In vitro Construction of Rice Artificial Chromosome via de novo Assembly

Huang LY1,2*, Wang WS2, Fu Q1, Zhang J1 and Fu BY2

1State Key laboratory for Conservation and Utilization of Bio-Resources in Yunnan, Research Center for Perennial Rice Engineering and Technology of Yunnan, School of Agriculture, Yunnan University, Kunming, Yunnan 650091, China
2Institute of Crop Sciences / National Key Facility for Crop Gene Resources and Genetic Improvement, Chinese Academy of Agricultural Sciences, Beijing 100081, China

Abstract
The assembly of rice artificial chromosomes (RACs) may be useful as a tool to study centromeric function and as a foundation for developing new high-capacity vectors for plant functional genomics and breeding. In this study, we constructed a series of RACs using shuttle vectors including native centromeric segments, origins of replication, selectable marker genes and telomeric repeats. Three main steps were taken to develop an artificial minichromosome. First, the bacterial artificial chromosome (BAC) library of rice Nipponbare genomic DNA (AGI) was screened for centromeric BACs. Second, the Tn5-transposon, which had two end-to-end telomeres flanked with I-CeuI, was constructed using the spectinomycin resistance gene (Sp) as a spacer and a fusional linker between the visualizable marker gene (gfp6) and the selectable marker gene (Hpt). Finally, the Tn5-transposon DNA was separately inserted into the 50 selected centromeric BAC backbones of rice.

Keywords: Rice Artificial Chromosome; Centromere; Telomere; Minichromosome; BAC

Introduction
Transgenic technologies are widely used in multiple fields, especially for resistance to herbicides, insects and viruses, and their entire biosynthetic pathways prove the potential of bioengineering. However, there are many technical challenges associated with these technologies. Some challenges in generating transgenic plant materials include an inability to control the location and nature of the integration of the transgenic DNA into the host genome and linkage of transformed genes to selectable antibiotic resistance genes used in the production of the transgenic cassette. In addition, successive transformations of multiple genes may require several selection genes. The coordinated expression of multiple stacked genes would be required for complex biosynthetic pathways or combined traits. Engineered non-integrating mini chromosomes can overcome these problems and are promising key players in the next generation of transgenic technologies for crop improvement [1,2].

Eukaryotic mini chromosomes can be generated via two approaches. One is the “top-down” approach by the induced truncation or fragmentation of native chromosomes; the other is the “bottom-up” approach by the de novo assembly of an artificial mini chromosome using appropriate sequences.

The “top-down” approach is based on chromosome fragmentation or truncation and can be achieved by irradiation or integration of telomeric repeats. The generation of truncated mini chromosomes has been achieved in humans, mammalian cells and plants [3-15]. The “bottom-up” approach has been used for the assembly of a human artificial chromosome (HAC) by transfection of the cell line HT1080 with a mixture of human centromere-specific alpha satellite, telomeric, and genomic carrier DNA [16] as well as either cloned centromere-specific tandem repeats of alphoid DNA or large centromeric DNA segments [17-23]. Plant mini chromosomes can be generated using techniques that are similar to those used for the construction of human mini chromosomes. Several examples of plant mini chromosomes generated via irradiation,
transposon-mediated breakage, or telomere-mediated truncation have been reported, eg: successful construction of stable artificial minichromosomes using de novo assembly has been reported in maize [24,25].

Although these recent reports are encouraging, the de novo assembly of a rice artificial chromosome has not yet been reported. In this publication, we report our selection strategy for bacterial artificial chromosome (BAC) clones containing centromeric sequences, BAC-based mini chromosome shuttle vector construction, and the formation of rice artificial chromosomes with functional de novo centromeres.

Materials and Methods

Vectors and strains

The pUC19 and pUC57 vectors were stored in our laboratory. The pMDC43 vector was kindly provided by Dr. Fu Yongfu (Institute of Crop Sciences, CAAS). The pH7WGI2D and EZ::TN™ vectors were obtained from Invitrogen (USA) and EpiCentre (USA), respectively.

DH10B and DH5α competent cells were obtained from Invitrogen and TransGen Biotech, respectively.

