OPEN

Removing PCR for the elimination of undesired DNA fragments cycle by cycle

Jiaojiao Huan[1]*, Kangkang Wan[1]*, Yunjun Liu[2], Wubei Dong[1] & Guoying Wang[2]

[1]Department of Plant Pathology, College of Plant Science and Technology and the Key Lab of Crop Disease Monitoring & Safety Control in Hubei Province, Huazhong Agricultural University, Wuhan 430070, Hubei Province, China, [2]Institute of Crop Sciences and National Center for Plant Gene Research, Chinese Academy of Agricultural Sciences, Southern Street of Zhongguancun 12, 100081 Beijing, China.

A novel removing polymerase chain reaction (R-PCR) technique was developed, which can eliminate undesired genes, cycle by cycle, with efficiencies of 60.9% (cDNAs), 73.6% (genomic DNAs), and ~ 100% (four DNA fragments were tested). Major components of the R-PCR include drivers, a thermostable restriction enzyme - *ApeK*I, and a poly(dA) adapter with mismatched restriction enzyme recognition sites. Drivers were generated from the undesired genes. In each cycle of R-PCR, drivers anneal to complementary sequences and allow extension by Taq DNA polymerase. Thus, *ApeK*I restriction sites in the undesired genes are recovered, and adapters of these undesired DNA fragments are removed. Using R-PCR, we isolated maize upregulated defense-responsive genes and *Blumeria graminis* specialized genes, including key pathogenesis-related effectors. Our results show that after the R-PCR reaction, most undesired genes, including very abundant genes, became undetectable. The R-PCR is an easy and cost-efficient method to eliminate undesired genes and clone desired genes.

P olymerase chain reaction (PCR), developed by Kary B. Mullis, is a technique used to amplify many copies of a region of DNA[1–3]. PCR is a fast, inexpensive and widely used technique to amplify desired sequence fragments. However, how undesired genes may be removed from a gene pool that includes both desired and undesired genes remains a challenge. Genes are turned on or off in different biological processes, such as cellular growth, organogenesis and disease development. To identify and clone specifically expressed genes is the first step and a key strategy to explore these biological processes[4]. To measure and isolate specifically expressed genes, a variety of methods have been developed including differential display PCR, RNA fingerprinting, serial analysis of gene expression (SAGE), real-time quantitative PCR, subtractive suppression hybridization (SSH), microarrays, and high-throughput next-generation sequencing technologies[5–9].

Differential display PCR is a highly sensitive method to investigate regulated genes; however, it generates a large number of false positives and requires a large number of primer pairs. SSH is a very popular subtraction method and is available as a kit. However, there are two potential problems with SSH. The first is gene redundancy. If there are a few genes that are highly differentially expressed, then they will appear as a large number in the SSH results. The second problem is generating many false positives[10]. DNA microarrays represent a high-throughput technique to measure a large number of genes within a single experiment. The use of DNA microarrays holds considerable promise in our understanding of genes and their impact on disease, drug discovery and development. The disadvantages of the microarrays include insufficient sensitivity because of hybridization; sequences must be known in advance; lack of reproducibility; lack of standardization; and expense. Advantages of high-throughput sequencing technologies include that they are highly efficient, and sequences do not need to be known in advance. Their disadvantages include sequencing only very short sequences, complicated post-sequencing data analysis and expense[11,12]. Each method has its own advantages and drawbacks, and no method can easily and efficiently remove undesired DNA fragments. The PCR method is a highly effective technique with few drawbacks. Restriction enzymes cleave DNA at specific nucleotide sequences. The R-PCR method proposed in this study makes use of a restriction enzyme that has only restriction activity and cuts in a predictable and consistent manner. As a time-saver, the enzyme *ApeK*I can digest one unit of assay DNA substrate in 5 min (New England Biolab Inc., USA). SSH is still a popular technique that allows isolation and cloning of differentially expressed genes. Here, we describe a novel method R-PCR, which inherits the merits of PCR, restriction enzymes and SSH.
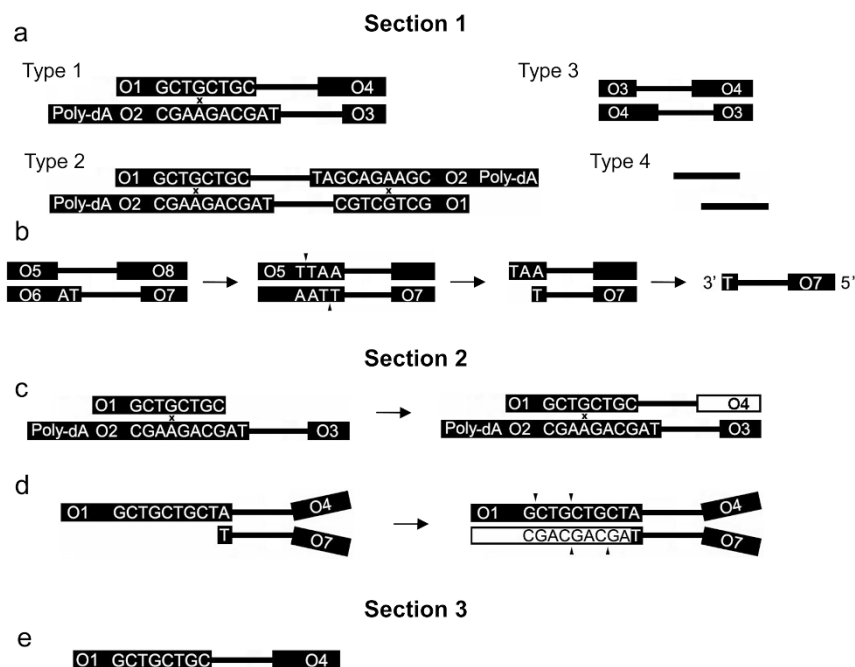
Figure 1 | Three main sections of the R-PCR method. (a). Section 1, tester preparation: DNAs were digested with *ApeK*I and *Mse*I, adapters of O1O2 and O3O4 were added, and four types of fragments were generated. Type 1 is the only expected fragment for R-PCR, which harbors one base mismatched *ApeK*I site; adapters of type 2 were cut by *ApeK*I in R-PCR cycles; type 3 was removed by oligo-dT spin column purification; type 4 may serve as a driver and remove a desired gene, fortunately, it was also removed by oligo-dT spin column purification. (b). Section 1, driver preparation: DNAs were digested with *ApeK*I and *Mse*I, adapters of O5O6 and O7O8 were added, PCR amplified with primers of O5 and O7, PCR products were digested with *Mse*I and were denatured to generate the single strand O7 driver. (c). Section 2, R-PCR reactions: O1 linear amplification. (d). Section 2, R-PCR reactions: *ApeK*I sites of undesired fragments were created and cut. (e). Section 3, recovery of the desired fragments and the products were cloned. After the R-PCR reactions, only the desired O1O4 strands survived due to no driver to match them. Recovery amplification of the desired fragments is carried out using primers of O1-short and O3. The original PolydA-O2O3 fragments could not get amplified because the 3′ end base of the O1-short primer mismatched them.

