Available online at www.sciencedirect.com
SCIENCE DIRECT
Plant Science 170 (2006) 889–896
PLANT SCIENCE
www.elsevier.com/locate/plantsci

Two highly representative rice BAC libraries of *japonica* cv Tainung 67 suitable for rice structural and functional genomic research

Yann-Rong Lin [a], Teh-Yuan Chow [b], Meizhong Luo [c], Dave Kudrna [c], Chih-Chi Lin [a], Rod A. Wing [c], Yue-Ie C. Hsing [b,*]

[a] *Department of Life Science, Fu-Jen Catholic University, Hsing Chuang, Taipei Hsien 24205, Taiwan, ROC*
[b] *Institute of Plant and Microbial Biology, Academia Sinica, Nankang, Taipei 11529, Taiwan, ROC*
[c] *Arizona Genomic Institute, University of Arizona, Tucson, AZ-85721, USA*

Received 28 September 2005; received in revised form 15 December 2005; accepted 16 December 2005
Available online 20 January 2006

Abstract

Two deep-coverage bacterial artificial chromosome (BAC) libraries of *Oryza sativa japonica* cv. Tainung 67 (TNG 67), a popular genetic stock in breeding programs and scientific research in Taiwan, have been constructed to facilitate positional cloning of rice genes and to analyze variety-specific genome composition, toward rice structural and functional genomic studies. Good high-molecular-weight DNA was produced well by nuclei preparation from two- to three-week-old seedlings, partially digested by either *Hin*dIII or *Eco*RI, and two cycles of size selection by pulsed-field gel electrophoresis (PFGE). The *Hin*dIII library consists of 45,312 clones, near lack of false positive clones, with an estimated average insert size of 138.4 kb and coverage of 15.1× haploid genome equivalents. The *Eco*RI library consists of 9984 clones, 2% of false positive clones, with an estimated average insert size of 137.8 kb and coverage of 3.2× haploid genome equivalents. Nine single-copy sequence tagged site (STS) markers, located on different chromosomes, were used to screen the two libraries, which a single BAC clone double-spotted on three filters, and hybridized 16–29, with an average of 21.2 BACs. As a result, these two libraries can complement each other and cover nearly 100% of the rice genome, consequently providing efficient tools to isolate any genes of interest. Resources, high-density filters, individual clones, and whole libraries, are available for public distribution and may be accessed at the Institute of Plant and Microbial Biology, Academia Sinica, or Arizona Genomic Institute (AGI).

Keywords: Bacterial artificial chromosome (BAC); Functional genomics; Genomics; Positional cloning; Rice; Tainung 67 (TNG67)

1. Introduction

Besides the two major cereals, wheat and maize, rice provides 23% of the calories consumed by the human population and feeds more than 3 billion people in the world [1]. About 90% of rice is produced in Asia, and the demand for rice in Asia will increase by 25% by 2010 [2]. Rice has been cultivated for more than 7000 years, and thousands of varieties of rice are farmed currently. A total of two cultivated and 22 wild species are found and classified in the Graminae family, *Oryza* genus [3]. *O. glaberrima* is planted only in Africa, while *O. sativa* is widely planted in other parts of the world. Two major subspecies, *indica* and *japonica*, were polyphyletic domesticated since they were derived from the wild species *O. rufipogon* and can be distinguished by hybrid sterility and molecular studies. Further classification by simple sequence repeat (SSR) markers and chloroplast genes revealed five distinct groups: *aus*, *indica*, *aromatic*, *tropical japonica*, and *temperate japonica* [4].

Rice is important, not only as one of the major foods but also as a model organism for scientific research. The comparative linkage mapping of the Poaceae family by use of orthologus DNA reveals a single genetic system in grasses, since the gene order and repertoire are relatively conservative. In addition to the micro-collinearity among these species, rice with the smallest genome of 389 Mb has been considered a model

* Corresponding author. Tel.: +886 2 27892496; fax: +886 2 27827954.
E-mail address: bohsing@gate.sinica.edu.tw (Y.C. Hsing).

0168-9452/$ – see front matter © 2006 Published by Elsevier Ireland Ltd.
doi:10.1016/j.plantsci.2005.12.016

organism in the crop circles [5]. Thus, the whole rice genome has almost been sequenced by the clone-by-clone hierarchical strategy [6]. The complete rice sequence will provide beneficial information to study the comparative genomics of the grasses and, by parallel walking, to assist in cloning orthologous genes in other species with complex genomes, such as maize and wheat.

In addition, the two high-density linkage maps, of restriction fragment-length polymorphism (RFLP) and SSR markers, with an average interval of 190 and 157 kb, respectively, accelerate mapping research resulting in the mapping of a tremendous number of loci corresponding to discrete or quantitative traits (ref: websites of RGP, Oryzabase; [7,8]). Those mapped genes/quantitative trait loci (QTLs) can be isolated by positional cloning [9,10], which has been successfully applied in cloning rice genes regulating heading date, plant height, and even yield [11–13]. Global gene profile studies by EST or SAGE [14–16], knocking out function genes through transposon tagging by exogenous T-DNA or endogenous *TOS17* and through sodium azide [17–20], and activation tagging by T-DNA [21] pave the way to deciphering the functional genomics of rice to understand the roles of single genes at the molecular, biochemical and physical level.