BAC libraries

Two public genomic BAC libraries from *Oryza sativa* L. japonica. cv. Nipponbare were constructed, as described in Chen, et al. in pBeloBAC11 or pBACIndigo, which were purchased from Arizona Genomics Institute (AGI) [26]. Nipponbare genomic DNA was partially digested with HindIII or EcoRI restriction enzymes. The DNA fragments were size-fractionated in agarose gel and cloned into the pBeloBAC11 Hind III site or the pBACIndigo EcoR I site. The Hind III library consists of 36,864 clones (OSJNba library) with an average insert size of 129 kb, whereas the EcoR I library consists of 55,296 clones (OSJNBb library) with an average insert size of 121 kb. The coverage of the Hind III and EcoR I BAC libraries is estimated to be 10.6 and 15.0 haploid genome equivalents, respectively, thus providing 25-fold redundant coverage when combined.

Probe labelling

RCS1 and RCS2 probes were synthesized and cloned into pUC57 [27]. Probe labelling by DIG-dUTP was performed according to the manufacturer’s instructions (DIG High Prime DNA labelling and Detection Starter Kit II, Roche).

Rice BAC library screening

BAC filter preparations and BAC library screening were performed as described by Wang, et al. [28,29]. Selected BAC plasmids were extracted using a BAC/PAC DNA Isolation Kit (OMEGA, D2156) and analysed by restriction digest fingerprinting as described by Mathewson, et al. [30]. Briefly, BAC DNA was fully digested with Hind III and separated by pulsed-field gel electrophoresis.

The cloning of two telomeres in end-to-end orientation by using Spr as a spacer

TEL fragment [Figure S1] flanked with HindIII, Bgl II and I-Ceu I, BamH1, and EcoRI restriction sites at both ends was synthesized and cloned into pUC57 (pUC57-TEL>) or pUC57-<TEL).

Sp was amplified using pH7WGI2D plasmids as a template by Sp-specific primers: Sp-F: 5'-GTACAAGCTTAGATCTAATTCGGCACGAACCCAG-3'; Sp-R: 5'-GATCGAATTCGGATCCGTCATGCATGATATATCTCCCAAT-3'. Then, the PCR products (1250 bp) and/or pUC19 plasmids were digested with both HindIII and EcoRI restriction enzymes and ligated with T4 ligase (pUC19-Sp).

The pUC57-TEL> plasmids were digested with HindIII, BamH1 and pUC19-Sp plasmids were digested with HindIII, Bgl II restriction enzymes. The TEL> DNA fragments and pUC19-Sp fragments were purified from agarose gel and ligated (pUC19-TEL>-Sp). Next, pUC57-<TEL plasmids were digested with BamH1, Bgl II restriction enzymes and pUC19-TEL>-Sp plasmids were digested with BamH1 restriction enzyme. The <TEL DNA fragments and pUC19-TEL>-Sp segments were purified from agarose gel and ligated (pUC19-TEL>-Sp-<TEL).

Selectable and visible marker constructs

Hpt was amplified using pH7WGI2D plasmids as a template by Hpt-specific primers: Hpt-F: 5'-GGGGACAAGTTTGTACAAA AAAGCAGGCTGCATGAAAAAGCCTGAACTCACCG-3'; Hpt-R: 5'-GGGGACCACTTTGTACAAGAAAGCTGGGTGAGTACTTCTACAGCCATCGG-3'. Then, the PCR products (960 bp) were purified from agarose gel, cloned into the pMDC43 vector (pMDC43-gfp6-Hpt) [Figure 1A] (Mark and Ueli, 2003) via Gateway Technology (Catalogue Numbers 12535-019 and 12535-027) [31].

