y, Nagamura Y, Saji S, Ohtsuki Y, Fujimura Y, Kurata N, Minobe Y. 1995. Construction and characterization of a rice YAC library for physical mapping. *Mol. Breed.* 1:79-89.

Notes

- Authors' address:* Rice Genome Research Program, National Institute of Agrobiological Research, Tsukuba, Ibaraki 305, Japan.
- Availability of the rice YAC library:* Sets of the high-density replica filters of the YAC library reported here are available for screening by hybridization-based method.
- Acknowledgment:* We thank Dr. Ilkka Havukkala for critically reading the manuscript.
- Citation:* [IRRI] International Rice Research Institute. 1996. Rice genetics III. Proceedings of the Third International Rice Genetics Symposium, 16-20 Oct 1995. Manila (Philippines): IRRI.