



OPEN

SUBJECT AREAS:

DNA
CHROMOSOMES
NUCLEIC ACIDS
GENOMEReceived
25 April 2013Accepted
25 November 2013Published
10 December 2013Correspondence and
requests for materials
should be addressed to
K.L.H. (Huangkl009@
sina.com) or Y.B.L.
(lyb@cau.edu.cn)* These authors
contributed equally to
this work.

Randomly broken fragment PCR with 5' end-directed adaptor for genome walking

Wentao Xu^{1,2*}, Ying Shang^{1*}, Pengyu Zhu¹, Zhifang Zhai¹, Jing He², Kunlun Huang^{1,2} & Yunbo Luo^{1,2}¹Laboratory of Food Safety, College of Food Science and Nutritional Engineering, China Agricultural University, Beijing 100083, China, ²The Supervision, Inspection & Testing Center of Genetically Modified Food Safety, Ministry of Agriculture, Beijing 100083, China.

Many genome walking methods based on polymerase chain reaction (PCR) are available, including those with and without restriction enzyme modification. Nevertheless, these methods suffer from low reproducibility, inefficiency, and non-specificity. Here, we present a traceable and efficient PCR strategy: randomly broken fragment PCR with 5' end-directed adaptor for genome walking. The genome is first fragmented randomly. After blunting ends, the fragments are ligated to the 5' end-directed adaptors. Semi-nested PCR is then performed. Thus, we can obtain an unknown sequence by cloning the fragments of interest, followed by sequencing. This method effectively bypasses the above-mentioned obstacles and offers the advances: 1) genome fragmentation without using restriction enzymes; 2) enhancement of primer specificity and the prevention of self-ligation between the adaptors by employing a 5' end-directed adaptor. All of the steps in this new method are straightforward, and the unknown sequence can be definitively obtained by merely applying the method once.

Genome walking (GW) is a basic molecular biology technique that involves a stepwise determination of an uncharacterized DNA sequence flanking a region with a known sequence¹. This technique has many applications, such as: 1) identifying insertion sites of transposons or T-DNA and exogenous genes in transgenic technology; 2) separating promoters and regulatory elements; 3) overlapping artificial chromosomes (PAC, YAC, BAC); 4) constructing map-based cloning of overlaps; and 5) discovering new genes related to diseases^{2–6}.

Current GW methods encompass two categories. The first category includes Inverse PCR (IPCR) and Cassette PCR, which require the use of restriction enzymes to fragment the genome. IPCR is the earliest flanking cloning technique based on conventional PCR⁷, whose rate of success and reproducibility are low because they depend on the efficiency of the endonucleases used and the rate of self-ligation. Cassette PCR techniques include vectorette PCR, single-specific-primer PCR, template-blocking PCR, RCA-GIP, straight-walk ligation-mediated genome walking, A-T linker PCR⁸, loop-linker PCR⁹, and others¹⁰. Most of the above-mentioned methods focus on enhancing the specificity of PCR, for instance, by retaining the ligation of the adaptors, de-phosphorylating the 5' end, or using a ddNTP to block adaptor self-ligation, or using an amino group to block the extension of the adaptor. However, these methods are ineffective if the genome is large or does not have appropriate restriction endonuclease sites. Moreover, these methods are dependent upon time-consuming and labor-intensive enzymatic modifications of the target DNA.

The second category of GW methods does not involve restriction enzymes to digest the genome first, and includes primer-based and extension-based methods¹¹, such as site-finding PCR, TAIL-PCR, semi-random primer PCR, and two-step gene walking PCR. However, the generation of non-specific DNA products is the main problem in this category, particularly in TAIL-PCR, and the specific product is seldom obtained after gel electrophoresis. Poor reproducibility is another problem due to the randomness of the primers utilized.

To circumvent the above-mentioned drawbacks, we developed a novel method, termed randomly broken fragment PCR with 5' end-directed adaptor for genome walking (RBF-PCR). RBF-PCR offers the following advantages: 1) the genome is randomly fragmented without the dependence on restriction enzymes; 2) a pair of partially reverse complemented adaptors can lead to T/A-specific ligation to the ends of the fragments, and the self-ligation between the adaptors is effectively avoided by a T overhang at the 3' end; 3) PCR can be conducted by combining different adaptor primers with specific primers; 4) the unknown sequence is efficiently obtained by