Results

Outline of the R-PCR method. The R-PCR method is essentially divided into three main sections (Fig. 1 and Supplementary Fig. 1) and includes specifically designed testers, drivers, a single primer, a thermostable restriction enzyme (*ApeK*I), a thermostable Taq DNA polymerase and dNTPs. A brief description of the R-PCR method is as follows: 1) Section 1, tester and driver preparation. The preparation of tester and driver starts from samples digestion with *ApeK*I and *Mse*I. The tester is made by ligation with an adaptor containing a polyA tail and then oligo-dT column purification. The driver is made by ligation of different adaptors, PCR amplification, and digestion with *Mse*I. 2) Section 2, R-PCR reactions. The tester and driver are mixed and subjected to R-PCR with a single primer in the presence of *ApeK*I, which results in linear amplification of the desired fragments without *ApeK*I digestion due to design of a mismatch in the adaptor. In contrast, common undesired fragments are extended from the 3′ end of the driver to create the *ApeK*I site, which is cut, removing those fragments from further amplification. 3) Section 3, recovery of the desired fragments and the products are cloned. Recovery of the desired fragments from linear amplification in the previous step is carried out using selective PCR primers, and the products are cloned by Invitrogen's TOPO TA cloning system. Detailed procedure refers to Figure 1 and Supplementary Figure 1.

The R-PCR design. A critical design aspect for R-PCR is the following nine oligonucleotides and related adapters.

O1: 5′-TTACCACGACCACCCTATTGCTGCTGC-3′
O1-short: 5′-TTACCACGACCACCCTATTGCTG-3′
O2: 5′-TAGCAGAAGCAATAGGGTGGTCGTGGTAAAAAAAA-
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA-3′

O3: 5′-ACGAGGTGCGGTCTTGGACTACTT-3′
O4: 5′-CWGAAGTAGTCCAAGACCGCACCTCGT-3′
O5: 5′-CGACATTCTGTAGGAAACACTAGGACTT-3′
O6: 5′-TAAAGTCCTAGTGTTTCCTACAGAATGTCG-3′
O7: 5′-GGGTTGCGATACGATTGTTATAGGTCAC-3′
O8: 5′-CWGGTGACCTATAACAATCGTATCGCAACCC-3′

Oligonucleotides O1 and O2 form adapter O1O2, which has an *Mse*I-compatible overhang at its end. Oligonucleotides O3 and O4 form adapter O3O4, which has an *ApeK*I-compatible overhang at its end. These two adapters are used for tester preparation (Fig. 1). Here, the tester represents a DNA population in which both desired and undesired genes are included. The role of the poly(dA)$_{36}$ adapter O1O2 is to remove fragments without adapters (Fig. 1, type 4 fragments) and fragments that have O3O4 adapters at both ends (Fig. 1, type 3 fragments) via oligo-dT spin column purification. It is essential that adapter O1O2 harbors two overlapping *ApeK*I recognition sites with one mismatched base pair (Fig. 1). The mismatched base pair can save this adapter from digestion, but once a correct match is recovered in R-PCR cycling, our results show that *ApeK*I can efficiently cut off the overlapping recognition sites. To construct the tester, DNA was digested by *ApeK*I and *Mse*I, and then two adapters, O1O2 and O3O4, were added (Fig. 1 and Supplementary Fig. 1).

