Methodology Study of Preparing High Quality Rice Genomic DNA for Shotgun Sequencing

YANG Ji-Liang[1)], WANG Hui[2)], WANG Qing-Hua[3)],
LI Jun[1)], LIU Bin[2)], JIN De-Min[1)], WANG Bin[1)*]

([1)] Institute of Genetics and Developmental Biology, The Chinese Academy of Sciences, Beijing 100101, China;
[2)] Beijing Genomics Institute, The Chinese Academy of Sciences, Beijing 100130, China;
[3)] School of Life Science, Shandong University, Jinan 250100, China)

Abstract The contamination of chloroplast and mitochondrial DNAs was a serious problem during genome sequencing of rice (*Oryza sativa* L. ssp. *indica*) by whole-genome shotgun strategy. Pulse field gel electrophoresis (PFGE) was utilized to purify rice genomic DNA, which could efficiently remove the organelles (chloroplast and mitochondrion) DNA and reduce the contamination ratio from 3% to 0.2%. At the same time, the rice DNAs yielded from yellow seedlings and green seedlings were compared, and the differences between HB method and NIB method in high molecular weight(HMW) DNA isolation were also studied. Finally, a set of methods for obtaining the whole and highly pure rice genomic DNAs were proposed, which included culturing rice yellow seedling, isolating, embedding and lysising rice nuclei, purifying and recovering rice genomic DNAs in low melting point (LMP) agarose gel by PFGE. Ultrasonic treatment on HMW DNAs in the melting LMP gel at 38℃ was reported at first time, it facilitated to obtain the desired DNA fragments for construction of shotgun library and gradient libraries.

Key words rice, genomic DNA, shotgun sequencing

The recent improvements in automated DNA sequencing have made whole-genome shotgun sequencing (WGS) an attractive approach for gene discovery in both small and large genomes[1, 2]. Yu *et al*[3] and Stephen *et al*[4] separately produced a draft sequence of the rice genome by WGS using different rice species. The major steps of WGS are as follows: rice genomic DNAs are sheared or restricted to yield random fragments of the desired size; the fragments are cloned into a universal vector; and the shotgun libraries (1.5~3 kb inserts) are constructed; sequencing reactions are performed with a pair of universal primers on randomly selected clones in the shotgun libraries; when the reads coverage is at 2X, these sequencing reads are assembled into contigs; subsequently the sequencing information gotten from gradient libraries is used to fill and span the different length of repeat gap, and construct desired scaffold (full sequenced molecule). One of the main limitations in the WGS approach is the generation of fully representative small insert libraries, so the preparation of high quality genomic DNAs is critical. When the normal method is used to prepare rice genomic DNAs for shotgun sequencing, two major problems are encountered. Firstly, the frequent existence of the serious contamination of organelle (chloroplast and mitochondrion) DNAs, which causes a lot of troubles in late sequencing. Secondly, how to obtain efficiently the needed large quantity (about 2 mg) of high quality rice genomic DNAs. In order to solve the above limitations and problems, many comparative experiments were done and some results were reported in this paper.

1 Material and methods

1.1 Seedling culture of rice

1 000 rice seeds were washed twice with distilled water, soaked in distilled water at 37℃ for 24 h, and then arranged evenly on the tela surface in a tray. The seeds were cultured at 30℃ for about 7 days. When the length of seedlings was between 8 ~ 10 cm, the seedlings were cut, grinded with liquid nitrogen or stored at −70℃ for late use.

The difference between the culture of yellow seedlings and green seedlings was that the former was done under the complete dark condition during the whole period of culture.

1.2 Extraction of rice genomic DNAs

Two different methods used to extract rice genomic DNAs were compared, one was homogenization buffer (HB) method modified from Zhang *et al*[5], the other was nuclei isolation buffer (NIB) method modified from "Molecular Cloning"[6]. The major reagents used in these two methods are listed in Table 1.