Bacterial artificial chromosome (BAC) vectors, F factor-based plasmids, have been used in the construction of large-insert genomic libraries because of their ability to carry DNA fragments up to 350 kb, near lack of chimerism, and easy manipulation [22,23]. Many BAC vectors, derived from pBeloBAC11 (7.4 kb, [22]), are modified for specific purposes such as an additional *Eco*RI cloning site (pECBAC1: [24]; pIndigoBac536: [25]), transformation (BIBAC: [26]; pBACwich: [27]; TAC: [28]), and high copy number for vector preparation (pCUGIBAC1: [29]; pAGIBAC1: M. Luo, S.S. Ammiraju, and R.A. Wing, unpublished data). By high throughput fingerprinting of BACs, which are fragmented by *Hin*dIII or *Eco*RI and then easily subjected to separation by agarose gel electrophoresis, BAC clones are very efficiently assembled into contigs for physical mapping prior to hierarchical genome sequencing [30]. BAC clones are also suitable for fluorescence in situ hybridization (FISH) for physical mapping and localizing genes on chromosomes [31]. Therefore, BAC vectors are extensively applied in genomic and functional studies of plants and animals. For example, the aims of The Green Plant BAC Library Project (www.greenbac.org), which constructs BAC libraries of green algae, non-seed land plants, and seed plants, and of The Oryza Map Alignment Project (OMAP, www.omap.org), which constructs BAC libraries of the *Oryza* genus, are to decipher critical genes of speciation and to study genome evolution since diversification.

Other than species included in the OMAP project, numerous rice BAC libraries have been constructed in the past decade to facilitate rice genome and functional genome research. The first rice BAC library was constructed by using the IR-BB21 line for positional cloning of *Xa-21* [32]. Later on, at least seven *indica* varieties, C101A51, Guang Lu Ai 4, Milyang23, Minghui, IR64, IR-BB56, and Tequing, and at least four *japonica* varieties, Azucena, Lemont, Nipponbare, and Shimokita, were constructed for cloning genes such as resistant genes for functional studies and for physical mapping and genome sequencing [23,28,33–38]. Even though more than 10 rice BAC libraries were established, genome-wide variation also was uncovered by comparative analysis of DNA sequences of *indica* and *japonica* [39]. The differences in the gene contents or the alleles of different varieties may account for the phenotypic variations corresponding to these varieties.

We constructed two BAC libraries of *O. sativa* ssp. *japonica* cv. Tainung 67 (TNG67) because TNG67 is not only an elite cultivar but also the major variety used in scientific investigation of physiology, genetics, functional genomics, and proteomics in Taiwan. TNG67, derived from *japonica*, *indica*, and even tropical *japonica* rice in breeding programs ([40]; ref: Taiwan Rice Information System, TRIS, http://tris.tari.gov.tw:8080/index.htm), possesses good agronomic characters such as high yield, resistance to rice blast and bacterial leaf blight and is photoperiodic insensitive. TNG67, with a blending genome composition of various subspecies and superior characteristics, has been widely used as germplasm for developing elite commercial varieties. Besides its significance in agriculture, TNG67 is the genetic stock for studies in physiology, genetics, and functional genomics. More than 50,000 mutant lines were generated by chemical mutagenesis or T-DNA insertion. Therefore, these two libraries would be very useful to facilitate rice research in the post-genomic era.

2. Materials and methods

2.1. Plant material and preparation of high-molecular-weight DNA

O. sativa ssp. *japonica* cv. Tainung 67 rice was grown in a growth chamber at 30 °C for two to three weeks. The leaf tissues of young seedlings were harvested and applied directly for megabase DNA preparation by the nuclei preparation method described by Luo and Wing, with minor modification [41]. About 80 g of fresh leaf tissue was grounded into powder with liquid nitrogen by mortar and pestle, and then was subjected to nuclei preparation [41]. Once the nuclei had been isolated, the nucleus pellet was gently suspended with about 1.2 ml of NIB (10 mM Tris–HCl, pH 8.0, 10 mM EDTA, pH 8.0, 10 mM KCl; 0.5 M sucrose, 4 mM spermidine, 1 mM spermine), then warmed up to 45 °C and mixed with pre-warmed 1.2 ml of 1% Seakem Gold agarose (FMC Bioproducts) by slow pipetting. The nuclei embedded in agarose plugs were lysis in proteinase K solution (1% *N*-lauroylsarcosine, 0.5 M EDTA, pH 9.2, 1 mg/ml proteinase K) at 50 °C for 48 h, with one change of fresh proteinase K solution, and the lysis nuclei were stored in fresh proteinase K solution at room temperature.

2.2. BAC library construction

Both *Hin*dIII and *Eco*RI BAC libraries were established by adopting the protocols described Luo and Wing [41]. The high-molecular-weight DNA was prepared from partial digested

either *Hin*dIII or *Eco*RI Tainung 67 genomic DNA fragments, which were consequently deployed by two cycles of size selection [41]. Large DNA fragments were electroeluted from the gel fractions by using Model 422 Electro-Eluter (Bio-Rad, USA) and used immediately for DNA ligation.

Large size-selected *Hin*dIII or *Eco*RI fragments were ligated into *Hin*dIII- or *Eco*RI-digested and dephosphorylated pAGI-BAC1. The ligation and transformation were carried out as described by Luo and Wing [41]. The transformants were screened by LB containing 12.5 μg/ml of chloramphenicol and X-gal and IPTG. The white colonies were picked robotically (Genetix Q-Bot, Genetix Ltd., Dorest, UK) into 384-well microtiter plates containing LB freezing media (LB, 36 mM K_2HPO_4, 13.2 mM KH_2PO_4, 1.7 mM sodium citrate, 0.4 mM $MgSO_4$, 6.8 mM $(NH_4)_2SO_4$, 4.4% (v/v) glycerol, 12.5 μg/ml chloramphenicol) and incubated in 37 °C overnight before storage at −80 °C.

2.3. Preparation of BAC DNA and estimation of insert size

BAC clones were randomly selected from the two libraries for estimation of insert size. Each BAC was inoculated in 5 ml LB medium with 12.5 μg/ml chloramphenicol overnight. The conventional alkaline lysis method of plasmid DNA preparation was used to isolate BAC DNA. The DNA pellets were dissolved in 30 μl of TE, and 10 μl of BAC DNA was digested with 10 units of *Not*I (NEB, USA). The digested BAC DNA was separated on 1% agarose gel by pulsed-field gel electrophoresis, PFGE (Rotaphor, Biometra, Germany), in 0.5× TBE buffer at 6 V/cm at 14 °C, with a linearly ramped pulse time of 2–10 s for 18 h. The insert size of BAC was estimated with reference to MidRange II PFG Marker, a 24.5 kb DNA ladder marker (NEB, USA).