Oligonucleotides O5 and O6 form adapter O5O6, which has an *Mse*I-compatible overhang at its end. Oligonucleotides O7 and O8 form adapter O7O8, which has an *ApeK*I-compatible overhang at its end. These two adapters are for the driver preparation (Fig. 1 and Supplementary Fig. 1). Here, the driver represents a DNA population in which only undesired genes (a reference) are included. To obtain the excess driver, we used primers O5 and O7 to amplify DNA fragments of undesired genes (genes from the control sample). The

driver was obtained after the PCR products were digested by *Mse*I and then purified by a commercial spin column.

Four types of DNA fragments were generated in the tester preparation (Fig. 1). Type 1 fragments: one end has the adapter O1O2 and the other end has the adapter O3O4. These are the only expected fragments for R-PCR. Type 2 fragments: both ends have adapter O1O2. Annealing and extension of primer O1 introduced the *ApeK*I recognition sites, which were digested. Thus, fragments with these adapters were removed. Type 3 fragments: both ends have adapter O3O4. As with random amplified polymorphic DNA (RAPD) and single primer amplification reaction (SPAR)[13,14], this type of fragment could also be exponentially amplified during PCR[15–17]. Fortunately, they are removed by the oligo-dT spin column purification. There are many fragments that do not have any adapters in Type 4 fragments. These non-adapter-ligated fragments can serve as the driver and remove desired genes. Fortunately, they are also removed by oligo-dT spin column purification.

Thermostable restriction enzymes and R-PCR reactions. In the R-PCR system, one or more thermostable restriction enzymes were used. We tested *Taq*I, *Tsp509*I, *Pho*I, *Tfi*I, *ApeK*I, and their different combinations. In our results, the *ApeK*I performed the best, and fits into the PCR system well. *ApeK*I is a highly thermostable restriction enzyme that can survive temperatures as high as 95°C. At that temperature, the half-life of the enzyme is 20 min. In addition, *ApeK*I performs complete cleavage in Phusion® High-Fidelity DNA Polymerase buffer (New England Biolab Inc., USA). We used a combination of *ApeK*I and Phusion DNA polymerase in the R-PCR system.

Reactions in each cycle of the R-PCR are more complicated than that in a normal PCR system. Denaturation and primer annealing (Fig. 1, section 2): primer O1 matches DNA fragment O2O3; drivers match DNA fragments O1O4; however, the ends of oligonucleotides O4 and O7 do not match each other, which is essential for the R-PCR system. If O7 and O8 were absent, the complementary strands of the drivers would also match O2O3 template and be extended to obtain an O4 end. In the case that *ApeK*I did not cut off an O1O2 end (also obtained by extension of drivers and the O1 primer), then it would be possible for excess drivers to be amplified in the last step of the R-PCR and produce many false positive results. Oligonucleotides O7 and O8 perform as a cap that blocks drivers (as a template) from being amplified in the final step of R-PCR. Extension and digestion (Fig. 1, section 2): all testers, including induced genes, are linearly amplified. The linear proportional augmentation of the testers by primer O1 makes the R-PCR method very sensitive to rare genes. Drivers make the *ApeK*I recognition sites of non-induced genes recoverable (induced genes do not have drivers), and then the *ApeK*I cuts off the adapters of non-induced genes and eliminates them from further amplification. Furthermore, the newly digested, non-induced genes can serve as novel drivers and, together with original drivers, they can eliminate residual non-induced genes from other R-PCR cycles. In each R-PCR cycle, extension of primer O1 generates one copy of single strand O1O4 (Fig. 1c), which is complementary to the driver with the O7 oligo sequence (O7 driver) (Fig. 1b). Annealing of O1O4 and O7 driver and subsequent extension of the O7 driver by DNA polymerase creates *ApeK*I recognition sites (Fig. 1d), which are cut.

A number of genes could be removed by R-PCR. We verified the removal efficiency of R-PCR using a number of DNA fragments. In Figure 2a, a wheat GTP-binding protein gene fragment was removed by gradually increasing the concentration of the driver from 2 to 20 ng. Our result shows that the R-PCR driver starts to work with an amount as low as 2 or 5 ng in a 20 μl R-PCR reaction. When the amount of the driver approaches that of the tester (15 ng in this case), the band almost disappeared (lane 6), which means that the driver had removed most of the fragments of the tester. This result

demonstrated that a driver concentration only a little higher than that of the tester is sufficient for the R-PCR system. In Figure 2b, four DNA fragments were mixed equally as testers, and certain fragments were removed, as planned, by adding related drivers in the R-PCR reactions. Figure 2a and 2b prove that the R-PCR can nearly remove 100% of undesired bands in a simple system with a number of DNA fragments.