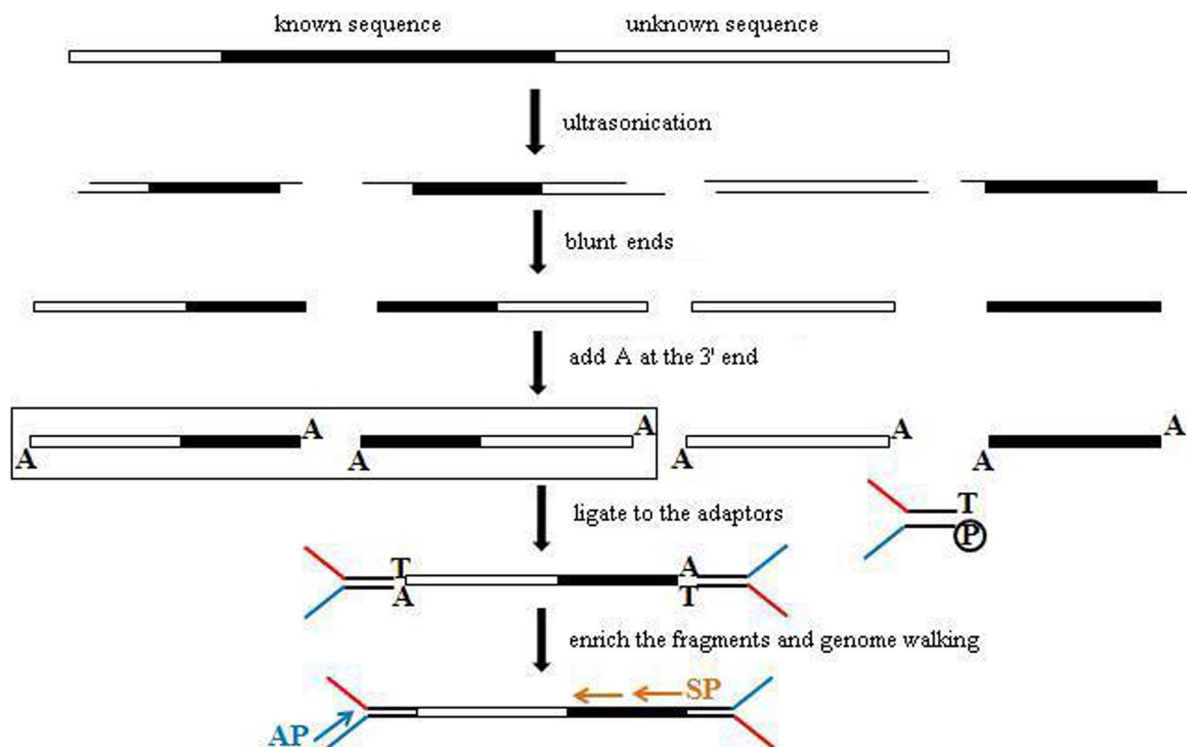


Figure 1 | The schematic diagram of RBF-PCR. The fragments in the box are the targets of interest that contain both known and unknown sequences.

cloning the PCR products of interest into vectors, followed by sequencing. Experimental results in genetically modified maize LY038 (GM LY038) showed that RBF-PCR is suitable for genome walking and can be applied on a large scale.

Results

Principle of RBF-PCR. The principle of RBF-PCR is illustrated in Figure 1. First, the genome is randomly fragmented using an ultrasonic processor. A series of additional steps were needed to obtain the fragments with blunt-ends. A T is overhung at the 3' end of the adaptor to avoid adaptor self-ligation. Therefore, after ends blunting, adenines are added to the 3' ends of the broken fragments for the ligation with adaptors. Semi-nested PCR is then conducted to obtain the flanking sequence of the known sequence. The semi-nested primers are designed to anneal to the region of the known sequence, and we named them the specific primers (SP); the other fixed primer is the adaptor primer (AP).

The first few steps are similar to the library preparation in some next-generation sequencing methods. However, to ensure that these steps were fit for identifying the flanking sequences, many simulation experiments and real-time PCR analyses were employed to test the feasibility and to optimize the conditions.

Fragmentation of genomic DNA. We employed ultrasonic probes with three different diameters to fragment the genomic DNA. As shown in Figure 2, when the diameter of the probe was 6 mm, the power of the ultrasound was easy to control, and the genomic DNA was uniformly fragmented into the 250–2000 bp range by using the following ultrasonic parameters: time, 1 s; interval, 10 s; repeat, 30 times. Meanwhile, the PCR amplification was performed with high fidelity when the length of the template was in the range of 250–2000 bp. In contrast, the DNA was not efficiently fragmented using the 2 mm diameter probe. However, increasing the length of the ultrasonication time helped splash the DNA liquid. The genome was severely fragmented using the 3 mm diameter probe, making later modifications difficult and resulting in a short genome walking sequence. Therefore, we chose the 6 mm diameter probe.

Meanwhile, we observed that by changing the length of the ultrasonic time, the degree of genomic DNA fragmentation can be controlled. Additionally, we found that the choice of probe diameter was not relevant to the liquid volume used in our research; we merely aimed to obtain the best fragmentation.

Adaptor primer choice. Due to the adaptors (Supplementary Table S2) being partially reverse complemented, there would be four kinds of amplification styles (Supplementary Figure S1), and two possible amplification products can be obtained for each single specific primer. We used real-time PCR to test which pair of primers could yield satisfactory amplification results. An A was overhung at the 3' end of the amplicon after PCR by the *rTaq* enzyme. It is precisely because the characteristic of *rTaq* enzyme that we could regard the PCR products as the genome fragments which have blunted ends and added A at the 3' ends. Therefore, we directly ligated the PCR

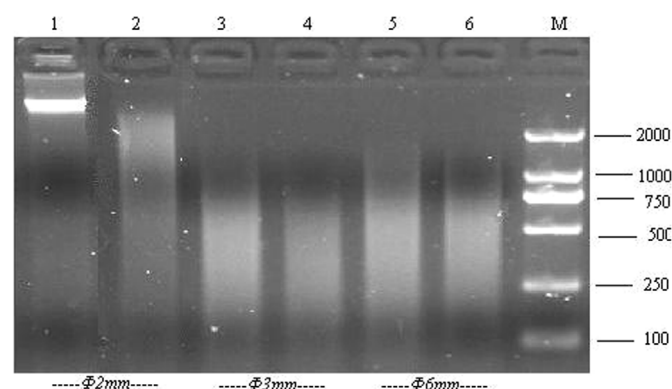


Figure 2 | The ultrasonication results of the GM LY038 genome with different probes. The diameters of the ultrasonic probes are noted under the figure. Each ultrasonic time pulse was repeated 30 times; the ultrasonic time for lanes 1, 3, and 5 was 1 s and lanes 2, 4, and 6 was 2 s. M, the marker DL2000.



products with the adaptors instead of the actual fragmented genomic DNA, and the ligation products will be tested as the template in real-time PCR. The direct use of PCR products rather than fragmented genomic DNA was time saving. In this study, an amplicon of 673 bp was chosen for the real-time PCR.