2.4. DNA probes and colony hybridization

Nine sequence tagged site (STS) markers were randomly selected from 12 chromosomes to test the genome coverage of the libraries. The location on linkage maps and primer sequences of rice STS are available at the Japanese Rice Genome Research Program Web site (http://rgp.dna.affrc.go.jp/). Oligo primers of each STS marker were synthesized by GENSET Singapore Biotech. Genomic DNA preparation of TN67 was as described [42]. The DNA fragments of STS markers were amplified by PCR with a final volume of 25 μl PCR reaction: 50 ng TN67 genomic DNA, 200 μM of each deoxynucleotide triphosphate, 0.2 μM of primers, 0.5 units *ProTaq* (Protech, Taipei, Taiwan), 10 mM Tris–HCl, 50 mM KCl, 1.5 mM $MgCl_2$, and 0.1% Triton X-100, and was carried out with use of a Biometra Model T1Thermocycler as follows: 94 °C for 5 min for 1 cycle; 94 °C for 1 min, 55 °C for 1 min, 72 °C for 2 min, for 35 cycles; 72 °C for 5 min for 1 cycle.

The chloroplast clone covering positions −715 to +48 of the rice *psaA* gene was kindly provided by Dr. S.G. Chen (Institute of Plant and Microbial Biology, Academia Sinica, Taipei, Taiwan; [43]) and was used to assess the proportion of BAC libraries containing chloroplast DNA. Clone inserts, putatively free of vector DNA, were used.

The BAC clones of these two libraries were double-spotted on 22.5 cm × 22.5 cm Hybond N+ nylon membranes (Amersham, Piscataway, NJ) in the format of 4 × 4 six fields by use of Genetix Q-Bot (Genetix Ltd.). The entire two libraries containing 55,296 (144 × 384) clones were gridded on three filters, each filter containing 18,432 (48 × 384) individual BAC clones. The three filters were named A, B, and C, and filters A and B each contain 18,432 (48 × 384) *Hin*dIII BAC clones and filter C contains 8448 (22 × 384) *Hin*dIII BAC clones and 9984 (26 × 384) *Eco*RI BAC clones. One set of three high-density filters with a macroarray of BAC clones underwent library screening by colony hybridization.

Colony hybridization was adopted to estimate the genome coverages of the two BAC libraries. The PCR-amplified DNA fragments of nine STS markers and the clone inserts of *psaA*, puatively free of vector DNA were labeled with [^{32}P] dCTP (Amersham, USA) by random priming with hexamers (Invitrogen, USA). The hybridization procedure followed the protocol of Arizona Genomic Institute (AGI) (http://genome.arizona.edu/).

3. Results

3.1. Construction of the O. sativa japonica cv TNG67 BAC libraries

We constructed two highly representative BAC libraries of TNG67 (*Hin*dIII and *Eco*RI) to facilitate the rice structural and functional genomic research in Taiwan by adopting the well-established procedures of vector DNA, high-molecular-weight DNA preparation, ligation, and transformation [41]. Because of the careful preparation of vector pAGIBAC1 digested with *Hin*dIII or *Eco*RI, none of the *Hin*dIII BACs and only 2% of *Eco*RI BACs were detected as false positive, white colonies without inserts (Table 1). Since the second size selection filters out small DNA fragments co-migrating with large DNA, we used two cycles of size selection for high-molecular DNA preparation in constructing both *Hin*dIII and *Eco*RI libraries. The average insert sizes of the *Hin*dIII and *Eco*RI BAC libraries were 138.4 and 137.8 kb, respectively (Table 1).

The *Hin*dIII library consists of 45,312 clones stored in one hundred and eighteen 384-well microtiter plates. A total of 382 clones were randomly selected for insert size estimation via mini-preparation, followed by *Not*I digestion and separation of insert and vector DNA subjected to PFGE. The common band of vector of around 7 kb and one or more bands of insert DNA, depending on the number of internal *Not*I restriction sites, were

Table 1
Characterization of TNG67 *Hin*dIII and *Eco*RI libraries by insert sizing

BACs	*Hin*dIII	*Eco*RI
Total number of clones	45312	9984
Average frequency of false positive clones	0	0.02
Number of clones for size estimation	384	178
Average insert size (kb)	138.4	137.8
Average number of *Not*I site	3.1	2.9
Genome coverage	15.1	3.2

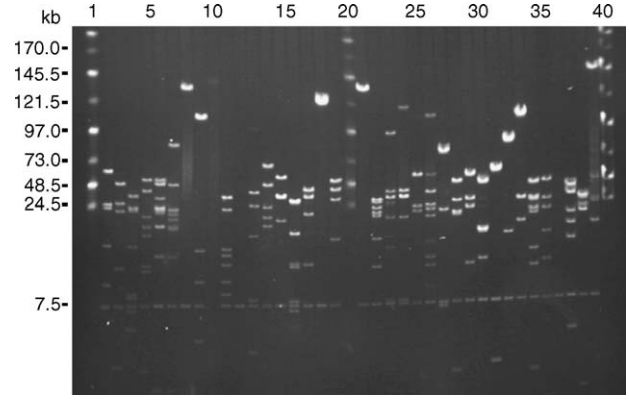

Fig. 1. Size determination of BAC clones randomly selected from the TNG67 *Eco*RI BAC library. DNA was isolated by alkaline lysis method and digested with *Not*I. The digested BAC DNA was separated 1% pulsed field gel with a linearly ramped pulse time of 2–10 s at 180 V/cm at 14 °C in 0.5× TBE buffer for 18 h. Lanes 1, 20 and 40 are the MidRange II PFG Marker. The 7.5-kb common band is the vector pIndigoBac536.