Gene populations could be removed by R-PCR. In the study of gene expression profiles, thousands of genes may be induced or repressed in particular cells. For example, thousands of stress-inducible genes were identified in *Arabidopsis*[18]. To determine the removal efficiency of R-PCR in a population that includes thousands of genes, maize leaf normal expressed genes, maize leaf genes induced after pathogen inoculation and maize leaf genomic DNAs were tested. Here, pathogen inoculation was carried out by inserting a piece of *Rhizoctonia solani* inoculum substrate into the basal leaf sheath of maize. In Figure 2c, lane 1 shows that drivers generated from untreated cDNAs removed untreated maize cDNAs (the testers and drivers were made from the same cDNA population). Lane 2 shows that drivers generated from maize cDNA derived from pathogen-inoculated leaves removed most of the untreated maize cDNAs. In Figure 2d, lane 2 shows that drivers derived from untreated maize leaf cDNA removed only some of the pathogen-inoculated maize leaf cDNAs. The surviving cDNAs are candidates for pathogen-induced genes, which were cloned and analyzed in subsequent experiments. Figure 2e shows that maize genomic DNAs could also be removed by their own drivers (lane 1); however, most of them could not be removed by drivers derived from maize expressed genes (lane 2). These results confirm that drivers work only on their own specific testers.

Verification of R-PCR by real-time PCR. The previous experiment demonstrated that R-PCR works well, as proved by agarose gel
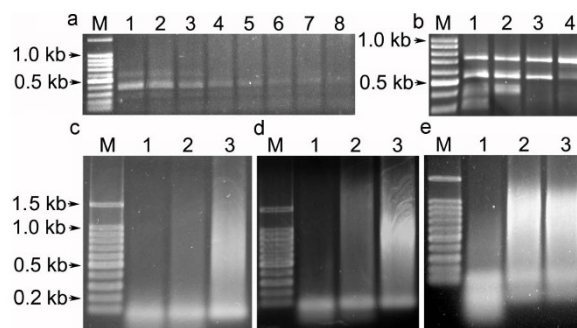
Figure 2 | Agarose gel electrophoresis showing the efficiency of R-PCR. Each R-PCR reaction was performed in a 20 μl system with 40 ng tester and 45 ng driver DNA, except reactions in (a). (a): a single DNA fragment was removed by its own driver. In each reaction, the tester DNA was 15 ng, but the driver was at different concentrations (lane 1 control: 0 ng; lane 2: 2 ng; lane 3: 5 ng; lane 4: 8 ng; lane 5: 11 ng; lane 6: 14 ng; lane 7: 17 ng; and lane 8: 20 ng). (b): four equally mixed DNA fragments were removed by: no driver as a control (lane 1), the smallest driver (lane 2), the smaller two drivers (lane 3), and the smaller three drivers (lane 4). (c): untreated maize leaf cDNAs were removed by: drivers derived from the same cDNA library (lane 1), drivers from pathogen-inoculated maize leaf cDNAs (lane 2), and no driver as a control (lane 3). (d): pathogen-inoculated maize leaf cDNAs were removed by: drivers derived from the same pathogen-inoculated maize leaf cDNAs (lane 1), drivers from untreated maize leaf cDNAs (lane 2), and no driver as a control (lane 3). (e): maize genomic DNAs were removed by: drivers derived from the same maize genomic DNAs (lane 1), drivers from maize cDNAs (lane 2), and no driver as a control (lane 3). Bands less than 0.2 kb are nonspecific amplification products and do not affect our results.

