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*III or *Eco*RI, and two cycles of size selection by pulsed-field gel electrophoresis (PFGE). The *Hin*III 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 Gramineae 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...
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 EcoRI cloning site (pECBAC1: [24]; pIndigoBac536: [25]), transformation (BIBAC: [26]; pBAC-wich: [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 HindIII or EcoRI 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 Teqing, 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 photoperiod 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 HindIII and EcoRI BAC libraries were established by adopting the protocols described Luo and Wing [41]. The high-molecular-weight DNA was prepared from partial digested
either HindIII or EcoRI 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 HindIII or EcoRI fragments were ligated into HindIII- or EcoRI-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 screened by LB containing 12.5 μg/ml of chloramphenicol and X-gal and IPTG. The white colonies were picked robotically.

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 samples were dissolved in 30 μl of TE, and 10 μl of BAC DNA was digested with 10 units of NotI (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 PFGE 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 rice and genomic 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) HindIII BAC clones and filter C contains 8448 (22 × 384) HindIII BAC clones and 9984 (26 × 384) EcoRI 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, putatively free of vector DNA were labeled with [32P] 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 (HindIII and EcoRI) 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 HindIII or EcoRI, none of the HindIII BACs and only 2% of EcoRI 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 HindIII and EcoRI libraries. The average insert sizes of the HindIII and EcoRI BAC libraries were 138.4 and 137.8 kb, respectively (Table 1).

The HindIII 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 NotI 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 NotI restriction sites, were

Table 1

| Characterization of TNG67 HindIII and EcoRI libraries by insert sizing |
|---------------------------|---------------------------|
| BACs | HindIII | EcoRI |
| 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 NotI site | 3.1 | 2.9 |
| Genome coverage | 15.1 | 3.2 |
detected for every HindIII 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 NotI sites ranged from 0 to 9, with an average of 3.1 sites (Table 1; Fig. 2).

The EcoRI 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 NotI sites (Fig. 2). About 2% of EcoRI 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 NotI 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 EcoRI library, since filter C comprises 8448 HindIII clones and 9984 EcoRI 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 HindIII library was 138.4 kb, with 45,312 BACs, and 0% of false positives contributing to the genomic coverage. The chloroplast DNA content in the EcoRI library was 6.2% of the total genome size.

Table 2

<table>
<thead>
<tr>
<th>Probe</th>
<th>Chromosome</th>
<th>No. of BACs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Filter A</td>
<td>Filter B</td>
</tr>
<tr>
<td>Nuclear DNA</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>S13528</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>E0437</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>E30531</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>R03879</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>S20922</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>R2118</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>C11021</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>C0011</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>Average</td>
<td>6.9</td>
<td>6.2</td>
</tr>
<tr>
<td>Chloroplast DNA</td>
<td>995</td>
<td>998</td>
</tr>
</tbody>
</table>

Fig. 1. Size determination of BAC clones randomly selected from the TNG67 EcoRI BAC library. DNA was isolated by alkaline lysis method and digested with NotI. 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 HindIII BAC and EcoRI BAC libraries. DNA was isolated from 384 clones, and 174 clones were randomly picked from HindIII BAC library and EcoRI BAC library, subjectively.
positives accounted for 15.1% of genome coverage. The average insert size of the EcoRI library was 137.8 kb, with 9984 BACs, and 2% of false positives accounted for 3.2% of genome coverage. In total, 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 HindIII library is named: OSJTBa, and the EcoRI 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, HsinChu, 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, HindIII and EcoRI. The genome coverage of HindIII and EcoRI 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 HindIII and EcoRI 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 HindIII BACs and 85% of EcoRI 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 HindIII BACs of TNG67 is false positive under the inspection of 384 clones (Table 1). However, 4 of the 178 EcoRI BACs, or consequently 200 (0.02 × 9984) in the EcoRI 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-covered 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 Shimokiko 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 HindIII and the japonica cv Lemont HindIII BAC libraries, respectively [23], and 20.6, 6.5, and 7.5% for the japonica cv Nipponbare BamHIII, EcoRIII, and HindIII 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 technical assistance. We also appreciate Dr. Wenming Wang, Chris Mueller and Samina Makda for technical assistance.

References