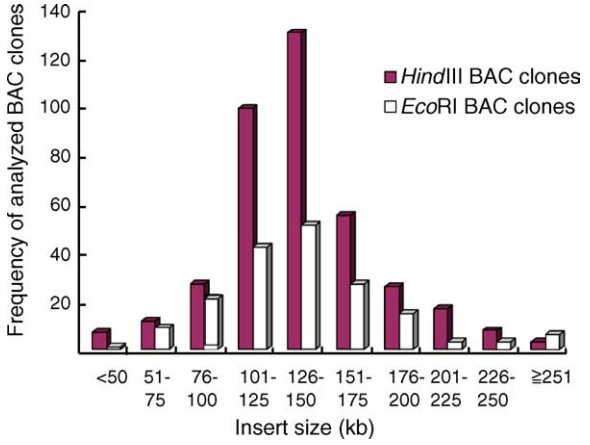

Fig. 2. Insert size distribution of the TNG67 *Hin*dIII BAC and *Eco*RI BAC libraries. DNA was isolated from 384 clones, and 174 clones were randomly picked from *Hin*dIII BAC library and *Eco*RI BAC library, subjectively.

detected for every *Hin*dIII clone (Fig. 1). The insert sizes of the BACs ranged from 20 to 273 kb, with an average of 138.4 kb (Table 1). The majority of insert sizes (82%) are 100–200 kb, and only 11% of insert sizes are smaller than 100 kb (Fig. 2). The number of *Not*I sites ranged from 0 to 9, with an average of 3.1 sites (Table 1; Fig. 2).

The *Eco*RI library consists of 9984 clones stored in twenty-six 384-well microtiter plates. A total of 178 clones were randomly picked to check the insert sizes and *Not*I sites (Fig. 2). About 2% of *Eco*RI BACs were vectors only, with no insert, thus false positives, resulting in white colonies caused by damage of multiple cloning site during vector preparation. The insert sizes of the BACs ranged from 40 to 303 kb, with an average of 137.8 kb (Table 1). Most of the insert sizes (78%) were 100–200 kb, and only 15% were smaller than 100 kb (Fig. 2). The number of *Not*I sites ranged from 0 to 8, with an average of 2.9 sites (Table 1; Fig. 2).

3.2. Screening of the BAC libraries

Colony hybridization of known STS markers and a chloroplast gene were used to estimate the genome coverage and chloroplast DNA content, respectively, of the two libraries. Nine STS markers distributed on nine different chromosomes were selected to hybridize the two libraries with a set of three filters, which a single BAC clone double-spotted on a 22.5 × 22.5 membrane. Sixteen to 29 BACs, with an average of 21.2, were identified (Table 2). The approximate average numbers of identified BAC clones were 6.9 and 6.2, from filter A and B, respectively; however, more BAC clones, 8.1 on average were identified on filter C (Table 2). The discrepancy is presumed to be the larger insert size in the *Eco*RI library, since filter C comprises 8448 *Hin*dIII clones and 9984 *Eco*RI clones.

Since it is difficult to eliminate all chloroplast DNA while preparing nuclei DNA with green young leaf tissues, the chloroplast DNA was also incorporated into vector pAGIBAC1 via size selection, ligation and transformation, together with nuclear DNA. The content of chloroplast DNA, which cannot account for nuclear genome coverage, must be estimated. The chloroplast-specific gene, *psaA*, with 763 bp [43], was used to screen the two whole libraries and identified 3420 clones. Hybridization results of one of the three high-density filters are shown in Fig. 3. The rice chloroplast genome of *O. sativa* spp. *japonica* cv Nipponbare is about 134,525 bp [44]. It is reasonable to assume that very few clones containing chloroplast DNA were undiscovered, because the average insert sizes of the BAC libraries are 138.4 and 137.8 kb, respectively. As a result, 6.2% of these libraries comprised chloroplast DNA.

3.3. Estimation of genome coverage

The genome coverage of the libraries was estimated on the basis of the average insert size; the clone number excluded the 6.2% of chloroplast DNA and false positives, and the genome size of rice, 389 Mb [6]. The average insert size of the *Hin*dIII library was 138.4 kb, with 45,312 BACs, and 0% of false

Table 2
Characterization of TNG67 *Hin*dIII and *Eco*RI libraries by colony hybridization with nuclear DNA and chloroplast DNA

Probe	Chromosome	No. of BACs			
		Filter A	Filter B	Filter C	Total
Nuclear DNA					
S13528	1	4	4	8	16
E0437	2	6	5	10	21
E30531	5	11	6	9	26
R03879	6	6	6	6	18
S20922	7	7	9	9	25
R2118	8	6	7	8	21
C11021	9	5	6	7	18
C0011	10	7	4	6	17
E3676	11	10	9	10	29
Average		6.9	6.2	8.1	21.2
Chloroplast DNA					
psaA		995	998	1427	3420