The 673 bp amplicon was amplified by SP 673-F/R, a segment of DNA located in the exogenous gene region of the GM LY038 genome; the detailed location is shown in Supplementary Figure S2 B. The 673 bp amplicons were ligated to the adaptors, and the ligation products were then 10-fold diluted for use as the template in real-time PCR. SP inner-F with AP1 or AP2 (Supplementary Figure S2 C) were taken as the primers for real-time PCR. As shown in Figure 3, the Ct values of AP1 and AP2 were comparable when a high concentration of template was used. However, as the concentration of the template decreased, primer dimers tended to appear in the reaction when AP2 was used, which may be attributed to the greater binding and amplification capacities of AP1 versus AP2. In addition, due to the sequence of AP2 was reverse complemented with that of Adaptor-2, AP2 could also bind to other non-specific ligation products during the amplification. As shown by the Ct and T_m values, the SP along with AP1 yielded a better amplification result than AP2, furthermore, the specificity of this method can be enhanced when we use AP1 with SP.

The specificity and detection limit of RBF-PCR. GM LY038 and its wild type were used to test the specificity of RBF-PCR. SP 673-F and

SP inner-F were chosen as the nested primers, two technical replicates were analyzed for each template. As the genome of wild type did not contain the exogenous gene, no amplicons should be obtained after the semi-nested PCR amplification because there are no binding sites for SP. As shown in Figure 4, dispersed bands (lanes 1 and 2) were observed after 2% agarose gel electrophoresis when the randomly broken GM LY038 genome was used as the template; and only primer dimers (lanes 3 and 4) were observed when the wild type genome was used as the template.

Two PCR products, fragments 673 bp and 218 bp with no identical sequence (Supplementary figure S2 B), generated by SP 673-F/R and SP 218-F/R using the GM LY038 genome as the template, were used to test the detection limit of RBF-PCR via real-time PCR. After common PCR, we obtained amplicons of 673 bp and 218 bp; these amplicons were ligated to adaptors separately, and the copy number of the ligation product was measured. We mixed these ligation products in different ratios to imitate the condition of DNA fragments after ultrasonication. The ligation products of 673 bp were regarded as the target templates and the 218 bp as the non-target templates. We only decreased the concentration of the 673 bp fragment until no amplicons were obtained. The primers used in real-time PCR were AP1 and SP inner-F, potentially amplifying a segment of 189 bp from the 673 bp fragment, which was nearly the same length as that of the 218 bp amplicon.

As shown in Table 1, when the ratio of 673 bp to 218 bp products was decreased to 0.0001 : 1 (a copy number equivalent of 6 : 6000),

