Transgenics, Vol. 4, pp. 151-156
Reprints available directly from the publisher
Photocopying permitted by license only

Rapid Cloning of Novel Genes and Promoters for Functional Analyses in Transgenic Cells

PEER M. SCHENK[1,2,*], CLAUDIA E. VICKERS[2,3] AND JOHN M. MANNERS[1,3]

[1]Cooperative Research Centre for Tropical Plant Protection, Level 5 John Hines Building, [2]Department of Botany, School of Life Sciences, The University of Queensland, St. Lucia, QLD 4072,

[3]CSIRO Plant Industry, 306 Carmody Rd, St Lucia QLD 4067, Australia

The availability of sequence information for thousands of genes for many organisms is currently unmatched by functional studies. A cost-effective and high-throughput cloning system for PCR products was therefore adopted to enable the rapid assessment of coding and promoter sequences in functional assays in transgenic cells. Unlike other systems that involve initial cloning into a specialized PCR fragment cloning vector, this method describes a rapid and cost-effective procedure for the amplification of a DNA fragment by PCR, its phosphorylation and its direct insertion into the vector of choice. Restriction enzymes are only required once for the preparation of the recipient vector, which is blunt-ended and dephosphorylated. No special primer designs (e.g. restriction enzyme sites or flanking homologous sequences) or subcloning steps are required. The turn-around time from source organism genomic DNA to new recombinant DNA is 24 hrs. It is particularly suitable for functional genomics projects or the generation of libraries from PCR products where a large number of fragments need to be cloned into the same vector. We have used this method to rapidly clone 72 full-length genes (ranging from 0.8 to 6.4 kb) and putative promoters (2 kb each) from *Arabidopsis thaliana* into plant cell expression cassettes for subsequent direct functional analyses in transgenic cells.

Key Words: amplification, functional genomics, gene mining, library construction, rapid cloning, transgenic expression

INTRODUCTION

The recent availability of large data sets of sequence information (e.g. large-scale EST and whole genome sequencing projects) is

*Corresponding author: E-mail: p.schenk@uq.edu.au; Fax: +61-7-33654771; Phone: +61-7-33658817

currently unmatched by data on their actual function. DNA microarray hybridisations and *in silico* sequence analyses have led to the identification of a large number of candidate genes and regulatory sequences that are likely to be involved in a given biological process (e.g. [1,2,3,4]). Further screening and functional analyses are therefore required to confirm and further elucidate the roles of these sequences. One of the key procedures that help to reveal the biological function of uncharacterised DNA sequences is the expression of those sequences in transgenic organisms. This can lead to interesting phenotypes displaying visible modifications or molecular alterations. In most cases, full-length coding sequences and regulatory sequences (such as promoters) of these genes are analysed in different heterologous expression systems, each requiring subcloning steps. Rapid, robust and cost-effective cloning procedures are therefore required to cope with the large number of candidate genes and regulatory sequences.

Many methods for cloning PCR products require the use of a special intermediate cloning vector where PCR products are inserted into a selectable marker (such as the lacZ gene, which utilises blue/white selection) or a lethal gene that will stop growth of false positives (e.g. [5,6,7,8]). For these methods, additional sub-cloning steps are required to insert the PCR product into the target vectors. Alternatively, PCR products can be inserted into the vector by adding suitable flanking restriction enzyme sites to the PCR primers and digesting the PCR product after amplification (e.g. [9]). A prerequisite for this method is the absence of these sites in the amplified DNA sequence. Other flanking sequences allow for homologous recombination events that lead to the integration of the PCR product [10,11,12]. However, the introduction of these additional flanking sequences may lead to unwanted side effects during transcriptional or post-transcriptional regulation. High costs associated with commercially available cloning kits often limit their use in high-throughput approaches. In this paper we present a rapid and very cost-effective method for direct cloning of PCR-amplified gene and promoter sequences into any target vector without the need of flanking restriction enzyme sites or other added sequences.