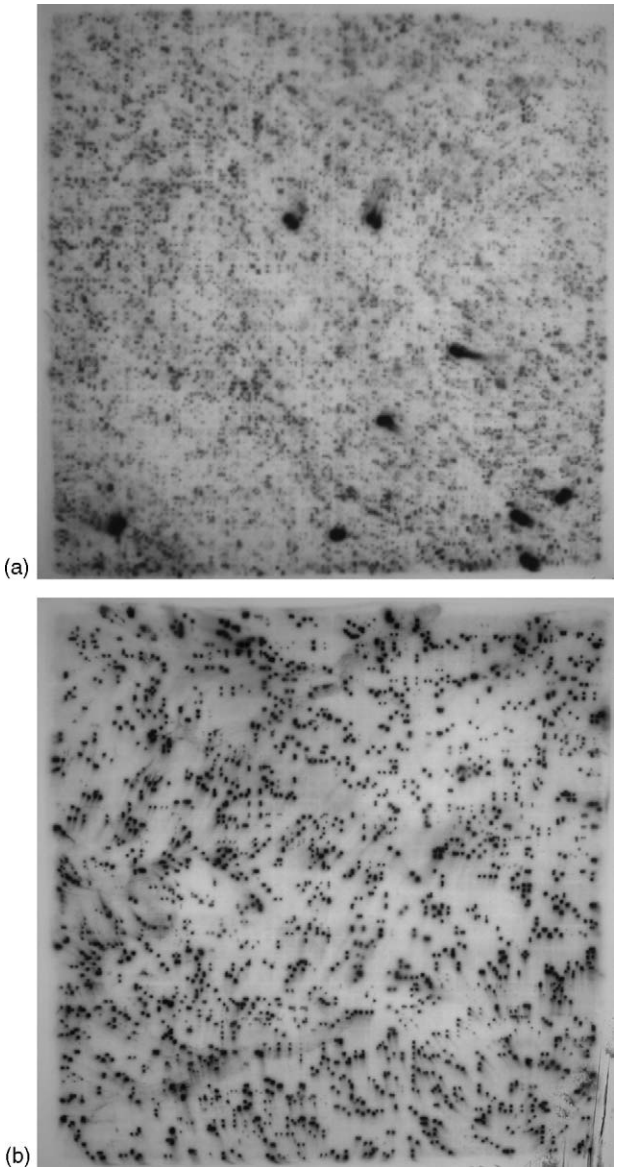

Fig. 3. Screening of TNG67 BAC libraries by colony hybridization. Autoradiogram images of the probe E30531 on chromosome 5 hybridized against filter C, colonies from 22 384-well plates of the *Hin*dIII BAC library and 26 384-well plates of the *Eco*RI BAC library (left), and of the probe chloroplast gene, *psaA*, hybridized against filter A, colonies from 48 384-well plates of the *Hin*dIII BAC library (right). Each colony was doubled-spotted in high density using a Genetix Q-Bot on a 22.5 × 22.5-cm Hybond N+ nylon membranes in the format of 4 × 4 six fields. Each filter contains 18,432 clones.

positives accounted for 15.1× of genome coverage. The average insert size of the *Eco*RI library was 137.8 kb, with 9984 BACs, and 2% of false positives accounted for 3.2× of genome coverage. In total, these two libraries comprised 18.3× of haploid genome equivalents. An alternative way to estimate genome coverage is to apply single-copy DNA for colony hybridization. Based on the single-copy STS markers, the average number of hit clones was 21.2, which indicates that the two libraries covered approximately 21.2× of the genome coverage.

The genome coverage estimation of the two libraries based on calculation and colony hybridization slightly differed: 18.3× and 21.2× for *Hin*dIII and *Eco*RI, respectively, of haploid genome equivalents. The discrepancy might be due to underestimation of insert sizes by rough calculation with reference to the MidRange II PFG Marker after PFGE analysis or overestimation of genome coverage by nonspecific binding of the short radioactive labeled DNA fragments of STS markers through colony hybridization. On the basis of the overall average insert size and the number of total BACs, these two TNG67 libraries, cover nearly 100% of the rice genome and provide very good tools, with nearly 100% probability of isolating a gene from a specific genomic region.

3.4. Public availability of the BAC libraries and resources

The two BAC libraries constructed as a result of this work were made available to the public from the Arizona Genomic Institute. Resources, in the forms of screening filters, individual clones and whole libraries, are distributed on a cost-recovery basis. The *Hin*dIII library is named: OSJTBa, and the *Eco*RI library is named OSJTBb. The resources may be ordered from the AGI website, http://genome.arizona.edu/orders/. Alternatively, one copy of the two libraries is kept in the Institute of Plant and Microbial Biology, Academia Sinica, and the other copy is deposited in the Bio-resource Collection and Research Center, Food Industry Research and Development Institute, HsingChu, Taiwan. We also made 15 sets of high-density filters. These filters can be hybridized dozens of times and are also available with agreement. To request the filters or BAC clones, please contact Yue-Ie C. Hsing, Institute of Plant and Microbial Biology, Academia Sinica, Taiwan (http://genome.sinica.edu.tw).

4. Discussion

Large-insert libraries, constructed with vectors of yeast artificial chromosome (YAC), P1-derived artificial chromosome (PAC), and bacterial artificial chromosome, are crucial for physical mapping of contig assembly, genome sequencing, and positional cloning. Because of the advantageous characteristics of clone stability, almost no chimeric clones, and very user-friendly manipulation, BAC has replaced YAC and PAC as the favored vector for constructing genomic libraries with average insert sizes larger than 100 kb and has been employed in numerous organisms. In this study, we constructed two deep-coverage BAC libraries of ssp. *japonica* TNG67 to accelerate rice structural and functional genomic studies. These two libraries can complement each other because insert DNA was prepared by partial digestion of two different enzymes, *Hin*dIII and *Eco*RI. The genome coverage of *Hin*dIII and *Eco*RI BAC libraries are 15.1× and 3.2×, respectively, of haploid genome equivalents (Table 1). As a result, the two libraries can cover nearly 100% of the rice genome and provide the opportunity to determine a contig encompassing specific genomic region and to clone interesting genes.

4.1. Assessments of the two TNG67 BAC libraries

The average insert sizes of *Hin*dIII and *Eco*RI BAC libraries are 138.4 and 137.8 kb, respectively, and are quite large among established BAC libraries. The insert size, which is important

for contig assembly and to encompass target genomic regions, depends on megabase DNA isolation of nuclei followed by partial digestion and size selection by separation of fragmented DNA through PFGE. Two cycles of size selection are recommended by many research groups to achieve higher average insert sizes, since the second cycle of size selection can eliminate trapped small-fragment DNA in the first size-selected DNA [34,45]. By following up the well-established size selection procedure of the Arizona Genomic Institute [41], the first size selection, with a pulse time of 1–50 s linear ramp, and the second size selection, with a pulse time of 4 s constant time, resulted in the insert size of 89% of *Hin*dIII BACs and 85% of *Eco*RI BACs being larger than 100 kb (Fig. 1).

One of the major obstacles encountered while making BAC libraries is false positive white clones with no inserts, resulted from the inactivated *lacZ* of multiple clone sites by vector DNA damage rather than insertion of genomic DNA. Many factors can result in false positives, such as vector storage, enzyme digestion and dephosphorylation of vector DNA, and electroporation [35,45]. None of the *Hin*dIII BACs of TNG67 is false positive under the inspection of 384 clones (Table 1). However, 4 of the 178 *Eco*RI BACs, or consequently 200 (0.02×9984) in the *Eco*RI BAC library, were found to be false positive during the insert size estimation (Table 1; Fig. 1). Nevertheless, the effect of false positive can be negligible as compared with the whole TNG67 BAC libraries, 55,296 clones.

