L-RCA (ligation-rolling circle amplification): a general method for genotyping of single nucleotide polymorphisms (SNPs)

Xiaoquan Qi*, Saleha Bakht, Katrien M. Devos¹, Mike D. Gale¹ and Anne Osbourn

Sainsbury Laboratory and ¹John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, UK

Received July 23, 2001; Revised and Accepted October 5, 2001

ABSTRACT

A flexible, non-gel-based single nucleotide polymorphism (SNP) detection method is described. The method adopts thermostable ligation for allele discrimination and rolling circle amplification (RCA) for signal enhancement. Clear allelic discrimination was achieved after staining of the final reaction mixtures with Cybr-Gold and visualisation by UV illumination. The use of a compatible buffer system for all enzymes allows the reaction to be initiated and detected in the same tube or microplate well, so that the experiment can be scaled up easily for high-throughput detection. Only a small amount of DNA (i.e. 50 ng) is required per assay, and use of carefully designed short padlock probes coupled with generic primers and probes make the SNP detection cost effective. Biallelic assay by hybridisation of the RCA products with fluorescence dye-labelled probes is demonstrated, indicating that ligation-RCA (L-RCA) has potential for multiplexed assays.

INTRODUCTION

Recent genome projects in humans, Arabidopsis, rice and other species have provided evidence that single nucleotide changes, referred to as single nucleotide polymorphisms (SNPs), are common among different genotypes of a species (1–3). SNPs are usually biallelic polymorphisms although there are four nucleotide possibilities in principle (4). The availability of a large number of SNPs in genomes enables construction of a highly saturated genetic map. Large-scale genome polymorphism studies will have great impact on human disease diagnosis and drug development. In plants, direct selection for agronomically important traits in large numbers of seedlings can be conducted by using SNPs in and around gene coding regions. These applications require technologies capable of genotyping thousands of SNPs from large numbers of individual DNA samples in an accurate, rapid and cost-effective manner (5).

A wide variety of approaches to genotype SNPs have been developed (6). Restriction endonuclease digestion, allele-specific oligonucleotide hybridisation (7,8), PCR primer mismatch (9), single-base extension (10), oligonucleotide ligation (11–13), pyrosequencing (14,15) and invasive complex cleavage (16,17) are often used to discriminate the point mutations. Targets or signals are usually amplified by PCR. Gel electrophoresis, mass spectrometry (5), HPLC (18) or special fluorescence detection systems are used to separate and detect signals. However, methods based on gel electrophoresis can hardly meet the requirements for high-throughput analysis. Facilities for mass spectrometry are often not readily available for extensive analysis of large numbers of samples. The newly developed Invader™ assay and pyrosequencing seem to be the promising technology for SNP genotyping (14,16). However, these methodologies require special facilities and reagents and, like mass spectrometry-based methods, cannot be used for multiplexed assays (19).

Circularisation of padlock probes (20) with T4 ligase or thermostable ligase (21,22) can specifically and sensitively discriminate point mutations in target DNA sequences. Circularised padlock probes can be efficiently amplified by single-primer rolling circle amplification (RCA) with φ29 DNA polymerase (23) or by two-primer hyperbranched-RCA (HRCA) with Vent® (exo–) DNA polymerase (24). The above technologies have potential for development of a new, high-throughput SNP genotyping system (25).

In this paper we describe a flexible, non-gel-based SNP genotyping method for complex plant genomes that involves thermostable ligation coupled with signal amplification by RCA. The method has been tested on sodium azide-generated saponin-deficient sad mutants of diploid oat (Avena strigosa accession S75) that we have generated as part of our analysis of plant secondary metabolite biosynthesis (26). These mutants are compromised in their ability to synthesise antimicrobial triterpenoid saponins known asavenacins that act as chemical defences against invading pathogens. Recently we have cloned the gene encoding A. strigosa β-amyrin synthase (AsbAS1), a novel oxidosqualene cyclase that catalyses the first committed step in avenacin biosynthesis (27). We have also identified two sad mutants (109 and 610) that are defective in the gene encoding this enzyme, providing a direct link between AsbAS1 and disease resistance. Sequence analysis of wild-type (S75) and the sad1 mutants 109 and 610 revealed a single base change of G to A in the AsbAS1 gene in each mutant (Table 1). These point mutations were used as test systems for the development of the L-RCA SNP genotyping method. In brief, the single copy gene was amplified from a small amount of genomic DNA (~50 ng) by PCR. Short padlock probes