electrophoresis. We then further verified R-PCR using real-time PCR. Fragments ranging from 300 to 1500 bp in lane 2 of Figure 2d were cut, purified, and cloned by T-A cloning system. Fifty two genes were cloned and sequenced (Supplementary Table 1). We randomly selected 30 genes for real-time PCR verification. Based on these DNA sequences, we designed 30 pairs of primers and amplification of 23 genes using real-time PCR. The remaining 7 genes were failed in the real-time PCR amplification. Among the 23 genes, maize wound-induced protein gene and a retro-element were induced by 53- and 64-fold, respectively (six replicates each in real-time PCR). Twelve other genes among the 23 genes were induced by more than 2- and less than 20-fold (Supplementary Fig. 2). The induced gene ratio was 60.9% (14/23). The sequences of the primer pairs of the 14 induced genes are shown in Supplementary Table 2. We identified genes including a chitinase gene, a wound-induced protein gene, retro-elements, phosphoglycerate mutase, anthranilate synthase and calmodulin. Chitinases are upregulated by a variety of stress conditions, both biotic and abiotic, and by phytohormones, such as ethylene, jasmonic acid, and salicylic acid. Like other PR proteins, chitinases play a role in plant resistance against distinct pathogens[19]. The wound-induced protein gene is related to the plant systemic defense responses[20]. The retro-elements constitute a large portion of genomes and are believed to have contributed extensively to genomic evolution and to participate in reprogramming of genetic programs during the course of development and pathogenesis[21]. Differential accumulation of retro-elements and diversification of nucleotide binding domain-leucine rich repeat (NB-LRR) disease resistance genes in duplicated regions follows polyploidy in the ancestor of soybean[22]. The phosphoglycerate mutase gene was identified as a pathogen responsive gene in soybean and *Arabidopsis*[23,24]. Anthranilate synthase is pathogen-inducible (or in a few cases, elicitor-inducible) in *Arabidopsis*[25,26]. Many pathogen-induced calmodulin isoforms are associated with basal resistance against bacterial and fungal pathogens in plants[27]. By a literature search, among the R-PCR isolated 52 genes, we found that at least 31 genes, about 60% (31/52), were reported that are related to known pathogen-induced genes (Supplementary Table 1). The literature review showed that the R-PCR cloned DNA sequences from *R. solani* treated sample are biased to known pathogen-induced genes. These isolated genes proved that the R-PCR is a simple and highly cost-efficient procedure.

The R-PCR removing mechanism at the sequence level. To explore the R-PCR removing mechanism at the DNA sequence level, we

identified *Blumeria* specialized genes and analyzed pairwise sequence comparisons of drivers and testers. *Blumeria graminis* f.sp. *hordei* (Bgh) and f.sp. *tritici* (Bgt) are serious pathogens of barley and wheat, respectively. Using R-PCR, we took Bgh as the tester and Bgt as the driver to identify Bgh-specific genes. At the same time, we took Bgt as the tester and Bgh as the driver to identify Bgt-specific genes. Using BLAST, we compared and verified our R-PCR identified sequences with a Bgh genomic library from BluGen (the *Blumeria* Genome Sequencing Consortium, UK) and a Bgt genomic library from the Genome Survey Sequences (GSS) Database (NCBI). We identified 39 verified *Blumeria*-specific DNA fragments (Supplementary Table 3), including 10 pathogenesis effector-related genes (Supplementary Fig. 3). In Bgh and Bgt speciation, most effectors represent species-specific adaptations[28]. Comparative sequence analysis revealed that Bgh and Bgt diverged 10 million years ago, and basic mutations would on average lead to approximately 13% nucleotide difference in sequences that were not subjected to a selection pressure[29]. If the value is less or greater than 13%, the sequences would be subjected to selection pressure. Thus, it is likely that sequence identities between Bgh and Bgt of less than 87% were under diversifying differentiation with a selection pressure. The sequence identities of R-PCR cloned genes between Bgh and Bgt were analyzed in Supplementary Table 4. We used 87% of identity as one of the three assessment criteria to analyze R-PCR cloned genes (Fig. 3 and Supplementary Table 4). The R-PCR efficiency was 73.6% (Supplementary Table 3).

We randomly selected 2000 Bgt sequences from the GSS database and did a blastn analysis against a Bgh genomic library from BluGen. There were 1834 matched sequences between these two libraries and with an average identity of 90.48%. If the R-PCR did not work efficiently, identities between Bgh and Bgt of our cloned sequences from the R-PCR should be about 90.48%. However, the average identity of our cloned 53 sequences by the R-PCR is 79.49% between Bgt and Bgh (Supplementary Table 4). Because data in our two samples are not normally distributed, we did a Wilcoxon Rank Sum test, and the p-value is $9.28e^{-8}$. This p-value is quite low enough to reject the null hypothesis of no difference in favor of the alternative that values of the two average identities (90.48% and 79.49%) are significantly different. The R-PCR can be used to isolate differential DNA sequences efficiently. To investigate the removal mechanism of R-PCR at the DNA sequence level, we compared tester and driver sequences and found in three cases that drivers could not remove the testers (Fig. 3 and Supplementary Fig. 3).
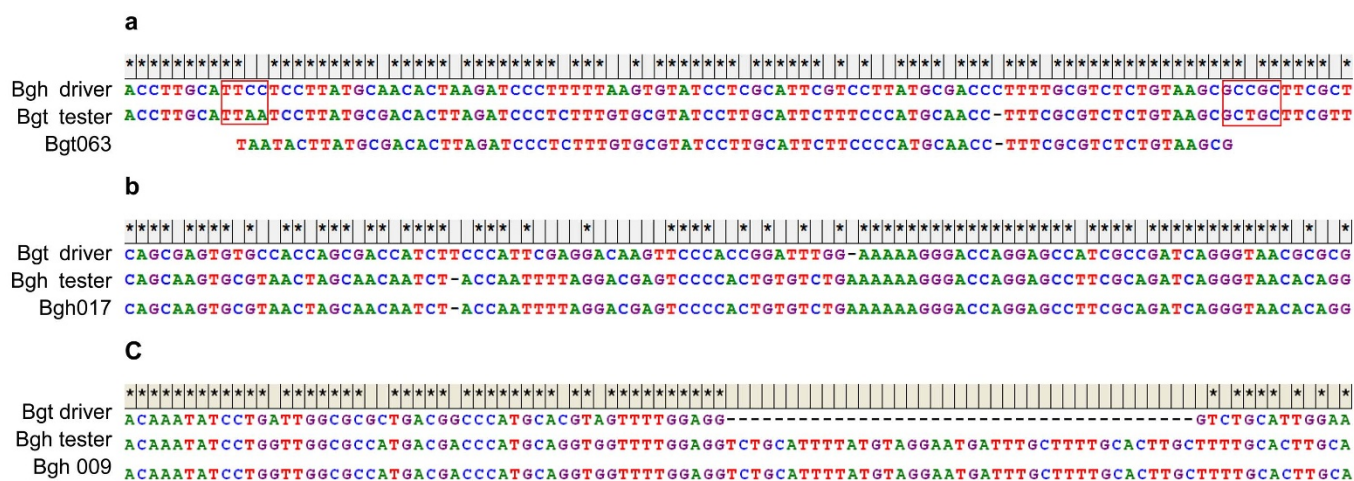
Figure 3 | Drivers could not remove testers in three cases. (a): case 1, the *Mse*I and *Ape*KI recognition sites (in red boxes) have mutations. (b): case 2, the tester contains enough bases that do not match the driver. (c): case 3, the sequence is partially absent in the tester or driver (or completely absent in the driver). Fragments of Bgt063, Bgh017, and Bgh009 were cloned and sequenced after R-PCR reactions. *: Identical.