* Corresponding author
Tel: 86-10-64889378, E-mail: bwang@genetics.ac.cn
Received: February 26, 2003 Accepted: May 28, 2003

Table 1 The buffer components of HB and NIB

Reagents		Component (HB)	Component (NIB)
Homogenization buffer	Tris	10 mmol/L	10 mmol/L
	KCl	80 mmol/L	100 mmol/L
	EDTA	10 mmol/L	10 mmol/L
	Spermidine	1 mmol/L	4 mmol/L
	Spermine	1 mmol/L	1 mmol/L
	Sucrose	0.5 mol/L	0.5 mol/L
	β-Mercaptoethanol	0.15%	0.1%
Washing buffer		Adjust pH to 9.4~9.5 with NaOH The above buffer plus 0.2% Triton X-100	
Lysis buffer	EDTA, pH 9.0~9.3	0.5 mol/L	The NIB buffer without β-mercaptoethanol
	Sodium lauryl sarcosine	1%	2.5%
	Proteinase K	0.5~1 g/L	1 g/L

1.3 Purification and recovery of rice genomic HMW DNAs by pulse field gel electrophoresis

1.3.1 Preparation of agarose gel for pulse field gel electrophoresis(PFGE): 150 ml of 1% agarose gel was dissolved in 0.5 × TBE, cooled to 60℃ and poured into gel mold (Bio-Rad) of 21 cm × 12.7 cm. A comb sealed with plastic stick was put into the sample position to form a long slot-well, the gel was solidified by nature. A trough of 20 cm × 2 cm size was made on the solidified gel by removing some gel with a scalpel, which was just 0.5 cm from the front edge of sample well. At the same time, 40 ml of 1% low melting point (LMP) agarose gel were prepared with 0.5 × TBE and poured into the prepared trough.

1.3.2 Purification of rice genomic HMW DNAs by PFGE: CHEF Mapper (Bio-Rad) was used for PFGE. The plugs containing rice genomic high molecular weight(HMW) DNAs were permeated into 0.5 × TBE for 1 h before electrophoresis, then the plugs were laid in the sample slot-well side by side and sealed with LMP agarose gel. Thus it was ready for PFGE. The conditions of PFGE were showed in Table 2, the total time was 24 h.

Table 2 Conditions of PFGE for purifying and recovering rice genomic HMW DNAs

		Voltage gradient/(V·cm^{-1})	Included angle/(°)	Initial switch time Final switch time/s	Run time /h	Other conditions
Block 1	State 1	6	60	60	15	0.5 × TBE 1% gel 14℃
	State 2		−60	60		
Block 2	State 1	6	60	90	9	
	State 2		−60	90		

1.3.3 Recovery of rice genomic HMW DNAs in LMP gel: When PFGE was finished, 1cm gel was excised at both edges and stained in 0.5 mg/L EB for 20~30 min. The exact position (the band width) of rice HMW DNAs in the gel was determined with a fluorescence ruler under the UV lamp, and then the remaining unstained gel containing HMW DNAs was excised from the corresponding position determined just before and collected in 1.5 ml tubes.

1.4 Ultrasonic treatment on rice genomic DNAs

1.4.1 In the LMP agarose: About 600 mg LMP gels containing rice HMW DNAs were put in each 1.5 ml tube, melted at 65℃ for 10 min, then transferred to a 38℃ water-bath for ultrasonic treatment. The tubes containing LMP gel with HMW DNAs were treated by specially designed ultrasonic equipment under number 1 power, one by one, and immediately put back to 38℃ water bath.

1.4.2 In TE solution: The rice genomic DNAs recovered by NIB method was dissolved in TE, and sonicated under number 1 power using specially designed ultrasonic equipment.

1.4.3 Checking the size of rice DNAs size after ultrasonic treatment: 4 μl liquid was taken from each eppendorf tube and mixed evenly with 1 μl of 10 × loading buffer separately, their size range was checked by common electrophoresis.

2 Results and discussion

2.1 Comparison of rice HMW DNAs obtained from yellow seedlings and green seedlings

Comparison of HMW DNAs obtained from yellow seedlings and green seedlings by the HB method was showed in Figure 1. It was clear that by PFGE purification, the content of HMW DNAs from yellow seedlings was 4 times higher than that from green seedlings. Rice yellow seedlings contain less chloroplast, only one centrifugation is needed, thus reduced the chance of losing and destroying the nuclei, so the recovery ratio of rice HMW DNA was high. By comparison, in order to remove organelle DNAs in the nuclei of rice green seedlings, three times lavation and three times centrifugation were needed, nuclei losses concurred during these processes.