Nuclei prepared from green leaves of seedlings are accompanied by organelles such as mitochondria, plastid, and chloroplast especially. To prevent the harvesting of plastids, Triton X-100 is added in the NIBT (10 mM Tris–HCl, pH 8.0, 10 mM EDTA, pH 8.0, 10 mM KCl; 0.5 M sucrose, 4 mM spermidine, 1 mM spermine, 10% Triton X-100) during nucleus preparation, which can disrupt the integrity of organelle membrane structure. Nevertheless, the content of BACs containing chloroplast DNA range from 0.1 to 2.9% in most constructions [22,28,31,36,37,44]. In the current study, about 6.2% of TNG67 BACs were hybridized to cp DNA, *psaA* (Table 2; Fig. 3), which is higher than the proportions in the above studies. One of the reasons could be that chloroplast DNA was integrated into the nucleus during evolution as reported [6], resulting in overestimating the number of clones containing chloroplast DNA. We performed a BLAST search of the 763 bp of *psaA*, used for labeling, against 12 rice pseudomolecules [6], and found five hits, two located on chromosome 10, and the others located chromosomes 1, 4, and 5. Thus, this 6.2% is overestimated and can be negligible in the two deep-coveraged BAC libraries. Another reason for the high proportion of cp clones could be that relatively less Triton X-100 was used (0.5%) in the working solution of NIBT or not enough times were repeated for washing. However, the proportion of chloroplast BACs included in the TNG67 BAC libraries is acceptable, as compared to the 7% of the cv Shimokito BAC library, for which HMW DNA was prepared from protoplasts [34]. Our libraries still represent the genome of TNG67 highly after subtracting chloroplast BACs (Table 1; Table 2).

To verify the genome coverage of the TNG67 BAC libraries, nine STS markers were hybridized to the three high-density filters and turned out the high hit numbers, with an average of 21.2 (Table 2, Fig. 3). Some discrepancy always exists between the estimations of genome coverage by numerical calculation of insert sizes and by colony hybridization. The genome coverage of the two libraries inferred by numerical calculation (18.3×) was lower than that inferred by colony hybridization (21.2×). This outcome is in contrast to that of the other rice BAC libraries, for which the genome coverage estimated by colony hybridization was less than that estimated by numerical calculation of insert sizes, at 15 and 12% for the *indica* cv Teqing *Hin*dIII and the *japonica* cv Lemont *Hin*dIII BAC libraries, respectively [23], and 20.6, 6.5, and 7.5% for the *japonica* cv Nipponbare *Bam*HIII, *Eco*RIII, and *Hin*dIII BAC libraries, respectively [38]. We might have underestimated average insert sizes or nonspecific binding of short fragments of STS probes in our study. The hits of STS markers ranging from 16 to 29 could be due to various specificity of STS sequences or to chromosome structures where STS are localized [33]. We still can draw the conclusion that the genome coverage of the TNG67 BAC libraries is at least 18.3× of haploid genome equivalents.

4.2. The utilization of the two TNG67 BAC libraries

Japonica Tainung 67, the leading commercial variety, possesses good agronomic characters such as photoperiod insensitivity; high yield; semi-dwarf stature; tolerance to lodging, brown plant hoppers, and sclerotial stem rot; and resistance to blast and bacterial leaf blight [40], and is widely used as germplasm in breeding programs. Besides its significance in agriculture, TNG67 has been the most favored material for rice research in physiology, in genetics, and in functional genomics in Taiwan.

Global studies of every gene at physiological, biochemical and molecular levels are the main aims in the functional genomic era. Genome-wide mutated genes, induced by radiation, chemical mutagens, or transposable elements resulting in various phenotypes (phenomics), can be isolated by forward genetics via positional cloning, also called map-based cloning, or by reverse genetics via gene tags by flanking sequences of transposable elements. Tainung 67 was mutated genome-wide by chemical mutagens or T-DNA insertion for functional genomic research in Taiwan. Approximately 3000 TNG67 mutants were induced by use of sodium azide [17]. More than 50,000 mutants were created by T-DNA insertion, and more than 12,000 flanking sequence tags (FSTs) were collected and annotated (Hsing et al., submitted for publication); the information/seeds of T-DNA knockout mutants will be available to the public in the near future.

Positional cloning is the strategy to isolated natural variation or induced mutation. In the post-genomic era, two steps of positional cloning, high-resolution mapping and identification of candidate genes by transformation, are critical in organisms in which genome sequences and annotation are accomplished [10]. The rice transformation techniques are sophisticated, so the decisive step would be the high-resolution mapping, which can be achieved by polymorphic markers and a large volume of

recombinants [9,10]. Since tremendous TNG67 mutants were generated for functional study and many genes are preferable to clone by positional cloning, we are starting to seek suitable varieties to cross with TNG67 to obtain sufficient recombinants and polymorphism for high-resolution mapping. We surveyed the polymorphism of the TNG67 vs. 38 *japonica* and 45 *indica* varieties originating from Taiwan, Japan, IRRI, and the United States against 106 PCR markers, covering 93% of the rice genome (Y.-R. Lin et al., unpublished data). The above information of TNG67 will be released on our Web site in the near future.

These two highly representative TN67 BAC libraries should be beneficial to rice structural and functional genomic studies. In addition, they specifically pave the way to determining TNG67-specific alleles. Resources, high-density filters, individual clones, and whole libraries are available to all researchers.

Acknowledgements

This work was supported by the Thematic Project, Academia Sinica, Taipei, Taiwan. The authors would like to thank Dr. Yong-Pei Wu for seeds of TNG67 and Dr. Susan McCouch for seed importation, and Ms. Laura Heraty for English editing. We also appreciate Dr. Wenming Wang, Chris Mueller and Samina Makda for technical assistance.