*To whom correspondence should be addressed. Tel: +44 1603 450412; Fax: +44 1603 450011; Email: xiaquan.qi@bbsrc.ac.uk
and 610 were described in Papadopoulou et al. (26). Genomic DNA was isolated from the wild-type (S75) and mutant 109. Amy109F and Amy109R were designed for amplification of the flanking region of the SNP between S75 and mutant 610 in AsbAS1. Amy109OCP_C44, Amy109OCP_C52 and Amy109OCP_C60 are padlock probes with 44, 52 and 60 nt, respectively, designed to target the G allele at position 1912 bp in the AsbAS1 gene in S75, and the padlock probe Amy109OCP_T60 was designed to target the A allele in mutant 109. Amy610OCP_C and Amy610OCP_T target the G and A alleles at position 3417 bp in the AsbAS1 gene in S75 and mutant 610, respectively. OCP_T and OCP_C are the generic padlock probes, designed for targeting A and G alleles, respectively. RCAP_A and RCAP_G are primers for RCA of the circularised padlock probes OCP_T and OCP_C, respectively. OCP_T_H and OCP_C_H were used either as the primers for HRCA or as hybridisation probes for detection of the single-stranded DNA products from RCA.

Amplification of targets by asymmetric PCR

The SNP region was amplified by asymmetrical PCR using the forward/reverse primer pair, Amy109F/Amy109R or Amy610F/amy610R, in the ratio 9:1. Each 20 µl PCR reaction contained 50 ng of oat genomic DNA, 0.45 µM forward primer, 0.05 µM reverse primer, 200 µM dNTPs, 10 mM Tris–HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl2 and 0.5 U of Taq DNA polymerase (Roche Molecular Biochemicals). After an initial denaturation at 94°C for 3 min, amplification was carried out for 35 cycles of 30 s denaturation at 94°C, 30 s annealing at 58°C and 45 s extension at 72°C. PCR products were split into two 10 µl aliquots for subsequent ligation with padlock probes.

Ligation of padlock probes

Aliquots (5 µl) of ligation mixture containing 15 mM MgCl2, 1.5 mM nicotinamide adenine dinucleotide (NAD), 2 U of Ampligase® Thermostable DNA Ligase (Epicient Technologies) and 0.6 µM allele-specific padlock probe (Table 1) were added to each 10 µl PCR reaction containing the amplified target DNA. Multiple cycle ligation was conducted with 15 cycles of 30 s denaturation at 94°C and 5 min ligation at 37°C.

RCA reaction

The ligation mixture was used directly for RCA. An aliquot of 5 µl of 0.6 µM allele-specific amplification primer (RCAP_G or RCAP_A) was added. The reactions were heated to 70°C and cooled to room temperature prior to the addition of 5 µl of RCA mix [800 µM dNTPs, 50 mM Tris–HCl (pH 8.3), 250 mM KCl, 7.5 mM MgCl2, 0.8 µg/µl BSA, 140 ng/µl phage T4 gene-32 protein and 12 ng/µl of φ29 DNA polymerase], followed by incubation at 37°C for 3 h. Reactions were terminated by incubation at 70°C for 10 min.

HRCA reaction

In this research, HRCA was also used for amplification of circularised padlock probes. A 10 µl aliquot containing 400 µM dNTPs, 50 mM Tris–HCl (pH 8.3 at 25°C), 25 mM KCl, 5 mM MgSO4, 25 mM (NH4)2SO4, 0.05% Triton X-100, 70 ng/µl phage T4 gene-32 protein, 0.25 U/µl of Vent (exo−) DNA polymerase (New England Biolabs) and 0.3 µM allele-specific amplification/hybridisation primer pair was added to the 15 µl ligation mixture for amplification of circularised padlock probes. The reaction was incubated at 92°C for 3 min, and then at 63°C for 3 h.

Figure 1. Design of short padlock probes. Arrows indicate the 5′→3′ direction of the oligonucleotides. Padlock probe Amy109OCP_C60 matches the target G allele (at 1912 bp) in the wild-type (S75) AsbAS1 gene leading to successful ligation and mismatches with the mutant A allele in mutant 109 causing ligation failure. RCAP_G is complementary to the padlock probes and OCP_C_H has an identical sequence to part of the padlock probe (Table 1).