2.2 Comparison between HB and NIB method for preparation of rice HMW DNAs

In isolating the rice nuclei, the reagents and the procedures were almost the same in HB and NIB method. The main difference between these two methods was as follows, the nuclei was embedded in LMP agarose gel in HB method, while the nuclei was suspended in liquid in NIB method, which led to the different operations of nuclei lysis, DNA extraction and purification in the following procedures. Therefore, the very different results were gotten for recovering rice HMW DNAs, as shown in Table 3.

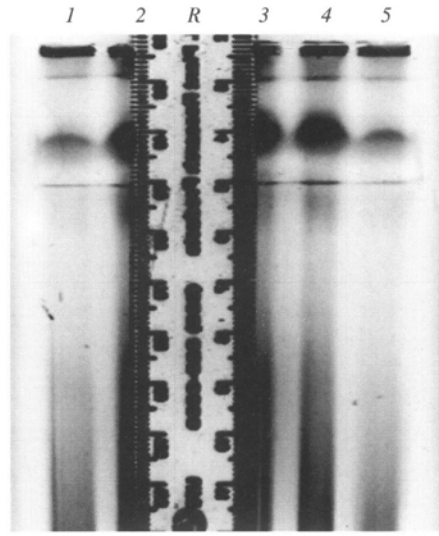

Fig. 1 Comparison of HMW DNAs obtained from yellow seedlings and green seedlings after purification of PFGF

1, *5*: rice green seedlings; *2*, *3*, *4*: rice yellow seedlings. Each sample is the same plug size that represents DNAs content from 20 g original rice seedlings. *R* is a fluorescence ruler. The arrow points the HMW DNAs bands.

Table 3 Comparison of HB and NIB method for preparation of rice HMW DNAs

	HB	NIB
Centrifugation condition for nuclei isolation	1 900 *g*, 4℃, 20 min	3 000 *g*, 4℃, 12 min
States of nuclei	Embedded in LMP agarose gel	Suspended in liquid
Condition of nuclei lysis reaction	50℃, 48 h, in lysis buffer	37℃, 1 h, in lysis buffer
Method of DNA recovery	LMP agarose gel pulse field gel electrophoresis	phenol/chloroform/isopentanol isopropanol centrifugation
Quality of DNA	Large DNA fragments (more than 1 000 kb); less than 0.2% chloroplast and mitochondrion DNAs; easy for ultrosonic treatment	30~50 kb DNA fragments; more than 3% contamination of chloroplast and mitochondrion DNAs
Quantity of DNA	100~200 μg DNA/20~40 g material	100 μg DNA/100 g material

2.2.1 Technique procedures: Technique procedures of NIB method included many steps under flowing condition, such as isolation, suspension and lysis of rice nuclei; purification of rice genomic DNAs by several centrifugations, which made the rice genomic DNA loss or destroy. So large quantities of rice samples were needed for DNA preparation by this method, 100 g per batch at least.

However, the rice genomic DNAs were treated mainly in LMP gel in the technique procedures of HB method, such as lysising rice nuclei, purifying and recovering HMW DNAs by PFGE. The DNAs released from nuclei could be protected by LMP gel, thus the DNAs sheared by foreign force were avoided, and the degradation and loss of the DNAs were greatly reduced.

2.2.2 Centrifugation conditions for nuclei isolation: To isolate rice nuclei, centrifugation was used in both HB and NIB methods, but their centrifugation conditions were different. In HB method, it was at 1 900 *g*, 4℃ for 20 min, while in NIB method, it was at 3 000 *g*, 4℃ for 12 min. HB method facilitated to isolate nuclei

from organelle DNAs, while NIB method used greater centrifugation force, was easy to cause organelle (chloroplast and mitochondrion) DNA contamination.

2.2.3 Quality and quantity of prepared HMW DNAs: The sizes of recovered DNA fragments were different between the two methods, generally, larger than 1 000 kb by HB method, while only 30~50 kb by NIB method. The large DNA fragments were good for getting the desired fragments sizes of 1.5~3 kb and 3~12 kb by the ultrasonic treatment, making the DNA fragments more representative.

According to the sequencing result, the organelle DNA contamination was less than 0.2% in HB method, while more than 3% in NIB method.

For shotgun sequencing 2 mg of rice DNAs were required, therefore, the DNA recovery rate should be considered. To obtain 100 μg of rice DNAs, 20 g of rice yellow seedling samples were needed when using HB method, while 100 g samples are needed when using NIB method. So the DNA recovery rate of HB method is 5 times greater than that of the NIB method.