Tn5-Transposon constructs and Tn5-Transposon DNA

The pMDC43-gfp6: Hpt and EZ::TN™ vectors were digested with HindIII and EcoRI restriction enzymes. The 35S:gfp6-Hpt expression cassette DNA fragments and EZ::TN™ vectors were purified from agarose gel and ligated (Tn5-1).
The pUC19-TEL>-Spr-<TEL and Tn5-1 plasmids were digested with the Hind III restriction enzyme. After purification of the TEL>-Spr-<TEL DNA fragments and Tn5-1 vector segments from agarose gel, Tn5-1 vector segments were dephosphorylated using CIP (Calf Intestinal Alkaline Phosphatase) (TaKaRa, China) and ligated to the TEL>-Spr-<TEL fragments to construct the Tn5-Transposon. Tn5-Transposon plasmids were digested with Ear I restriction enzyme, and Tn5-Transposon DNA fragments were purified from low-melting agarose gel.

Rice minichromosome vector construction

In vitro transposition of the Tn5-Transposon DNA into the centromeric BAC backbones was performed according to the manufacturer's recommendations (EpiCentre). After transformation of DH10B competent cells, the positive RACs were tested by four different pairs of primers [Table S1]: Sp-F/R (the same as above), gfp6-F/R (F: 5’-GATGGTGATGTTAATGGGCAC-3’, R: 5’-GCGCCTTTGTATAGTTCATCC-3’), 35S-F/R (F: 5’-AAGCTTTGGCGTGCCTGC-3’, R: 5’-GATAGTGGATTGTGCGTCAT-3’) and Hpt-F/R (F: 5’-ATGAAAAAGCCTGAACTCACCG-3’, R: 5’-GAGTACTTCTACACAGCCATCGG-3’).

Recombinant retrofitted BAC DNA was digested with a homing restriction enzyme (I-CeuI) converting the circular DNA into a linear minichromosome shuttle vector flanked with telomeric sequences in the correct orientation and removing the fragment comprising Sp<sup>r</sup>.

Results

Rice centromeric BAC clones

To obtain centromeric DNA, 92160 rice BAC clones on 5 high-density filters were screened with RCS1 and RCS2 probes specific to two centromere-specific repeats: CRM and CentC (CentO). According to the Southern hybridizations [Figure 2A and 2B], 50

Figure 1: Screening of the positive centromeric BAC clones

A: One high-density filter (Hybond-N+) after hybridization; B, A: schematic array of the BAC library of rice Nipponbare genome DNA (AGI); C: Analysis of 50 centromeric BAC clones by fingerprinting. Centromeric BACs 1-50

![A: One high-density filter (Hybond-N+) after hybridization; B, A: schematic array of the BAC library of rice Nipponbare genome DNA (AGI); C: Analysis of 50 centromeric BAC clones by fingerprinting. Centromeric BACs 1-50](image)
The rice telomere (TEL) is composed of a repetitive sequence 5'-TTTAGGG-3' [32,33]. Approximately 51 repetitive 5'-TTTAGGG-3' sequences were synthesized in series flanked by homing restriction sites (HRS, I-Ceu I) [Figure 3A,3B and 3F]. The Spr gene was cloned as a spacer for the two telomeres in end-to-end orientations [Figure 3]. Briefly, the forward telomere (TEL>) and reverse telomere (<TEL) were successively inserted at both ends of Spr [Figure 3E].

The pUC19-TEL->Spr<TEL construct has two telomeres in end-to-end orientations using Spr as a spacer

The rice telomere (TEL) is composed of a repetitive sequence 5'-TTTAGGG-3' [32,33]. Approximately 51 repetitive 5'-TTTAGGG-3' sequences were synthesized in series flanked by homing restriction sites (HRS, I-Ceu I) [Figure 3A,3B and 3F]. The Spr gene was cloned as a spacer for the two telomeres in end-to-end orientations [Figure 3]. Briefly, the forward telomere (TEL>) and reverse telomere (<TEL) were successively inserted at both ends of Spr [Figure 3E].

B, Diagram of the first transposon vector named Tn5-1 containing an (A) expression cassette. C, A schematic diagram of the integrated Tn5-Transposon construct. Electrophoretogram of pMDC43-gfp:Hpt and Tn5-1 vectors digested with HindIII/ EcoRI (D) and the Tn5-Transposon construct digested with Ear I (E). 1, pMDC43-gfp: Hpt; 2, Tn5-1; 3, Tn5-Transposon DNA segment (5100 bp); M, DNA ladder.