Discussion

The R-PCR method is a restriction enzyme-based PCR, in which some genes are not amplified, but are removed by means of removing drivers. The gene removing mechanism works by drivers annealing to complementary sequences and being elongated by the polymerase. Then, the specifically designed mismatched overlapped restriction sites are recovered and digested by *ApeK*I. After digestion, this gene fragment will lose its adapter and cannot be further amplified by regular primers. Furthermore, this adapter-lost gene fragment may serve as a fresh driver to remove residual adapters of the same gene fragment in subsequent R-PCR cycles.

The R-PCR design is different from that of suppression subtractive hybridization (SSH). SSH is widely used to selectively amplify target cDNA fragments (differentially expressed) and simultaneously suppress non-target DNA amplification. In SSH, the driver is in excess, but in the R-PCR system, the driver is not necessarily in excess. In SSH, two rounds of hybridization are normally performed, by hand[5,30]. The R-PCR is performed automatically in a PCR machine for about 12 cycles, and, importantly, drivers matching non-specifically expressed genes are removed in each cycle. In some SSH experiments, the number of background clones might considerably exceed the number of target clones in the subtracted libraries[31]. Many false positive results were generated in our earlier versions of the R-PCR method. We then focused on to how to prevent this problem and increase removal efficiency in the current version of the R-PCR design. The key strategies in the R-PCR design were: (i) design and use of removing drivers to eliminate undesired genes, cycle by cycle; (ii) the use of a Poly(dA)$_{36}$ adapter to retain only the adapter-ligated fragments in oligo-dT spin column purification step (non-adapter-ligated fragments could serve as drivers); (iii) the design of the overlapping two *ApeK*I recognition sites and placing them at the ends of adapter O1O2 to increase the removal efficiency and to eliminate nonspecific gene removal in the R-PCR reaction; (iv) design of a mismatched base pair in the overlapping *ApeK*I recognition sites to save original O1O2 adapters from digestion by *ApeK*I, the mismatched base pair also stops amplification of the original O2O3 DNA strand by primer O1-short during the last step of R-PCR amplification of desired genes; (v) the use of adapter O7O8 that serves as a blocking cap, which means that drivers can only can serve as primers to obtain extension and cannot serve as templates (after extension) for amplification by regular primers O1 and O3; and (vi) using the combination of the *ApeK*I and the Phusion® High-Fidelity DNA Polymerase in the R-PCR system. These key design strategies produced an efficient R-PCR system.

The R-PCR design is based on PCR; however, in each R-PCR cycle, genes are removed, whereas, in each cycle of regular PCR, genes are amplified. The R-PCR is a counterpart of PCR, and it has similar advantages and drawbacks to PCR. A common concern for PCR is nonspecific amplification. In the R-PCR system, we used the Phusion® High-Fidelity DNA Polymerase, which shows high fidelity, speed, and an error rate 50-fold lower than that of Taq DNA Polymerase and 6-fold lower than that of *Pyrococcus furiosus* DNA Polymerase (New England Biolab Inc., USA). In our experiments, the Phusion® High-Fidelity DNA Polymerase performed well. To avoid reducing representatives of cloned genes, the number of R-PCR cycles is generally less than 15.