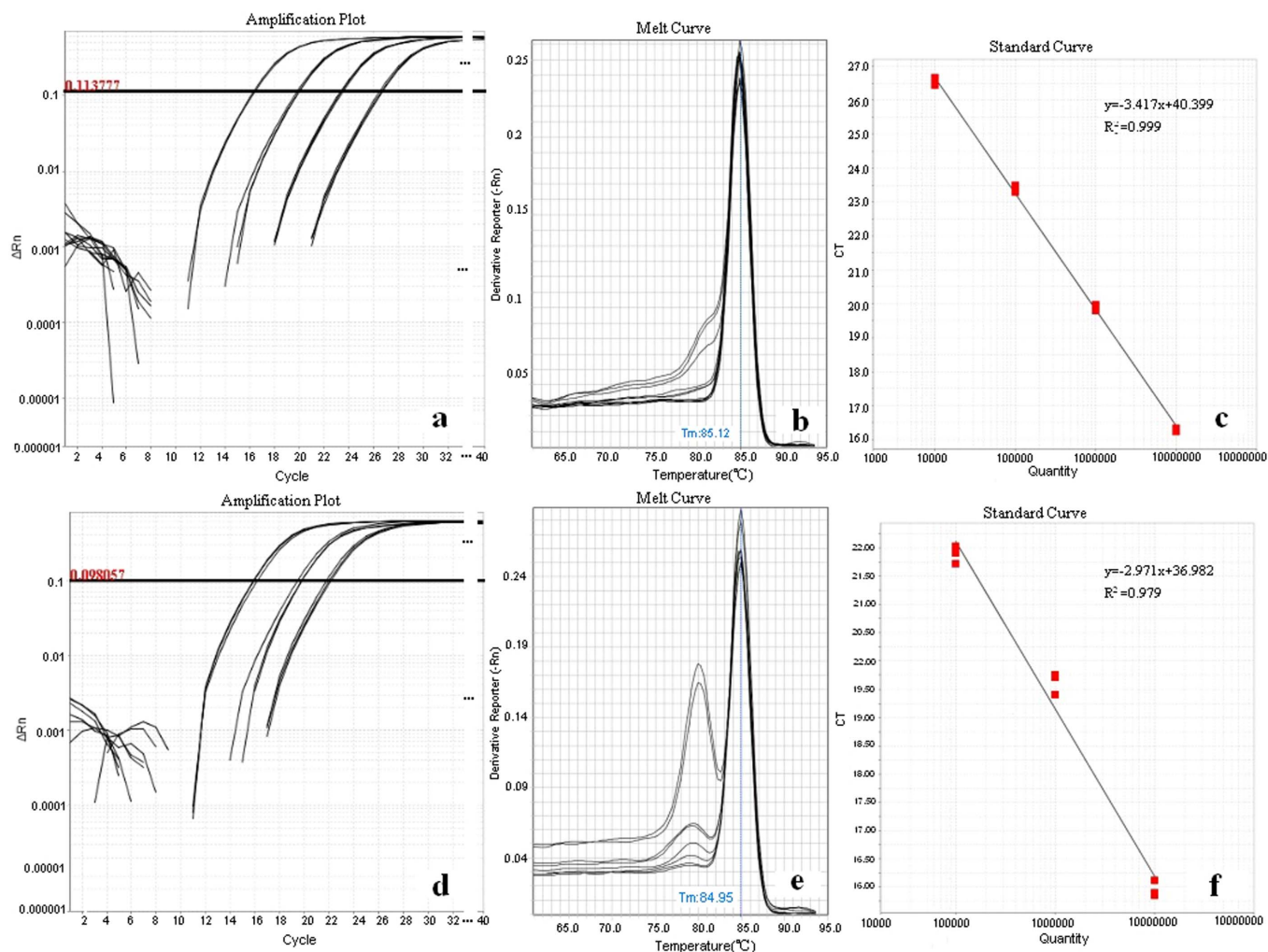


Figure 3 | Real-time PCR results for APs. Figure 3 (a), (b), and (c) show the real-time PCR amplification curve, melting curve and standard curve of AP1, respectively. Figures (d), (e), and (f) show the respective curves for AP2.

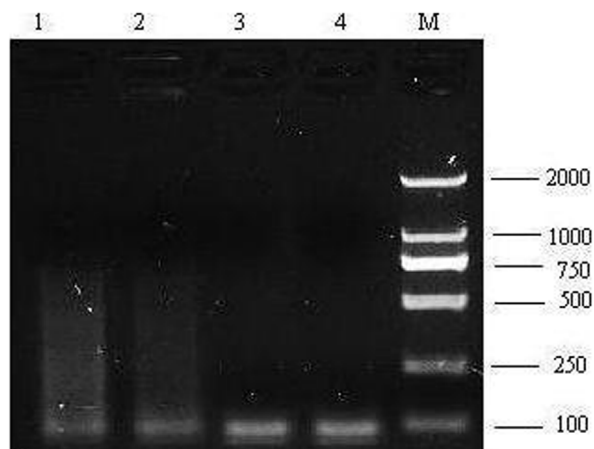


Figure 4 | The specificity of RBF-PCR in GM LY038 and its wild type. The template was the GM LY038 genome in lanes 1 and 2; the template was wild type genome in lanes 3 and 4. M, the marker DL2000.

the target templates were still detectable. Based on the melting curve (Supplementary Figure S3), there was no amplification product when the non-target 218 bp fragment was used as the template. Thus, flanking sequences can definitely be obtained using our developed method.

Identity of amplification between AP/SP and SP/SP. The amplicons generated using SP 673-R and SP inner-F were located within the regions corresponding to SP 673-R and AP1. We tested whether the AP/SP primer pair had the same amplification ability as the SP/SP pair under the same thermal cycle conditions. As shown in Figure 5, under the same concentration of the templates, the Ct value for SP/SP was slightly smaller than that for AP/SP, which may be because both of the SP primers could work together to amplify the target. However, there was no binding site for AP in the first few PCR cycles of the AP/SP reaction, because the sequence of AP1 is the same as Adaptor-1. The SP primer first bound to the target segment, and after few cycles of amplification, the AP1 primer has its binding sites and plays its role. Primer dimers appeared at the same concentration; thus, the detection limit of AP/SP was regarded to be the same as that of SP/SP.

The ratio of AP to SP. We optimized the ratio of AP to SP in semi-nested PCR. The 218 bp and 673 bp fragments were mixed together to simulate the detection. The ratios of AP1 to SP inner-F were 10 : 1, 4 : 1, 2 : 1, 1 : 1, 1 : 2, 1 : 4, 1 : 6, and 1 : 10 (1 equals 100 nM), and the amplification curves were comparable according to the real-time PCR results (Figure 6 and Supplementary Table S1). When the primer ratios were 1 : 1, 1 : 2 and 1 : 4 (highlighted in red), the height of the peaks on the melting curves were higher than the others,

and the Ct values were correspondingly lower. In spite of these observations, the differences between the primer ratios were slight, even when the template was at a low concentration. This finding suggested that the range of AP/SP ratios was wide, and optimized result was obtained when the concentration of SP was higher than AP.