References

[1] G.S. Khush, Productivity improvements in rice, Nutr. Rev. 61 (2003) S114–S116.
[2] K.S. Fisher, J. Barton, G.S. Khush, H. Leung, R. Cantrell, Genomics and agriculture. Collaborations in rice, Science 290 (2000) 279–280.
[3] G.S. Khush, D.S. Brar, Rice, in: V.L. Chopra, S. Prakash (Eds.), Evolution and Adaptation of Cereal Crops, Science Publishers, Inc., New Hampshire, USA, 2002, pp. 1–41.
[4] A.J. Garris, T.H. Tai, J. Coburn, S. Kresovich, S. McCouch, Genetic structure and diversity in *Oryza sativa* L, Genetics 169 (2005) 1631–1638.
[5] K.M. Devos, Updating the 'crop circle', Curr. Opin. Plant Biol. 8 (2005) 155–162.
[6] International Rice Genome Sequencing Project, The map-based sequence of the rice genome, Nature 436 (2005) 793–800.
[7] Y. Harushima, M. Yano, A. Shomura, M. Sato, T. Shimano, Y. Kuboki, T. Yamamoto, S.Y. Lin, B.A. Antonio, A. Parco, H. Kajiya, N. Huang, K. Yamamoto, Y. Nagamura, N. Kurata, G.S. Khush, T. Sasaki, A high-density rice genetic linkage map with 2275 markers using a single F2 population, Genetics 148 (1998) 479–494.
[8] S.R. McCouch, L. Teytelman, Y. Xu, K.B. Lobos, K. Clare, M. Walton, B. Fu, R. Maghirang, Z. Li, Y. Xing, Q. Zhang, I. Kono, M. Yano, R. Fjellstrom, G. DeClerck, D. Schneider, S. Cartinhour, D. Ware, L. Stein, Development and mapping of 2240 new SSR markers for rice (*Oryza sativa* L.), DNA Res. 9 (2002) 199–207.
[9] S.D. Tanksley, M.W. Ganal, G.B. Martin, Chromosome landing: a paradigm for map-based gene cloning in plants with large genomes, Trends Genet. 11 (1995) 63–68.
[10] G. Jander, S.R. Norris, S.D. Rounsley, D.F. Bush, I.M. Levin, R.L. Last, Arabidopsis map-based cloning in the post-genome era, Plant Physiol. 129 (2002) 440–450.
[11] M. Ashikari, J. Wu, M. Yano, T. Sasaki, A. Yoshimura, Rice gibberellin-insensitive dwarf mutant gene *Dwarf1* encodes the α-subunit of GTP-binding protein, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 10284–10289.
[12] M. Ashikari, H. Sakakibara, S. Lin, T. Yamamoto, T. Takashi, A. Nishimura, E.R. Angeles, Q. Qian, H. Kitano, M. Matsuoka, Cytokinin oxidase regulates rice grain production, Science 309 (2005) 741–745.
[13] Y. Takahashi, A. Shomura, T. Sasaki, M. Yano, Hd6, a rice quantitative trait locus involved in photoperiod sensitivity, encodes the α-subunit of protein kinase CK2, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 7922–7927.
[14] S. Kikuchi, K. Satoh, T. Nagata, N. Kawagashira, K. Doi, N. Kishimoto, J. Yazaki, M. Ishikawa, H. Yamada, H. Ooka, et al., Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice, Science 301 (2003) 376–379.
[15] M. Gowda, C. Jantasuriyarat, R. Dean, G.-L. Wang, Robust-longSAGE: a substantially improved longSAGE method for gene discovery and transcriptome analysis, Plant Physiol. 134 (2004) 890–897.
[16] C. Jantasuriyarat, M. Gowda, K. Haller, J. Hatfield, G. Lu, E. Stahlberg, B. Zhou, H. Li, H. Kim, Y. Yu, R.A. Dean, R.A. Wing, C. Soderlund, G.L. Wang, Large-scale identification of expressed sequence tags involved in rice and rice blast fungus interaction, Plant Physiol. 138 (2005) 105–115.
[17] C.S. Wang, T.H. Tseng, C.Y. Lin, Rice biotech research at the Taiwan Agricultural Research Institute, Asia Pac. Biotech. 6 (2002) 950–957.
[18] G. An, S. Lee, S.H. Kim, S.R. Kim, Molecular genetics using T-DNA in rice, Plant Cell Physiol. 46 (2005) 14–22.
[19] A. Miyao, K. Tanaka, K. Murata, H. Sawaki, S. Takeda, K. Abe, Y. Shinozuka, K. Onosato, H. Hirochika, Target site specificity of the *Tos17* retrotransposon shows a preference for insertion within genes and against insertion in retrotransposon-rich regions of the genome, Plant Cell 15 (2003) 1771–1780.
[20] C. Sallaud, C. Gay, P. Larmande, M. Bes, P. Piffanelli, B. Piegu, G. Droc, F. Regad, E. Bourgeois, D. Meynard, C. Perin, X. Sabau, A. Ghesquiere, J.C. Glaszmann, M. Delseny, E. Guiderdoni, High throughput T-DNA insertion mutagenesis in rice: a first step towards in silico reverse genetics, Plant J. 39 (2004) 450–464.
[21] C. Wu, X. Li, W. Yuan, G. Chen, A. Kilian, J. Li, C. Xu, X. Li, D.X. Zhou, S. Wang, Q. Zhang, Development of enhancer trap lines for functional analysis of the rice genome, Plant J. 35 (2003) 418–427.
[22] H. Shizuya, B. Birren, U.J. Kim, V. Mancino, T. Slepak, Y. Tachiiri, M.I. Simon, Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector, Proc. Natl. Acad. Sci. U.S.A. 89 (1992) 8794–8797.
[23] H.B. Zhang, S. Choi, S.S. Woo, Z. Li, R.A. Wing, Construction and characterization of two rice bacterial artificial chromosome libraries from the parents of a permanent recombinant inbred mapping population, Mol. Breeding 2 (1996) 11–24.
[24] A.C. Frijters, Z. Zhang, M.V. Damme, G.L. Wang, P.C. Ronald, R.W. Michelmore, Construction of a bacterial artificial chromosome library containing large *Eco*RI and *Hin*dIII genomic fragments of lettuce, Theor. Appl. Genet. 94 (1997) 390–399.
[25] K. Wang, C. Boysen, H. Shizuya, M.I. Simon, L. Hood, Complete nucleotide sequence of two generations of a bacterial artificial chromosome cloning vector, BioTechniques 23 (1997) 992–994.
[26] C.M. Hamilton, A. Frary, C. Lewis, S.D. Tanksley, Stable transfer of intact high molecular weight DNA into plant chromosomes, Proc. Natl. Acad. Sci. U.S.A. 93 (1996) 9975–9979.
[27] S. Choi, D. Begum, H. Koshinsky, D.W. Ow, R.A. Wing, A new approach for the identification and cloning of genes: the pBACwich system using *Cre/lox* site-specific recombination, Nucleic Acids Res. 28 (2000) e19.
[28] S. Qu, G. Coaker, D. Francis, B. Zhour, G.L. Wang, Development of a new transformation-competent artificial chromosome (TAC) vector and construction of tomato and rice TAC libraries, Mol. Breeding 12 (2003) 297–308.
[29] M. Luo, Y.H. Wang, D. Frisch, T. Joobeur, R.A. Wing, R.A. Dean, Melon BAC library construction using improved methods and identification of clones linked to the locus conferring resistance to melon Fusarium Wilt (*Fom-2*), Genome 44 (2001) 154–162.
[30] M.A. Marra, T.A. Kucaba, N.L. Dietrich, E.D. Green, B. Brownstein, R.K. Wilson, K.M. McDonald, L.W. Hillier, J.D. McPherson, R.H. Waterston, High throughput fingerprint analysis of large-insert clones, Genome Res. 7 (1997) 1072–1084.