MATERIALS AND METHODS

Plant materials

The diploid A. strigosa accession S75 and the sad1 mutants 109 and 610 were described in Papadopoulou et al. (26). Point mutants in the AsbAS1 gene of mutants 109 and 610 are shown in Table 1. Scoring of saponin-deficient seedlings was conducted under UV illumination according to Papadopoulou et al. (26). Genomic DNA was isolated from the wild-type S75, sad1 mutants 109 and 610, and the 83 and 96 F2 progenies derived from crosses between S75 and mutant 109, and between S75 and 610, respectively, by using the DNeasy 96 Plant Kit (Qiagen).

Primers and probes

Primers and probes were synthesised by Life Technologies or MWG-Biotech. Sequences are given in Table 1. Primers Amy109F and Amy109R were designed for amplification of the flanking region of the SNP between S75 and mutant 109, and Amy 610F and Amy610R were designed to amplify the matching target DNA sequences at both the 3′ and 5′ ends were designed. A perfect match at the target region between the target DNA and the padlock probe allows circularisation of the padlock probe in the presence of thermostable ligase, while a single base mismatch at the 3′ end of the probe causes failure of ligation (Fig. 1). The circularised padlock probe is then amplified with φ29 DNA polymerase using primers derived from the central region (backbone) of the padlock probe. Amplified DNA is detected by UV illumination after staining with Cybr-gold® (Molecular Probes). Alternatively, the biotin 5′-end-labelled primers are used in RCA, and biallelic assay is conducted by fluorescence scanning after hybridisation of RCA products with fluorescent dye-labelled oligonucleotide probes. Compatible buffers are used for these reactions and the whole process is carried out in a single tube or microplate well. This method therefore has potential for high throughput automated analysis at low cost.
Signal detection under UV illumination

Prior to detection, 100 µl of 1000-fold diluted Cybr-Gold (Molecular Probes) in TAE (pH 7.6) was added to the reactions. After brief mixing, stained DNA was visualised under UV illumination (maximum peak at 302 nm), and the fluorescence was quantified using the Gel Doc 1000 system (Bio-Rad).

Biallelic assay by hybridisation of RCA products with fluorescence dye-labelled probes

Biotinylated and fluorescence dye-labelled primers and probes were obtained from MWG-Biotech. The biotin 5′-end-labelled primers RCAP_G and RCAP_A were bound to Dynabeads M-280 Streptavidin according to the manufacturer’s instructions (Dynal Biotech). PCR, ligation and RCA were as above, except that Amy109OCP_C60 and Amy109OCP_T60 were added into the same ligation reaction, and Dynabeads–RCAP_G and Dynabeads–RCAP_A complexes were used simultaneously in RCA. After RCA using φ29 DNA polymerase, the DNA–Dynabeads complex was washed once with 100 µl 10 mM Tris–HCl (pH 8.5) and resuspended in 100 µl 3.5× SSC/0.1% SDS. Then, 10 pmol of the Cy3- and Cy5-labelled OCP_C_H and OCP_T_H were added to the solutions, mixed, and incubated at 42°C for 15 min. The DNA–Dynabeads complex was then washed once at room temperature with 2× SSC/0.1% SDS and twice in 0.2× SSC. Fluorescence was detected using the Typhoon 8600 System (Amersham). A red laser (633 nm) and 670 BP30 emission filter were used for detecting Cy5, and a green laser (532 nm) and 555 BP20 filter for Cy3.

Southern blot analysis

RCA products were separated by agarose gel electrophoresis and transferred to Hybond-N+ nylon membranes (Amersham) using alkaline lysis transfer (28). The 24mer probes for hybridisation with RCA products were end-labelled using [γ-33P]ATP (NEN Life Science Products). Hybridisation was conducted using Church–Gilbert solution [0.5 M sodium phosphate buffer (pH 7.2), 7% SDS, 1 mM EDTA (pH 8.0)] (29) at 58°C for 16 h. After hybridisation, membranes were washed twice for 10 min at low stringency (2× SSC, 1% SDS at 58°C) and then subjected to autoradiography.
RESULTS