Analysis of the flanking region of the GM LY038 genome. Through the series of RBF-PCR steps, we retrieved the products of semi-nested PCR within 500–2000 bp, which were abundant as evidenced by gel electrophoresis. The T-easy vector transformation and cloning were conducted in triplicate. We prepared 10 plating media for each replicate to perform blue-white selections. One hundred white colonies were selected for culturing and then sequenced (Sangon, Shanghai, China). The numbers of successful sequencing results are listed in Table 2; however, sequencing was unsuccessful in every case, which might be attributed to the following reasons. First, some of the vectors were false positive empty vectors. Second, due to a few non-specific amplifications, the sequencing result neither matched the sequence of the adaptor nor the LY038 exogenous gene. Nonetheless, more than 65% of the sequencing results were definite flanking sequences of the LY038 exogenous gene.

After sequencing, we used BLAST to analyze the obtained sequences, and only small part of them could match the genome of B73 maize, which has been completely sequenced and assembled in maize first. The wild type of LY038 is maize inbred line H99 which has not been sequenced, and the difference may exist between the two genomes. Then, we submitted the obtained flanking sequences to GenBank, and the GenBank Accession number is KF425531; the sequence will be released to public database once they are processed. The obtained sequences are shown in Supplementary Materials S1 and S2.

The fragments between 1000–1500 bp were the most abundant (Table 2), and the length distribution of the sequences exhibited a bell-shaped curve. Based on the results generated from this method, we conclude that an unknown sequence can be precisely and confidently obtained by applying the method just once to continue further research. Therefore, to guarantee obtaining the unknown sequence and also avoid excessive sequencing times, the number of sequencing reactions should be lowered to 30, or even to 10, in future studies.

Discussion

GW is a molecular technique for direct identification of nucleotide sequences from target genomes. The only requirement is the availability of a known nucleotide sequence from which to start. Although several technical improvements have been reported for improving GW methods over the last 20 years, the lack of reproducibility, sensitivity, and specificity are major concerns.

In comparison to other GW techniques, RBF-PCR has striking advantages. It has been proved to be a very simple but strict gene walking method with high sensitivity, and all of the steps adopted are easy to operate, such as ends blunting, adaptor ligating and the T-vector cloning. The genome is fragmented randomly by the ultrasonic processor, and the degree of broken segments can be controlled through altering the ultrasonic time and frequency. This method of genomic DNA preparation does not require high DNA integrity or purity. The use of an ultrasonic processor produces fragments of desired length without enzymatic digestion.

Another innovation in this method is the 5' end-directed adaptor. The partially reverse complemented sequences make the adaptors bind directly to the ends of target fragments. The T overhang at the 3' end of the adaptor can prevent self-ligation between the adaptors. Because the PCR products are dispersed in the agarose gel, the most promising PCR products can be retrieved and cloned into a vector for

Table 1 Real-time PCR results showing the specificity of the primers						
673 bp:218 bp	Ct				Ct Mean	Tm value
	1	2	3	4		
0:1	30.29	30.76	30.09	30.19	30.33	none
0.00001:1	29.90	29.96	29.67	29.70	29.81	81.50
0.0001:1	29.44	29.82	29.90	29.74	29.73	84.50
0.001:1	28.25	28.26	28.38	28.35	28.31	84.50
0.01:1	24.84	24.97	25.05	25.01	24.97	84.50
0.1:1	21.12	21.24	21.46	21.30	21.28	84.50
1:1	17.28	17.60	17.49	17.48	17.46	84.50
1:0	17.70	18.86	17.82	17.61	18.00	84.50

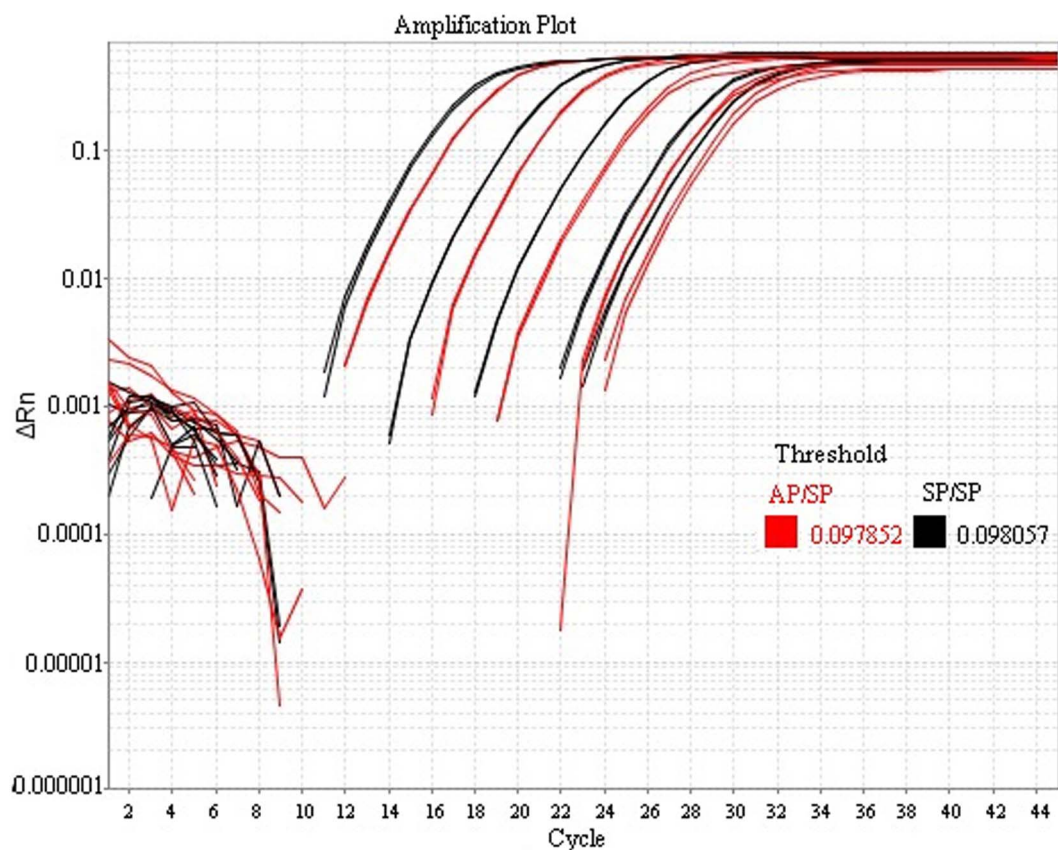


Figure 5 | Amplification identity of the primers AP/SP and SP/SP.