MATERIALS AND METHODS

DNA amplification and phosphorylation of 5'ends

PCR to amplify genes and promoters from *Arabidopsis thaliana* Columbia was carried out using enzymes with proofreading activity that generate blunt ends (Expand High Fidelity PCR System, Roche). Primers were designed according to annotated Genbank sequence information including start and stop codons for genes and covering approximately 2 kb upstream of the genes for promoter sequences. PCR products were then subjected to electrophoresis and purified (Qiaquick Gel Extraction Kit, Qiagen). 5' ends of PCR products (1-5 µg) were phosphorylated in a 20 µl reaction using 2 µl polynucleotide kinase (10 U/µl; Roche), 2 µl 10x buffer (50 mM Tris-Cl pH7.5; 10 mM $MgCl_2$; 5 mM DTT) and 0.2 µl 10 mM ATP and by incubating at 37 °C for 1 h. Alternatively, phosphorylated primers were used for PCR. These were prepared prior to

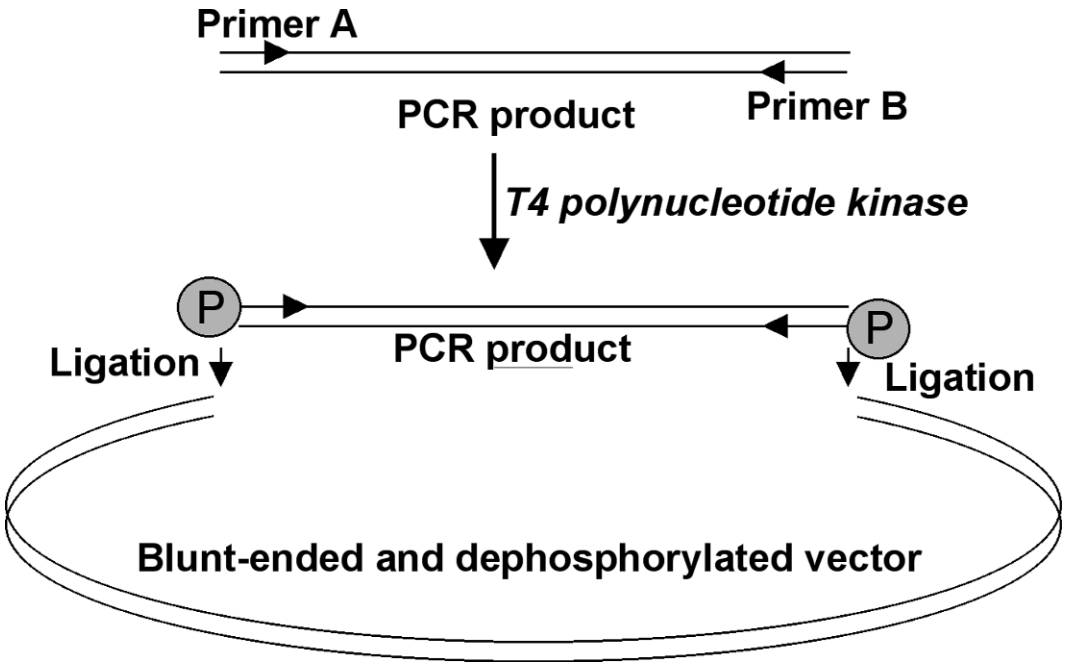

FIGURE 1
Schematic representation of the rapid cloning procedure of PCR products.

PCR by adding 0.2 µl polynucleotide kinase (10 U/µl), 1 µl 10x buffer (as above) and 1 µl 10 mM ATP to 300 pmol primers in a 10 µl reaction followed by incubation at 37 °C for 1 h and inactivation at 65 °C for 5 min.

Vector preparation

To allow for repeated use of vector fragments a large amount of plasmid DNA (> 10 µg) was prepared by cleaving with a restriction enzyme that generates blunt ends. Alternatively, if restriction enzymes were used that left overhanging 5' ends, these were subsequently filled in by adding the Klenow fragment (1U/µg DNA) and 1/10 volume 1mM dNTPs and by incubating for 15 min at room temperature. The vector fragment was then subjected to electrophoresis, purified (Qiaquick Gel Extraction Kit, Qiagen) and dephosphorylated (shrimp alkaline phosphatase, Roche) to prevent recircularisation.

The enzymes and reagents listed here have proven to work reliably for the preparation of new plasmid constructs and are suggestions only. The use of an enzyme with proofreading activity (e.g. Expand High Fidelity PCR System, Roche) is essential to reduce errors during PCR and to generate products with blunt ends. If a library is prepared or the vector will be used more than once, a large amount of vector (> 10 µg) should be used.