In R-PCR cycling, the tester could be progressively eradicated by the homologous drivers from the reference sample, even if these were initially present in low quantities. A potential weakness of the R-PCR technique, compared for example to RNA-seq, is that it might not detect genes that are over-represented (though not entirely specific) to a given experimental sample, compared to the reference sample used. The extent of this weakness is related to the number of R-PCR cycles. To reduce the number of R-PCR cycles might reduce the effect caused by this weakness. In our real-time PCR verification experiments, among 23 genes, there were 6 genes that expressed less than

four-fold (Supplementary Fig. 2). These genes were isolated under 12 R-PCR cycles. It proves that even with the above potential weakness, the R-PCR may isolate differently expressed genes. Although it has the weakness, the current version of the R-PCR system works well, and it represents an alternative to several existing techniques for the identification of differentially expressed genes, including the RNA-seq. In the future, drivers in the R-PCR system may be made shorter and become commercially available for specific samples or species.

Methods

Growth conditions of plants and fungi. Seedlings of a maize inbred line, Ye478, were grown in a growth chamber with day/night temperatures of 28°C/25°C, a relative humidity of 60% and 15-h photoperiods (light intensity: 800 μmol/m²s). *Rhizoctonia solani* AG-1-IA was cultured in the dark on PDA plates for 2–3 days at 26°C. Barley and wheat seedlings were grown in pots of compost soil in a growth chamber (16 h of light, 8 h of dark, 70% relative humidity, 20°C constant temperature). Bgh and Bgt were maintained at 22°C under a 16-h-light/8-h-dark cycle by weekly transfer to fresh barley and wheat seedlings.

Inoculation of *R. solani* and sample collection. Mycelial disks (4-mm diameter), cut from the edges of an actively growing colony, were used as an inoculum. Inoculation was carried out by inserting a piece of inoculum substrate into the basal leaf sheath[32]. High relative humidity was maintained by regularly misting the plants each day. Equal pieces of leaves were collected at 6, 12, 24, 32, 40, 48 and 72 h after inoculation. The leaf samples were then mixed for RNA extraction and cDNA synthesis.

DNA and RNA extraction and cDNA synthesis. DNA of *Blumeria* spores was extracted using the method of Robinson *et al.*[33]. Maize total RNA was extracted according to Logemann *et al.*[34]. Double-stranded cDNA was synthesized and amplified using SMARTer PCR cDNA synthesis kit (Clontech, Inc. CA, USA).

Real-time PCR. Quantitative PCR (SYBR Green PCR kit, Qiagen, Germany) was performed according to the manufacturer's protocols.

BLAST analysis. Using standalone BLAST analysis, we compared our R-PCR cloned sequences with Bgh and Bgt genomic libraries one by one by paying attention to the three cases in Figure 3. The Bgh genomic library was downloaded from BluGen (the *Blumeria* Genome Sequencing Consortium, UK). The Bgt genomic library was collected from the Genome Survey Sequences Database (NCBI). Standalone BLAST software was downloaded from website of the NCBI BLAST.

1. Mullis, K. *et al.* Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb Sym* **51**, 263–273 (1986).
2. Mullis, K. B. & Faloona, F. A. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Method Enzymol* **155**, 335–350 (1987).
3. Saiki, R. K. *et al.* Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**, 1350–1354 (1985).
4. Rosok, O. & Sioud, M. Discovery of differentially expressed genes: technical considerations. *Methods Mol Biol* **360**, 115–129 (2007).
5. Gadgil, C., Rink, A., Beattie, C. & Hu, W. S. A mathematical model for suppression subtractive hybridization. *Comp Funct Genomics* **3**, 405–422 (2002).
6. Ji, W., Wright, M. B., Cai, L., Flament, A. & Lindpaintner, K. Efficacy of SSH PCR in isolating differentially expressed genes. *BMC genomics* **3**, 12 (2002).
7. Li, L., Techel, D., Gretz, N. & Hildebrandt, A. A novel transcriptome subtraction method for the detection of differentially expressed genes in highly complex eukaryotes. *Nucleic Acids Res* **33**, e136 (2005).
8. Montgomery, S. B. *et al.* Transcriptome genetics using second generation sequencing in a Caucasian population. *Nature* **464**, 773–777 (2010).
9. von Stein, O. D., Thies, W. G. & Hofmann, M. A high throughput screening for rarely transcribed differentially expressed genes. *Nucleic Acids Res* **25**, 2598–2602 (1997).
10. Goetz, F. W. The "ups" and "downs" in Using Subtractive Cloning Techniques to Isolate Regulated Genes in Fish. *Integr Comp Biol* **43**, 786–793 (2003).
11. Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* **10**, 57–63 (2009).
12. Baginsky, S., Hennig, L., Zimmermann, P. & Gruissem, W. Gene expression analysis, proteomics, and network discovery. *Plant Physiol* **152**, 402–410 (2010).
13. Bhattacharya, E., Dandin, S. B. & Ranade, S. A. Single primer amplification reaction methods reveal exotic and indigenous mulberry varieties are similarly diverse. *J Biosci* **30**, 669–677 (2005).
14. Sharma, S. K., Kumaria, S., Tandon, P. & Rao, S. R. Single primer amplification reaction (SPAR) reveals inter- and intra-specific natural genetic variation in five species of Cymbidium (Orchidaceae). *Gene* **483**, 54–62 (2011).
15. Ferreira, E., Gontijo, C., Cruz, I., Melo, M. & Silva, A. Alternative PCR protocol using a single primer set for assessing DNA quality in several tissues from a large variety of mammalian species living in areas endemic for leishmaniasis. *Mem Inst Oswaldo Cruz* **105**, 895–898 (2010).