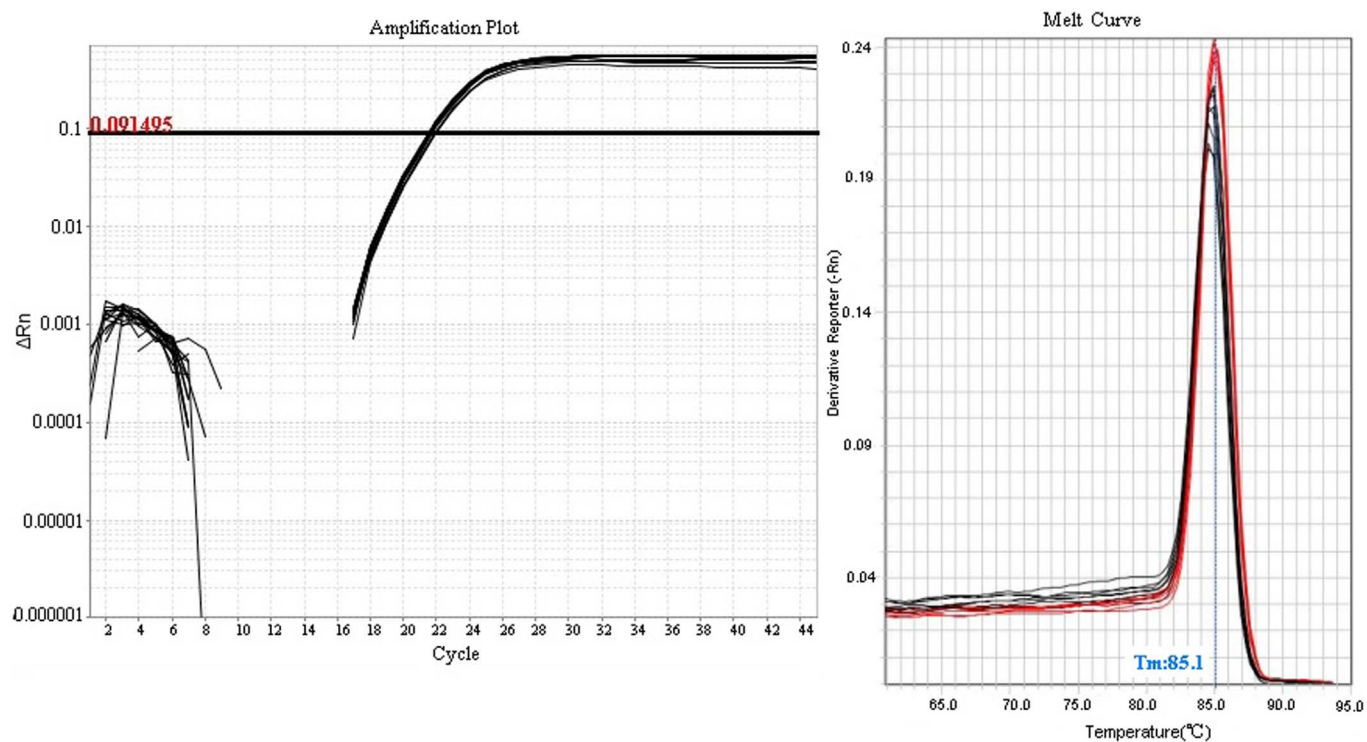


Figure 6 | Real-time PCR results of AP/SP ratios. Figure 6 (a), the amplification curve; (b), the melting curve. The red curves represent the AP/SP ratios for 1:1, 1:2, and 1:4.


Table 2 | The sequencing results of the flanking region of the investigated sequence in the LY038 genome

The distribution of the sequencing length (bp)	Sequencing results that contained the flanking region of the LY038 genome			Mean success rate (%)
	1	2	3	
0–500	8	12	9	9.67
500–1000	19	18	17	18.00
1000–1500	31	29	35	31.67
1500–2000	13	10	16	13.00
2000–	2	0	1	1.00
Total success rate	73.00%	69.00%	78.00%	73.33%

sequencing. The sequencing results are then compared with the semi-nested PCR primers and known sequences, and those which match the two primers and known sequences represent the desired flanking sequences.

Following the serial of steps, we can achieve the traceability through the simulated experiment. All steps are easy to perform, and no expensive instrument is needed. We obtain the unknown sequences through a single run of RBF-PCR; however, the lengths of the flanking sequences primarily depend on two factors. The first factor is the degree of fragmentation of the DNA genome in the ultrasonic processing: the more fragmented the genome is, the shorter is the flanking sequence we can retrieve. The second factor is related to the number of sequencing reactions. Based on our experience, additional sequencing reactions are more likely to generate longer flanking regions.