[31] Z. Cheng, G.G. Presting, C.R. Buell, R.A. Wing, J. Jiang, High-resolution pachytene chromosome mapping of bacterial artificial chromosomes anchored by genetic markers reveals the centromere location and the distribution of genetic recombination along chromosome 10 of rice, Genetics 157 (2001) 1749–1757.
[32] G.L. Wang, T.E. Holsten, W.Y. Song, H.P. Wang, P.C. Ronald, Construction of a rice bacterial artificial chromosome library and identification of clones linked to the *Xa-21* disease resistance locus, Plant J. 7 (1995) 525–533.
[33] G. Hong, Y. Qian, S. Yu, X. Hu, J. Zhu, W. Tao, W. Li, C. Su, H. Zhao, L. Qiu, D. Yu, X. Liu, B. Wu, X. Zhang, W. Zhao, A 120 kilobase resolution contig map of the rice genome, DNA Seq. 7 (1997) 319–335.
[34] S. Nakamura, A. Asakawa, N. Ohmido, K. Fukui, N. Shimizu, S. Kawasaki, Construction of an 800-kb contig in the near centromeric region of the rice blast resistance gene *Pi-ta*^{2} using a highly representative rice BAC, Mol. Gen. Genet. 254 (1997) 611–620.
[35] D. Yang, A. Parco, S. Nandi, P. Subudhi, Y. Zhu, G. Wang, N. Hung, 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 (1997) 1147–1154.
[36] N. Liu, Y. Shan, F.P. Wang, C.G. Xu, K.M. Peng, X.H. Li, Q. Zhang, Identification of an 85-kb DNA fragment containing pms1, a locus for photoperiod-sensitive genic male sterility in rice, Mol. Genet. Genomics 266 (2001) 271–275.
[37] W. Wang, W. Zhai, M. Luo, G. Jiang, X. Chen, X. Li, R.A. Wing, L. Zhu, Chromosome landing at the bacterial blight resistance gene *Xa4* locus using a deep coverage rice BAC library, Mol. Genet. Genomics 265 (2001) 118–125.
[38] Q. Tao, A. Wang, H.B. Zhang, One large-insert plant-transformation-competent BIBAC library and three BAC libraries of Japonica rice for genome research in rice and other grasses, Theor. Appl. Genet. 105 (2002) 1058–1066.
[39] B. Han, Y. Xue, Genome-wide intraspecific DNA-sequence variations in rice, Curr. Opin. Plant Biol. 6 (2003) 134–138.
[40] C.S. Huang, Development of rice variety Tainunng 67, Chin. Agron. Res. 28 (1979) 57–66 (in Chinese).
[41] M. Luo, R.A. Wing, An improved method for plant BAC library construction, in: E. Grotewold (Ed.), Plant Functional Genomics: Methods and Protocols. Methods in Molecular Biology, 236, Human Press, Inc., Totowa, New York, 2003, pp. 3–19.
[42] Z.K. Li, S.R.M. Pinson, J.W. Stansel, W.D. Park, Identification of quantitative trait loci (QTL) for heading date and plant height in cultivated rice (*Oryza sativa* L.), Theor. Appl. Genet. 91 (1995) 374–381.
[43] M.C. Chen, M.C. Cheng, S.G. Chen, Characterization of the promoter of rice plastid *psaA-psaB-rps14* operon and the DNA-specific binding proteins, Plant Cell Physiol. 34 (1993) 577–584.
[44] J. Hiratsuka, H. Shimada, R. Whittier, T. Ishibashi, M. Sakamoto, M. Mori, C. Kondo, Y. Honji, C.R. Sun, B.Y. Meng, The complete sequence of the rice (*Oryza sativa*) chloroplast genome: intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals, Mol. Gen. Genet. 217 (1989) 185–194.
[45] Y.R. Lin, L. Zhu, S. Ren, J. Yang, K.F. Schertz, A.H. Paterson, A *Sorghum propinquum* BAC library, suitable for cloning genes associated with loss-of-function mutations during crop domestication, Mol. Breeding 5 (1999) 511–520.