Ligation and transformation

Vector and PCR product were purified (Qiaquick PCR Purification Kit, Qiagen),

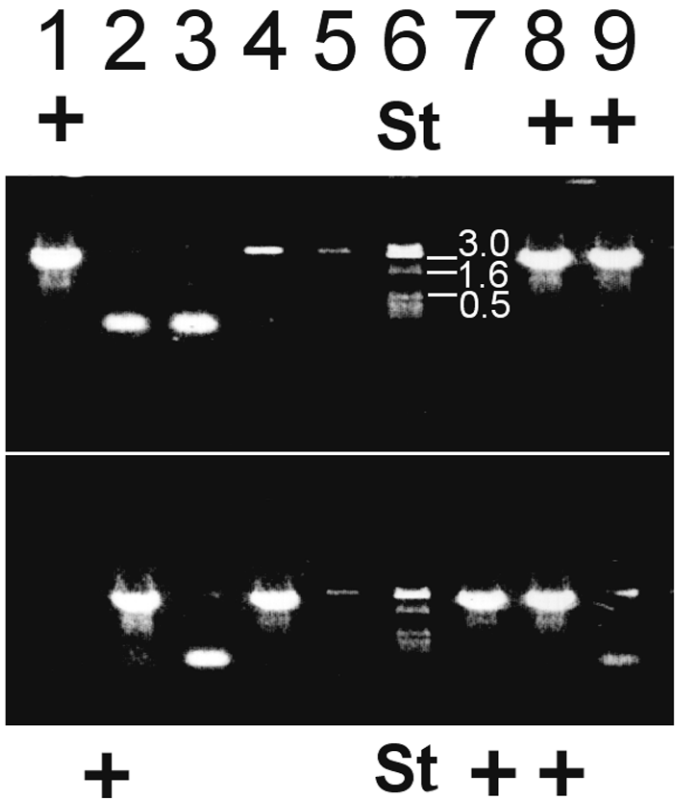

FIGURE 2
DNA agarose gel electrophoresis showing a typical result from screening of 16 bacterial colonies for the presence of a construct containing a PCR amplified *Arabidopsis* gene in the correct orientation cloned into a plant expression vector downstream of the CaMV 35S promoter. A forward primer at the 3' end of the CaMV 35S promoter was used in combination with the reverse primer of the amplified gene. Seven bright bands with the expected size of 4 kb indicate positive clones (+). Other bands typically present are primer dimers (e.g. lane 2 and 3) or weak bands of the right size which indicate false positives that are usually caused by traces of the ligation mix present on the bacterial plate (e.g. lane 4 and 5). St = 1 kb ladder (Invitrogen).

vacuum-concentrated and ligated (Rapid DNA Ligation Kit, Roche). Typically, only a small amount of vector (200-500 ng) and up to ten times more insert DNA were used for the ligation. Chemically competent *E. coli* cells (OneShot Top10, Invitrogen) were used for transformation following manufacturer's instructions. Plasmids were prepared from liquid cultures from five colonies for each construct and were screened by sequencing for the presence and correct orientation of the inserted sequence. Alternatively, colonies were screened directly by PCR using a primer within the inserted sequence together with a primer in the flanking vector sequence.

RESULTS

The availabity of sequence information from large DNA databases makes it possible to specifically design primers targeted to a large number of sequences that are of particular interest. We have developed a rapid and cost-effective way to directly clone and functionally analyse novel gene and

promoter sequences, without the addition or modification of flanking nucleotides (see Figure 1 and Materials and Methods section). To test whether this method can be used in a simple, robust and efficient way to simultaneously analyse a large number of sequences, we have used sequence information that was obtained from microarray analyses. Previously, we have studied transcriptional responses in *Arabidopsis thaliana* by cDNA microarray analysis and we have identified a set of 705 genes that were up- or down-regulated as a result of either pathogen attack or treatment with plant defense signaling compounds [3]. Based on specific expression profiles and cluster analyses, a subset of 72 potential key regulatory genes and promoters of special interest was selected for subsequent functional studies. These were amplified in volumes of 50 µl each using genomic template DNA (10-50 ng), previously isolated from *Arabidopsis* leaves following the CTAB method [13] and then cloned into three different plant expression vectors to test the above cloning protocol on a large scale. Putative promoter sequences from *Arabidopsis* genes were cloned into pGreenII binary vector derivatives [14] using either the *uid*A reporter gene [15] or the synthetic green fluorescent protein reporter gene s*gfp*(S65T) [16] fused to the *Agrobacterium tumefaciens* nos terminator. The plasmid p35SERFI [17], a derivative of pBI221 (Stratagene) containing the cauliflower mosaic virus (CaMV) 35S promoter [12], the *ERF-I* gene [18] and the nos terminator, was used to clone full-length coding sequences from *Arabidopsis* by replacing the *ERF-I* gene. Screening of bacterial colonies by PCR and sequencing revealed that on average 88 % of the tested colonies contained the correct PCR product with half of these in the desired orientation (see Figure 2 for example). All cloned coding sequences from *Arabidopsis* were functionally assessed for the presence of cell signaling activity by using a transient expression system in transgenic *PDF1.2* promoter-reporter plants described previously [17]. All putative promoter-reporter constructs were tested in transient expression assays in *Arabidopsis* leaves and/or stable transformation of *Arabidopsis* plants and lead to functional expression of the reporter transgenes (data will be described elsewhere). This allowed a rapid assessment of candidate gene and promoter sequences and demonstrated the functionality of the generated gene constructs.