In particular, this method can be applied to a variety of experiments and would function well in the identification of a multiple insertion site. Directed against specific sequences, particularly those with complex secondary or tertiary structure, the digestion of the genome may fail due to restriction endonuclease limitations. In contrast, fragmentation of the genome in RBF-PCR method would not be limited by such reasons. Therefore, RBF-PCR has an extensive application capacity and many potential advantages.

Methods

DNA extraction. The seeds of genetically modified maize LY038 and wild type were ground into powder with a blender (Supplementary Figure S2 A). A amount of 120 mg sample was measured for genomic DNA extraction using the CTAB method¹². The DNA concentrations were determined spectrophotometrically at 260 nm using a UV/VIS spectrometer (Kontron, Neufahrn, Germany). The DNA purity was determined by calculating the A260/A280 ratio.

Genomic DNA fragmentation and modification. An ultrasonic processor (Scientz JY92-IIN, Ningbo, China) was used to randomly fragment the genomic DNA^{13,14}. The diameters of the ultrasonic probes were 2, 3, and 6 mm. Each group of genomic DNA fragmented with different probes was processed under the same ultrasonic conditions, including the ultrasonic interval, ultrasonic pulse repeats, and the total amount of DNA. The only difference between the groups was the ultrasonic time, from 1 s to 2 s.

The total amount of genomic DNA was no less than 500 ng, and the ideal volume was 100 μ L at least for sufficient ultrasonication of the genome. The sample was then purified by the TIANquick Midi Purification Kit and dissolved in 30 μ L of ddH₂O (Tiangen, China).

Three enzymes were added to blunt the ends of the fragments: T4 DNA polymerase, Klenow DNA polymerase, and T4 PNK (Takara, China)^{15,16}. T4 DNA ligase buffer with 10 mM dATP, dNTP mix, and ddH₂O were added to a total volume of 100 μ L. The reaction was incubated in a thermal cycler for 30 minutes at 20°C. The blunt ends fragments were purified by the TIANquick Midi Purification kit (Tiangen, China) and dissolved in 35 μ L of ddH₂O. Klenow exo (3' to 5' exo minus) and dATP were added according to the manufacturer's instructions to produce A overhang at the 3' ends, and the mixture was incubated for 30 minutes at 37°C. The products were purified by the MinElute PCR Purification Kit (Qiagen, Germany) to eliminate all residual reagents, and dissolved in 15 μ L ddH₂O.

The RBF-PCR adaptor design and preparation. The RBF-PCR adaptors consisted of 33 base pairs, partially reverse complemented with each other for 11 bp; the remaining 22 bp were not. A T was overhung at the 3' end of Adaptor-1 to avoid the ligation between the adaptors. In the ligation reaction, Adaptor-1 can ligate to the 5'

ends of the fragments, and Adaptor-2 can bind to the 3' ends of the fragments. In the subsequent semi-nested PCR, the amplification styles were different through combining different APs with SPs. In the genome walking method, specificity was the most important consideration. The adaptor sequence was not expected to match with species that are commonly searched by BLAST in the nucleotide collection database. SPs were designed using ABI PRISM Primer Express Version 2.0 software (Applied Biosystems, California, USA).

In this study, 20 μ M of Adaptor-1 and Adaptor-2 were used to generate the RBF-PCR adaptor. The adaptors were denatured at 96°C for 2 min and slowly cooled to room temperature, and then the formed RBF-PCR adaptor was diluted with annealing buffer containing 10 mM Tris-HCl (pH 8.0) to 20 μ M.

Ligation with adaptor. A 10 : 1 molar ratio of adaptor to genomic DNA was used, in order that all fragments were ligated to the adaptors. We incubated the mixture of fragments and adaptors in a thermal cycler for 6 hours at 16°C, and unligated adaptors were removed by purification. The product was analyzed by 2% agarose gel electrophoresis, followed by the recovery of fragments ranging from 750–2000 bp.

The recovered fragments were used as templates for amplifications by semi-nested PCR. The adaptors at the 5' and 3' ends were different: the sequence of primer AP1 was the same as the Adaptor-1 at the 5' end, and the sequence of AP2 was the reverse complement of the Adaptor-2 at the 3' end. Thus, AP1 and AP2 were tested separately to decide which one would be ideal for semi-nested PCR. A series of pilot experiments aimed at identifying the specificity and detection limit of AP and the ratio of AP/SP concentration were performed to confirm and optimize this novel method.

PCR condition. Real-time PCR reactions were run on an ABI 7500 (Applied Biosystems, California, USA) with the following program: 2 min at 50°C, 10 min at 95°C, 40 cycles of 15 s at 95°C and 1 min at 60°C. The PCR reaction was performed in final volumes of 25 μ L, containing 11.25 μ L 2.5 \times SYBR Green Master Mix (Tiangen, Beijing, China), 3 μ L template DNA, and 100 nM of each primer.

For semi-nested PCR, the reactions were carried out in final volumes of 30 μ L, containing 1 \times Ex Taq buffer, 0.2 mM dNTPs, 0.3 μ M of each primer, 2.5 units of Ex Taq DNA polymerase (TaKaRa Biotechnology Co. Ltd., China), and 3 μ L DNA template. All of the amplifications were performed using an ABI2720 thermal cycler (Applied Biosystems, U.S.A.) with the following parameters: one step of 5 min at 95°C, 35 cycles of 30 s at 95°C, 30 s at 60°C, and 30 s at 72°C, one step of 5 min at 72°C.