Use of short padlock probes

Utilisation of long padlock probes (~80 nt) is expensive and problematic because of synthesis failures at the 3′ or 5′ ends or even internally. In our experiments, shorter padlock probes were tested (Table 1 and Fig. 2). Ligation was conducted using multiple cycles (see Materials and Methods) instead of at a constant temperature, since we found that ‘multiple-cycle’ ligation gave more effective circularisation of padlock probes and showed higher specificity in the discrimination of mismatches (data not shown). Efficient circularisation occurred for padlock probes Amy109OCP_C52 and Amy109OCP_C60, but not for Amy109OCP_C44. Ligation was more efficient for the 60 nt padlock probe than for the 52 nt padlock probe (Fig. 2). Ampligase® Thermostable DNA Ligase produced more circularised padlock probe at 37°C than at 34 or 42°C under the conditions tested (Fig. 2). The lower amount of circularised padlock probe obtained when ligation was carried out at 42°C compared to 37°C may be associated with the length of the complementary sequences. The padlock probes used in the above experiment contained 11 and 12 nt of complementary sequence at the 3′ and 5′ termini, respectively. Higher temperatures may prevent the padlock DNA probe annealing to the target and thus prevent the formation of catenate circles. On the other hand, the shorter 44 nt padlock probe is too short to form catenate circle. Padlock probes between 52 and 60 nt in length with 11–13 nt of complementary sequences at the 3′ and 5′ termini are preferable for SNP detection. A set of universal probes and primers matching the nucleotide sequences of the backbones of the padlock probes was designed (Table 1 and Fig. 1). The designed padlock probes are aimed to detect SNPs in PCR fragments (<10 kb) rather than total genomic DNA. Therefore, shorter complementary sequences (11–13 nt) at both 3′ and 5′ termini can be applied and, subsequently, a shorter backbone is required for the formation of a catenate circle. The 34 nt backbones contain a 19 nt binding site for RCA primers and a 24 nt sequence as a hybridisation probe for the detection of RCA products, which is also used as second primers in HRCA with a 9 nt overlap between the two regions (Fig. 1).

RCA with φ29 DNA polymerase and HRCA with Vent (exo−) DNA polymerase

Detection of SNPs can be achieved by separation of the ligation products on denatured polyacrylamide gels followed by silver staining (Figs 2 and 3A). However, gel-based SNP detection is labour intensive and time consuming, and is also difficult to adapt for automation to achieve high throughput. To circumvent this step, circularised padlock probes were further amplified with one primer using φ29 DNA polymerase (Fig. 3B) or with two primers using Vent (exo−) DNA polymerase (Fig. 3F). In both cases circularised padlock probes were amplified and readily detected in the microplate wells under UV illumination after staining with Cybr-Gold (Fig. 3E and G). The Amy109OCP_C60 padlock probe yielded a signal with S75 (GG) and the mixture S75/mutant 109 (GA), and Amy109OCP_T60 with the mixture S75/mutant 109 (GA) and mutant 109 (AA), indicating that the two alleles were correctly discriminated.

Linear amplification of the allele-specific circularised padlock probes using φ29 DNA polymerase produced long single-stranded DNA strands that tended to remain in the loading wells of a 0.6% agarose gel after electrophoresis (Fig. 3B). Southern hybridisation with OCP_C_H and OCP_T_H hybridisation probes confirmed that the amplified DNA was derived from the allele-specific circularised padlock probes (Fig. 3C and D). However, HRCA with Vent (exo−) DNA polymerase produced smaller DNA fragments and showed less specificity in allele discrimination (Fig. 3F).
Biallelic assay with complementary two-colour fluorescent dye-labelled probes

The L-RCA SNP detection method has potential for detecting multiple alleles or loci in a homogeneous assay. RCA primers were attached to a biotin–bead complex and the two hybridisation probes used for detection, OCP_C_H and OCP_T_H, were labelled with the fluorescence dyes Cy3 and Cy5, respectively. Signals were detected by scanning on the Typhoon 8600 system. Homozygous G and A alleles and the heterozygous GA mixture were clearly distinguished (Fig. 4).