16. Ma, J. *et al*. Single-primer PCR correction: a strategy for false-positive exclusion. *Genet Mol Res* **10**, 150–159 (2011).

17. Screaton, G. R., Bangham, C. R. & Bell, J. I. Direct sequencing of single primer PCR products: a rapid method to achieve short chromosomal walks. *Nucleic Acids Res* **21**, 2263–2264 (1993).

18. Shinozaki, K. & Yamaguchi-Shinozaki, K. Gene networks involved in drought stress response and tolerance. *J Exp Bot* **58**, 221–227 (2007).

19. Kasprzewska, A. Plant chitinases--regulation and function. *Cell Mol Biol Lett* **8**, 809–824 (2003).

20. Yang, Y., Shah, J. & Klessig, D. F. Signal perception and transduction in plant defense responses. *Genes Dev* **11**, 1621–1639 (1997).

21. Ramos, K. S. Unraveling genetic regulatory networks of mammalian retroelements. *BMC Proc* **3 Suppl 2**, S3 (2009).

22. Innes, R. W. *et al*. Differential accumulation of retroelements and diversification of NB-LRR disease resistance genes in duplicated regions following polyploidy in the ancestor of soybean. *Plant Physiol* **148**, 1740–1759 (2008).

23. Mazarei, M., Lennon, K., Puthoff, D., Rodermel, S. & Baum, T. Expression of an Arabidopsis phosphoglycerate mutase homologue is localized to apical meristems, regulated by hormones, and induced by sedentary plant-parasitic nematodes. *Plant Mol Biol* **53**, 513–530 (2003).

24. Wan, J. *et al*. Proteomic analysis of soybean root hairs after infection by Bradyrhizobium japonicum. *Mol Plant Microbe Interact* **18**, 458–467 (2005).

25. Zhao, J. & Last, R. L. Coordinate regulation of the tryptophan biosynthetic pathway and indolic phytoalexin accumulation in Arabidopsis. *Plant Cell* **8**, 2235–2244 (1996).

26. Niyogi, K. K. & Fink, G. R. Two anthranilate synthase genes in Arabidopsis: defense-related regulation of the tryptophan pathway. *Plant Cell* **4**, 721–733 (1992).

27. Takabatake, R. *et al*. Pathogen-induced calmodulin isoforms in basal resistance against bacterial and fungal pathogens in tobacco. *Plant Cell Physiol* **48**, 414–423 (2007).

28. Spanu, P. D. *et al*. Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. *Science* **330**, 1543–1546 (2010).

29. Oberhaensli, S. *et al*. Comparative sequence analysis of wheat and barley powdery mildew fungi reveals gene colinearity, dates divergence and indicates host-pathogen co-evolution. *Fungal Genet Biol* **48**, 327–334 (2011).

30. Diatchenko, L. *et al*. Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc Natl Acad Sci USA.* **93**, 6025–6030 (1996).

31. Rebrikov, D. V. *et al*. Mirror orientation selection (MOS): a method for eliminating false positive clones from libraries generated by suppression subtractive hybridization. *Nucleic Acids Res* **28**, E90 (2000).

32. Li, H. R., Wu, B. C. & Yan, H. Q. Aetiology of Rhizoctonia in sheath blight of maize in Sichuan. *Plant Pathol* **47**, 16–21 (1998).

33. Robinson, H. L., Ridout, C. J., Sierotzki, H., Gisi, U. & Brown, J. K. Isogamous, hermaphroditic inheritance of mitochondrion-encoded resistance to Qo inhibitor fungicides in Blumeria graminis f. sp. tritici. *Fungal Genet Biol* **36**, 98–106 (2002).

34. Logemann, J., Schell, J. & Willmitzer, L. Improved method for the isolation of RNA from plant tissues. *Anal Biochem* **163**, 16–20 (1987).

Author contributions

J.H. and K.W. carried out most of the experiments described in this paper; Y.L. verified the R-PCR experiments; W.D. and G.W. designed the R-PCR; and W.D. wrote the paper.

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: Huan, J., Wan, K., Liu, Y., Dong, W. & Wang, G. Removing PCR for the elimination of undesired DNA fragments cycle by cycle. *Sci. Rep.* **3**, 2303; DOI:10.1038/srep02303 (2013).