Figure 3: Construction of the Tn5-Transposon. A, The expression cassette containing a fusion marker between the gene encoding green fluorescent protein (GFP6) and the Hyg B resistance gene (HPT).
The pMDC43-gfp6-Hpt construct containing the gfp6-Hpt fusion gene expression cassette under the control of the CaMV 35S (35S) promoter

To enable both the visualization of a fluorescent marker and chemical selection of minichromosomes, an expression cassette was assembled containing a fusion between the genes encoding green fluorescent protein (GFP, gfp6) and Hygromycin B resistance (Hpt) (Figure 1A). This fusion gene was cloned behind the 35S promoter using nos as the terminator (Figure 1A and 1D).

Tn5-Transposon constructs and Tn5-Transposon DNA

The EPICENTRE EZ::TN™ pMOD™-2<MCS> Transposon Construction Vector system (EpiCentre, Madison, WI, USA) was used to retrofit genes and/or DNA fragments of interest into selected BAC clones.

Based on the EZ::TN™ pMOD™-2<MCS> vector, we constructed a custom transposon, Tn5-Transposon, which consists of gfp6 and Hpt marker genes under the control of the 35S promoter [Figure 1B]. The plasmid also contained two telomeric repeat fragments in an inverted orientation, each of which was composed of recognition sites for the homing restriction enzyme (I-Ceu I) and separated by Sp [Figure 1C and 1E].

Rice minichromosome shuttle vector construction

As shown in Figure 4A, Tn5-Transposon DNA was inserted into the 50 selected centromeric BACs, which were verified by four pairs of specific primers [Figure 4C].

The recombinant retrofitted BAC DNA which was named rice artificial chromosomes (RAC1-50) was digested with I-Ceu I converting the circular DNA into a linear minichromosome shuttle vector flanked with telomeric sequences in the correct orientation and removing the fragment containing Sp [Figure 4B].

Discussion

While several examples of plant mini chromosomes generated via irradiation, transposon-mediated breakage, or telomere-mediated truncation have been reported [13,14], mini chromosomes constructed via de novo assembly are poorly reported. The cause might due to the difficulty in getting repeat sequences from centromere and telomere. Centromeres are composed of long arrays of centromere-specific tandem repeats, with unit lengths from 150 to 180 bp, which are interspersed with multiple copies of centromere-specific retrotransposons, as well as other retro transposable elements found in plant genomes [34-36]. This complexity suggests that the identification of DNA segments responsible for kinetochore formation may be a prerequisite for developing vectors that will undergo de novo assembly of centromeres, and therefore, artificial chromosomes. In this study, RCS1, which can identify centromere CRR (centromere-specific retro transposon of rice) and RCS2 hybrids with centromere-specific tandem repeats (satellite repeat, CentO or CentC), were used to mine functional centromeres. Eukaryotic chromosomes have multiple origins of replication. Although 50 RACs were de novo assembled, some BACs containing different centromeres may only have partial function in the replication of chromosomes. In eukaryotic organisms, multiple origins of replication are located across the chromosomes. Therefore, we believe that most of the BACs selected by RCS1 and RCS2 probe have enough roles in regulating artificial chromosome replication.
The maize mini chromosomes were reported to be engineered by the “bottom-up” approach. Recent reports by Carlson, et al. and Ananiev, et al. described the in vitro construction of vectors containing centromeric sequences and selectable markers, the subsequent introduction of these vectors into maize, and the inheritance of these circular vectors [24,25]. Carlson, et al. introduced circular DNA construct with maize centromeric repeat sequences, a selectable marker gene, a phenotypic marker gene, and filler DNA that functioned in vivo to form autonomous circular chromosomes referred to as “Maize minichromosomes” (MMCs). Ananiev, et al. transformed maize with artificial mini chromosomes assembled with native centromeric segments, origins of replication, selectable marker genes, and telomeric repeats. However, as described in Houben, et al. these results are controversial and require additional experimental confirmation. Therefore, further studies are needed to confirm the feasibility of this approach in plant systems [37].

Financial support

Supported by the National Natural Science Foundation of China (Grant No. 31601274)

Acknowledgement

We thank Dr. WU Chengcang for his instructions.

Supplemental Data

References