Available online at www.sciencedirect.com
ScienceDirect
Genomics 89 (2007) 532–540
GENOMICS
www.elsevier.com/locate/ygeno

Method

An efficient method for producing an indexed, insertional-mutant library in rice

Chengkun He, Moul Dey, Zhihong Lin, Faping Duan, Fengling Li, Ray Wu *

Department of Molecular Biology and Genetics, Cornell University, 316 Biotechnology Building, Ithaca, NY 14853, USA

Received 9 August 2006; accepted 8 November 2006
Available online 16 January 2007

Abstract

Generation of an indexed, saturated, insertional-mutant library is an aid to understanding the functions of genes in an organism. However, 10 years of work by many investigators have not yet yielded such a library in rice. The major reason is that determining the chromosomal locations of a very large number of random insertion mutants by flanking sequence analysis is highly labor intensive, and therefore, libraries that do exist have not been indexed. We report here an efficient procedure to construct an indexed, region-specific, insertional-mutant library of rice. The procedure makes use of efficient long-PCR-based high-throughput indexing, coupled with a random but anchored population of *Ds* transposants. Long-PCR indexing allows rapid and simultaneous determination of the chromosomal locations of a large number of mutants that surround a particular anchor line, thus converting a random library into an indexed one. Such a library can be used directly, without the need to screen a large random library for a desired mutant plant.

Keywords: Rice; Long PCR; Indexing transposants; *Ac/Ds* transposon; Localized transposition; Anchor lines

Rice is arguably the most important cereal crop worldwide. Moreover, rice has become a model plant for molecular studies of monocot crops because of its relatively small genome and the availability of both an efficient genetic transformation method and the DNA sequence of the entire genome. The next major phase of research is to determine the function of each of the approximately 40,000 genes. A key functional-genomics tool is an indexed, saturated, insertional-mutant library in which the chromosomal location of each mutant is known. The maize *Ac/Ds* transposon system has been used to generate a relatively small number of anchor plants from which a large number of secondary insertional-mutant sublines in *Arabidopsis* have been derived [1]. Saturation of the genome by random insertional mutagenesis requires about 220,000 mutants for *Arabidopsis* and 660,000 mutants for rice, to ensure with 99% probability that an insertion occurs every 3 kb [2].

Ac/Ds transposon-based tagging is one of the major approaches used for constructing insertional-mutant libraries of rice. For rice, close to 50,000 *Ac/Ds*-based insertional mutants have been produced [3–11]. One advantage of an *Ac/Ds*-based library is that revertants can be readily obtained and easily identified [12,13], allowing verification of the function of a given gene. The second advantage is that the transposon-based mutagenesis may provide the most effective tool for obtaining gene disruptions and for generating gene traps [14]. In contrast, revertants cannot be readily obtained in direct T-DNA-based insertional mutants or by using the endogenous transposon tos17, even though several large, random mutant populations have been generated [5,15]. The chromosomal locations of the majority of rice insertion sites, regardless of type, are not known. Thus it would be very time consuming to screen even a partial library of 100,000 mutants, searching for mutants that correspond to specific genes of interest. This is because the entire mutant population would need to be screened using one of the three available methods [2,16,17]. The currently available approaches make it impractical to analyze the functions of all 40,000 rice genes.

In reverse genetics as applied to plants, the process of identifying a specific mutant to help determine the function of a gene requires two major steps. The first is to produce a saturated,

* Corresponding author. Fax: +1 607 255 2428.
E-mail address: ray.wu@cornell.edu (R. Wu).

0888-7543/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2006.11.014

random, insertional-mutant library. The second is to screen the random library to identify the desired insertional mutant and, when possible, allelic mutants as well [18,19]. An attractive alternative approach is first to produce an indexed, insertional-mutant library so that the chromosomal location of each insertional mutant is already known. With such a library, steps 1 and 2 are combined. In principle, the flanking sequences of each currently available insertional-mutant library can be determined to produce an indexed library. However, it will require too much effort and time to determine the sequences of the currently available insertional mutants in rice by flanking sequence analysis (FSA). It would be even more difficult to sequence 660,000 insertional mutants, one at a time, when such a population becomes available. Moreover, for rice, only approximately 70% of the samples used for FSA gave sufficient information for matching a sequence in the DNA database [8]. Thus, 30% of the chromosomal locations would not be identifiable.