2.3 PFGE is an efficient technique to remove the organelle DNA contamination

When preparing rice HMW DNAs by NIB method, the lavation with washing buffer and differential centrifugation could remove part of mitochondrial DNAs and chloroplast DNAs, but the contamination of organelle DNAs was still serious. The 3% organelle DNAs not only hindered the sequencing speed, but also was time-consuming and costly. In order to solve this tough problem, PFGE technique was applied to purify and recover rice nuclei DNAs, the result was encouraging, only less than 0.2% organelle DNA contamination was found during DNA sequencing. The reason was that the DNA fragments with sizes less than 1Mb were removed (Figure 1) at the PFGE condition of 6 V/cm, 120°, 60 s switch time for 15 h, and 90 s switch time for 9 h (Table 1), which included the rice chloroplast DNAs (135 kb) and mitochondrial DNAs (492 kb) (http://www.esb.utexas.edu/bot350m/lecnotes/MITOCHONDRIA).

2.4 Difference between ultasonic treatment in TE buffer and in LMP gel

Generally, the ultrasonic treatment on genomic DNAs was done in TE, while the HMW DNAs obtained by HB method was done in LMP gel. At first, agarase was utilized to digest the LMP gel to release the rice DNAs, which not only required operation steps, but also increased cost. In order to reduce cost and save time, the melting LMP gel containing rice HMW DNAs was directly sonicated at 38℃. The result showed that the melting LMP gel could protect rice DNAs from ultrasonic treatment. To obtain 1.5~3 kb DNA, 10~13 s treatment was needed in LMP gel, while only 1~2 s treatment was needed in TE. It was clear that the ultrasonic treatment in TE was very short, which was difficult to operate to get DNA fragments larger than 5 kb in size. Therefore, ultrasonic treatment on rice DNAs in melting LMP gel facilitated to obtain the larger DNA insert fragments, such as 3~5 kb, 5~7 kb, 7~9 kb and 9~11 kb gradient size ranges, which were used to construct the gradient libraries.

2.5 Development of plant DNA extraction methods

DNA extraction is the basic and first step to study plant genome and molecular characteristics at DNA level. A large number of DNA extraction protocols can be found in the literature, many of these are slight variation of other procedures [7~9]. The widely used traditional plant DNA preparation methods are CTAB and SDS. As the development of marker assistant selection in plant breeding, the simple, rapid and high throughput methods to extract plant DNAs are required [10~13]. On the other hand, specific and consistent detection is required for accurate identification of a particular genetic modification in a plant sample, especially in the transgenic plant in which the modification is present in trace amounts [14]. So, the different requirements for plant DNA lead to the different plant DNA extraction methods.

The isolated plant nuclei were embedded in LMP plugs to prepare plant HMW DNAs, and PFGE was used to isolate large DNA fragments during the construction of bacterial artificial chromosome (BAC) libraries [5]. The improved method was specially designed from these references to meet our goal of preparing high quality rice genomic DNAs for shotgun sequencing. We mainly concerned the following three aspects: the contamination of organelle (chloroplast and mitochondrion) DNAs was as low as possible, the amount of DNA was higher than 1 mg, and the total DNAs should represent the whole genome. The result showed that our purpose was realized.

3 Conclusion

Through above comparative study, the following conclusions were drawn: a. Using rice yellow seedlings could get high yield of rice HMW DNAs with less contamination of chloroplast DNAs. b. HB method largely reduced the ratio of damaged, degraded and lost rice genomic DNAs, due to the protection of LMP agarose. c. PFGE was an efficient technique to purify and recover rice HMW DNAs, especially to reduce the contamination of chloroplast and mitochondrial DNAs. d. Ultrasonic treatment on rice HMW DNAs in LMP facilitated to get the desired DNA fragments.

To combine these techniques together could produce enough DNAs with high quality in relatively short time. In establishment of the draft sequence of *Oryza sativa* L. ssp. *Indica*, a quarter of rice genomic

DNAs was prepared by the present method

Acknowledgments We are grateful to Dr Jie Shang for critical reading of the manuscript

Reference

1 Adams M D, Celniker S E, Holt R A, *et al*. The genome sequence of *Drosophila melanogaster*. Science, 2000, **287** (5461): 2185~2195

2 Venter J C, Adams M D, Myers E W, *et al*. The sequence of the human genome. Science, 2001, **291** (5507): 1304~1351

3 Yu J, Hu S N, Wang J, *et al*. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science, 2002, **296** (5565): 79~92