Genotyping of SNPs in F2 populations using L-RCA

Having established the methodology, we applied it to genotype F2 progeny derived from crosses between the wild-type A.strigosa accession S75 and each of the two Sad1 mutants 109 and 610. Eighty-three F2 progeny from a cross between the wild-type S75 and mutant 109, and 96 F2 progeny from a cross between S75 and mutant 610, were assessed for the saponin-deficient phenotype by screening seedlings for the presence or absence of fluorescence in the root tips as described in Papadopoulou et al. (26). SNP analysis using the developed L-RCA method was conducted in both populations. Of the 61 F2 progeny derived from the cross S75 × 109 that contained saponins in the root tip, 44 were heterozygous GA, and 17 were homozygous GG. The 22 saponin-deficient individuals were homozygous AA (Fig. 5). Analysis of the 96 F2 individuals derived from the cross S75 × 610 showed that 14 of 66 wild-type individuals were homozygous GG and 52 were heterozygous GA. All 30 saponin-deficient F2 individuals were of the homozygous AA genotype (data not shown). All SNP scores were confirmed using mini-sequencing based on single primer chain extension (10).

DISCUSSION

We have demonstrated that SNPs can be efficiently detected using the L-RCA method and simple UV detection. The use of short padlock probes (52–60 nt) and a set of generic amplification primers and hybridisation probes reduces the cost and simplifies the detection. Although long padlock probes can be obtained by PCR (30,31), direct synthesis of short padlock probes (52–60 nt) is simpler and more cost effective. All experimental procedures including PCR amplification, thermostable ligation, RCA and final signal detection can be conducted in the same tube or microplate well using appropriate buffers. The methodology can be readily applied to the analysis of multiple samples using 384-well microplates and
liquid handling systems. Introduction of PCR amplification of target DNA has a number of advantages. First, it narrows down the target regions so that short padlock probes with short complementary 3′ and 5′ termini can be used. Secondly, the concentration of target DNA is substantially increased after PCR, so only a small amount of genomic DNA (e.g. ≤50 ng of oat genomic DNA) is required per assay. Thirdly, use of the asymmetric PCR can normalise the initial targets when a high-throughput DNA isolation method is used and measurement and adjustment of DNA concentrations are impracticable.

Standardised reaction conditions must be used and both positive and negative controls must be included in the experiment in order to achieve reliable allelic discrimination. In the experiments described here, DNA samples were prepared using the Qiagen DNeasy 96 Plant Kit, and aliquots of 5 μl were directly used as templates in asymmetric PCR. The normalised PCR products are crucial for achieving uniformity of signals and reliable allelic discrimination.

Use of the amplified DNA as reporter and detection by UV illumination do not allow the simultaneous assay of multiple alleles in a single tube. However, homogeneous reactions in PCR and ligation can be conducted. In the case of detection of two alleles, homogeneous PCR amplification and subsequent splitting of the products for separate ligation of allele-specific padlock probes will help to identify the PCR failures, which are probably the major causes of false-negative errors in this detection system. Simultaneous ligation of two allele-specific padlock probes can also be carried out, followed by separate RCA or HRCA. Our experiment indicates that RCA with 32P-labeled DNA polymerise shows higher allelic specificity in signal amplification than HRCA with Vent (exo−) DNA polymerase.

The L-RCA SNP detection method is sensitive enough to be applied directly to genomic DNA (24). Using the circularised probes as templates, the signals can be amplified several thousand-fold in a few hours by RCA or HRCA. However, we were unable to detect the SNP within the AsbAS1 gene of the oat lines S75, 109 and 610 by L-RCA using long padlock probes without prior amplification of the target sequences. This may be because the circularised probes are not released from the long target stands, resulting in inhibition of the RCA (32). On the other hand, the UV-detection system is not sensitive, and the template, padlock probes and primers will result in high background signal. To achieve a high signal-to-noise ratio, a substantial amount of amplification product is required. A preamplification step achieved this objective.

We have shown that by hybridisation of the RCA products with fluorescence dye-labelled probes, biallelic assay has been achieved. Moreover, use of fluorescent dyes as reporter systems will improve sensitivity and specificity, enabling direct use of genomic DNA in SNP detection. Multiplexed assay for several loci or alleles in the whole genome can be achieved with the L-RCA method followed by hybridisation of the RCA products with multiple fluorescence dye-labelled probes and detection using appropriate wavelength laser sources and filters in a comprehensive fluorescence system.

ACKNOWLEDGEMENTS

We are indebted to M. Salas for kindly providing φ29 DNA polymerase. This work was supported by grants from the Biotechnology and Biological Sciences Research Council (BBSRC) and the Department for International Development (DFID), UK. The Sainsbury Laboratory is supported by the Gatsby Charitable Foundation and the John Innes Centre is grant aided by the BBSRC.

REFERENCES