In this paper, we describe an efficient procedure to construct an indexed, insertional-mutant library such that, in principle, a specific chosen chromosomal region can be truly saturated with insertional mutants. This is possible because the *Ds* elements prefer to transpose to genetically linked sites in rice [7,20]. In our approach, the entire rice chromosomal region is arbitrarily divided into some 1200 anchor regions (launch pads), such that specifically chosen regions can be saturated and studied one at a time. More importantly, we developed an efficient high-throughput method to determine simultaneously the chromosomal location of each mutant in a large population, generated from a specific anchor plant, without the need for sequencing each mutant by FSA. This long-PCR procedure to determine rapidly and simultaneously the chromosomal locations of mutants is a unique aspect of our approach.

Results

Overall design

To produce an indexed, region-specific, saturated, insertional-mutant library with the use of the *Ac/Ds* gene tagging system, the following steps are required, as shown in Fig. 1. First, rice calli are transformed with suitable plasmids to generate transgenic rice plants that harbor the *Ds* element and separately the *Ac* transposase-containing plasmid (*Ac-TPase*). The chromosomal location is determined for each *Ds*-containing plant that harbors a single copy of the T-DNA integration. Those *Ds*-containing plants that are suitably distributed over the rice chromosome, for example, every ca. 400 kb, are considered anchor plants (launch pads) and referred to as Ds plants (Fig. 1, top line). Next, a Ds plant is crossed with an *Ac-TPase* plant to allow transposition to occur. Then, a large number of localized F_2 and F_3 transposants (transposed *Ds* elements) are collected. Once transposition is confirmed, the chromosomal locations of up to 5000 localized transposants are simultaneously determined by combining an efficient long-PCR-based high-throughput procedure and a DNA-pooling strategy. In practice, one procedure is to use information from DNA databases to determine the chromosomal locations of all genes within the 400-kb region and then design primers that correspond to each gene and screen the mutant library to find the specific, desired insertional mutants (Fig. 1, Option A). Alternatively, to make sure that the gene prediction programs do not miss any genes, especially small ones, the entire 400-kb region can be saturated with one insertion every kilobase. In this case, the transposition distance of a specific transposant is estimated by comparing the size of the PCR product with DNA size markers. The chromosomal location of the specific transposant

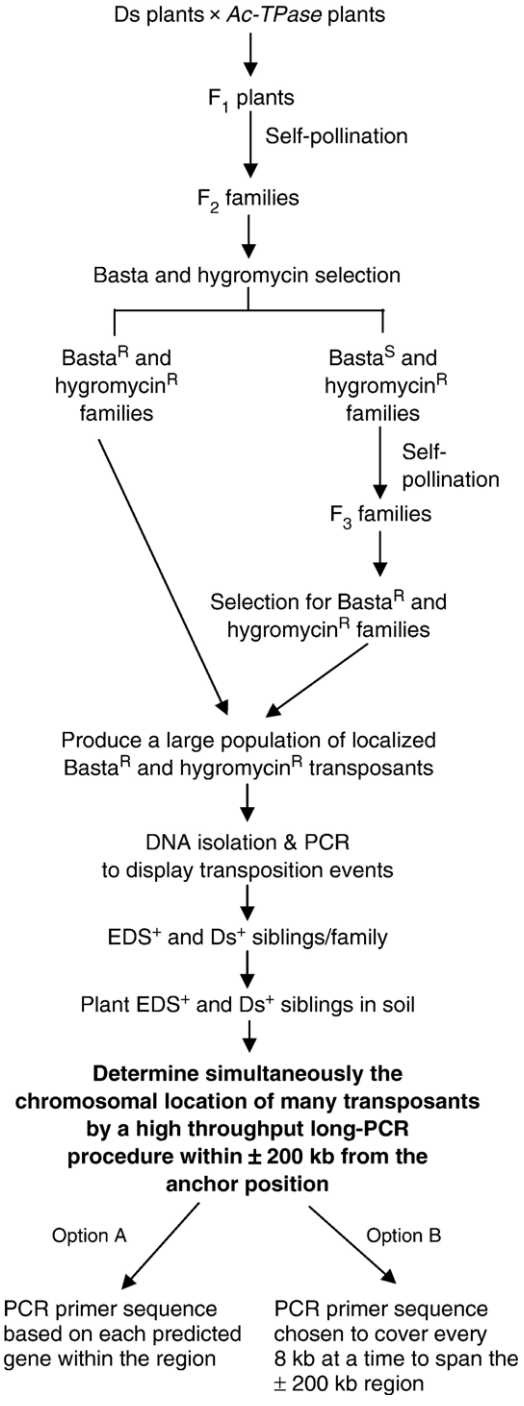

Fig. 1. Flow chart of our protocol to rapidly produce an indexed, insertional-mutant library in rice.