4 Goff S A, Ricke D, Lan T H, *et al*. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). Science, 2002, **296** (5565): 92~100

5 Zhang H B, Zhao X, Ding X, *et al*. Preparation of megabase-size DNA from plant nuclei. Plant J, 1995, **7** (1): 175~184

6 Sambrook J, Russell D W. Molecular Cloning. 3rd. New York: Cold Spring Harbor Laboratory Press, 2001. 408~483

7 Graham G, Henry R. Preparation of fungal genomic DNA for PCR and RAPD analysis in fingerprinting methods based upon arbitrarily primed PCR. In: Micheli M R, Bova R eds. Springer Laboratory Manual. 29~34. Springer Heidelberg. 29~34

8 Gang D R, Weber D J. Preparation of genomic DNA for RAPD analysis from thick-walled dormant teliospores of *Tilletia* species. Biotechniques, 1995, **19** (1): 92, 94, 96, 97

9 罗志勇, 周 钢, 陈湘晖, 等. 高质量植物基因组 DNA的分离. 湖南医科大学学报, 2001, **26** (2): 178~180
Luo Z Y, Zhou G, Chen X H, *et al*. Bulletin of Hunan Medical University, 2001, **26** (2): 178~180

10 Dilworth E, Frey J E. A rapid method for high throughput DNA extraction from plant material for PCR amplification. Plant Molecular Biology Reporter, 2000, **18** (1): 61~64

11 Paris M, Carter M. Cereal DNA: A rapid high-throughput extraction method for marker assisted selection. Plant Molecular Biology Reporter, 2000, **18** (4): 357~360

12 Ikeda N, Bautista N S, Yamada T, *et al*. Ultra-simple DNA extraction method for marker-assisted selection using microsatellite markers in rice. Plant Molecular Biology Reporter, 2001, **19** (1): 27~32

13 Burr K, Harper R, Linacer A. One-step isolation of plant DNA suitable for PCR amplification. Plant Molecular Biology Reporter, 2001, **19** (4): 367~371

14 Holden M J, Blasic J R Jr, Bussjaeger L, *et al*. Evalution of extraction methodologies for corn kernel (*Zea mays*) DNA for detection of trace amounts biotechnology-derived DNA. J Agric Food Chem. 2003, **51** (9): 2468~2474

Shotgun法测序制备高纯度水稻基因组 DNA的方法探讨

杨继良[1)] 王 辉[2)] 王庆华[3)] 李 军[1)] 刘 斌[2)] 金德敏[1)] 王 斌[1)*]

([1)]中国科学院遗传与发育生物学研究所, 北京 100101; [2)]中国科学院北京基因组学研究所, 北京 101300;
[3)]山东大学生命科学院, 济南 250100)

摘要 在用散弹 (shotgun) 法测定水稻 (*Oryza sativa* L. ssp. *indica*) 基因组全序列的过程中, 叶绿体和线粒体 DNA的污染问题非常严峻. 应用脉冲场电泳 (PFGE) 技术对水稻基因组 DNA进行纯化, 结果表明它能够有效去除叶绿体和线粒体 DNA, 使其污染率从 3%降低到 0.2%. 同时, 比较了水稻绿苗和黄化苗的 DNA得率, 以及 HB法和 NIB法制备大分子质量 (HMW) DNA的异同. 最后提出一套制备水稻基因组 DNA的方法, 包括黄化苗培养; 细胞核的分离、包埋和裂解; 脉冲场电泳纯化、回收聚集在低熔点 (LMP) 胶中的水稻 HMW DNA. 用该方法所得的水稻基因组 DNA, 纯度高 (无叶绿体和线粒体 DNA污染)、基因组完整 (机械剪切和降解少)、回收率高 (操作过程中 DNA损失少). 另外, 首次报道在融化的低熔点 (LMP) 胶中对水稻 HMW DNA于 38℃进行超声波处理, 能够获得用于 shotgun文库和梯度文库构建所需要的各种 DNA片段 (1.5~3 kb, 3~12 kb), 并且效果优于在 TE中进行操作.

关键词 Shotgun测序, 水稻, 基因组 DNA

学科分类号 Q94

* 通讯联系人. Tel: 010-64889378, E-mail: bwang@genetics.ac.cn

收稿日期: 2003-02-26, 接受日期: 2003-05-28