The sequences used for the imitation experiments are listed in Supplementary Table S2. The SPs were designed based on the exogenous genes of GM LY038 genome and constituted the semi-nested primers; the other fixed primer was AP.

Genome walking and sequence analysis. For genome walking, the nested primers were designed based on the exogenous genes of the LY038 genome. R1–R5 were used to obtain the right flanking region of the GM LY038 genome, and designed based on the sequence in the right border region of the exogenous genes introduced in the LY038 genome. L1–L5 were used obtain the left flanking region of the LY038 genome, and designed based on the sequence in the left border region of the exogenous genes. The sequences used in our experiments are listed in Supplementary Table S3.

The PCR products were analyzed on a 2% agarose gel; and the products showing the most intense staining were recovered. The fragments were then ligated into the T-easy vector; the plasmid was transformed into DH5 α competent cells and sequenced. The sequencing primers were designed based on the sequence of the vector¹⁷, and the results were analyzed by the DNAMAN software, which searched for similarities between the sequences and GM LY038 genome.

- Guo, H. & Xiong, J. A specific and versatile genome walking technique. *Gene* **381**, 18–23 (2006).
- Yuanxin, Y. *et al.* T-linker-specific ligation PCR (T-linker PCR): an advanced PCR technique for chromosome walking or for isolation of tagged DNA ends. *Nucleic Acids Res.* **31**(12), e68 (2003).
- Wahl, G. M. *et al.* Cosmid vectors for rapid genomic walking, restriction mapping, and gene transfer. *Proc. Natl. Acad. Sci. USA* **84**(8), 2160–2164 (1987).
- Berriman, M. *et al.* The genome of the African trypanosome *Trypanosoma brucei*. *Science* **309**(5733), 416–422 (2005).
- Acevedo, J. P. *et al.* Cloning of complete genes for novel hydrolytic enzymes from Antarctic sea water bacteria by use of an improved genome walking technique. *J. Biotechnol.* **133**(3), 277–286 (2008).
- Terauchi, R. & Kahl, G. Rapid isolation of promoter sequences by TAIL-PCR: the 5'-flanking regions of Pal and Pgi genes from yams (*Dioscorea*). *Mol. Gen. Genet.* **263**(3), 554–60 (2000).
- Ochman, H., Gerber, A. S. & Hartl, D. L. Genetic applications of an inverse polymerase chain reaction. *Genetics* **120**(3), 621–623 (1988).
- Trinh, Q., Xu, W. T., Shi, H., Luo, Y. B. & Huang, K. L. An A-T linker adapter polymerase chain reaction method for chromosome walking without restriction site cloning bias. *Anal. Biochem.* **425**(1), 62–67 (2012).
- Trinh, Q. *et al.* Loop-linker PCR: An advanced PCR technique for genome walking. *IUBMB Life* **64**(10), 841–845 (2012).
- Thirulogachandar, V., Pandey, P., Vaishnavi, C. S. & Reddy, M. K. An affinity-based genome walking method to find transgene integration loci in transgenic genome. *Anal. Biochem.* **416**(2), 196–201 (2011).



11. Leoni, C., Volpicella, M., De Leo, F., Gallerani, R. & Ceci, L. R. Genome walking in eukaryotes. *FEBS J.* **278**(21), 3953–3977 (2011).
12. Xu, W. T. *et al.* A novel universal primer-multiplex-PCR method with sequencing gel electrophoresis analysis. *PLoS One* **7**(1), e22900 (2012).
13. Elsner, H. I. & Lindblad, E. B. Ultrasonic degradation of DNA. *DNA* **8**(10), 697–701 (1989).
14. Grokhovsky, S. L. Specificity of DNA cleavage by ultrasound. *Mol. Biol.* **40**(2), 317–325 (2006).
15. Clark, J. M., Joyce, C. M. & Beardsley, G. P. Novel blunt-end addition reactions catalyzed by DNA polymerase I of *Escherichia coli*. *J. Mol. Biol.* **198**(1), 123–127 (1987).
16. Briggs, A. W. *et al.* Removal of deaminated cytosines and detection of in vivo methylation in ancient DNA. *Nucleic Acids Res.* **38**(6), e87 (2010).
17. Kim, B. H. *et al.* Enrichment of microbial community generating electricity using a fuel-cell-type electrochemical cell. *Appl. Microbiol. Biotechnol.* **63**(6), 672–681 (2004).

Acknowledgments

The study was funded by the National GMO Cultivation Major Project of New Varieties (Nos. 2013ZX08012-001 and 2014ZX08012-001). We thank Dr. Wen-Hsing Cheng for critical comments.

Author contributions

W.X. designed and performed experiments and wrote the manuscript. Y.S. designed and performed experiments and wrote the manuscript. P.Z. analyzed results and critically read the manuscript. Z.Z. advised on the experimental design and commented on the manuscript. J.H. analyzed results and commented on the manuscript. K.H. and Y.L. arranged funding, designed experiments, and supervised and mentored all work.

Additional information

Supplementary information accompanies this paper at <http://www.nature.com/scientificreports>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Xu, W.T. *et al.* Randomly broken fragment PCR with 5' end-directed adaptor for genome walking. *Sci. Rep.* **3**, 3465; DOI:10.1038/srep03465 (2013).



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported license. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0>