DISCUSSION

It can be expected that rapid functional characterisation of genes and their regulatory sequences will be essential for the overall understanding of complex biological processes. The availability of sequence data makes it possible to amplify a large number of sequences from cDNA or genomic DNA. These can be directly cloned into suitable vectors for functional analysis using the method described here. The simple step of PCR product phosphorylation [18] in combination with a rapid and direct cloning system offers a rapid and cost-effective alternative to other commonly used DNA cloning procedures. A significant reduction of cloning costs is achieved (approximately US $ 5-10; calculated per PCR product based on a total of 100 reactions) in comparison to commonly used modern methods for cloning of PCR products, such as cloning kits with

supplied cloning vectors (approximately US $ 20-25) or systems that use homologous flanking sequences for recombination (approximately US $ 40-50). In addition, products can be directly inserted into any position of the target vector where no flanking artificial sequences need to be introduced. Potential large-scale automation of this method could be established by using software-assisted primer design and gene identification of annotated sequence. PCR amplification, phosphorylation, ligation and transformation steps can be carried out in 96-well microtiter plates (96-well columns for DNA purification steps). For this purpose the electrophoresis step could be replaced by a simple purification step, and the correct sequence could be confirmed alternatively by direct sequencing of bacterial colonies.

Acknowledgements

Part of this research was supported by the Grains Research and Development Corporation (GRDC) of Australia (to J. Manners). The authors wish to thank Drs Kemal Kazan, Paul Ebert, Tony Remans, Jonathan Anderson and Kenneth McGrath for independently testing of the method and useful discussions.

REFERENCES

1. Alon UN, Barkai N, Notterman DA, Gsh K, Ybarra S, Mack D, Levine AJ. Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. Proc Natl Acad Sci U S A 1999;96:6745-6750.
2. Odell JT, Nagy F, Chua N-H. Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. Nature 1985;313:810-812.
3. Schenk PM, Kazan K, Wilson I, Anderson JP, Richmond T, Somerville SC, Manners JM. Coordinated plant defense responses in Arabidopsis revealed by microarray analysis. Proc Natl Acad Sci U S A 2000;97:11655-11660.
4. White KP, Rifkin SA, Hurban P, Hogness DS. Microarray analysis of *Drosophila* development during metamorphosis. Science 1999;286:2179-2184.
5. Bernard P, Kezdy KE, Van Melderen L, Steyaert J, Wyns L, Pato ML, Higgins PN, Couturier M. The F plasmid CcdB protein induces efficient ATP-dependent DNA cleavage by gyrase. J Molec Biol 1993;234:534-541.
6. Brandt ME, Gabrik AH, Vickery LE. A vector for directional cloning and expression of polymerase chain reaction products in *Escherichia coli*. Gene. Gene 1991;97:113-117.
7. Chaffin DO and Rubens CE. Blue/white screening of recombinant plasmids in Gram-positive bacteria by interruption of alkaline phosphatase gene (phoZ) expression. Gene 1998;219:91-99.
8. Zhou MY and Gomez-Sanchez CE. Universal TA cloning. Curr Issues Mol Biol 2000;2:1-7.
9. Patel M, Johnson JS, Brettell RIS, Jacobsen J, Xue G-P. Transgenic barley expressing a fungal xylanase gene in the endosperm of the developing grain. Mol Breed 2000;6:113-124.
10. Boyd A. Turbo cloning: a fast, efficient method for cloning PCR products and other blunt-ended DNA fragments into plasmids. Nucl Acids Res 1993;21:817-821.
11. Hartley JL, Temple GF, Brasch MA. DNA cloning using in vitro site-directed recombination. Genome Res 2000;10:1788-1795.
12. Ohara O and Temple G. Directional cDNA library construction assisted by the *in vitro* recombination reaction. Nucl Acids Res 2001;29:E22-
13. Graham G, Mayers P, Henry RJ. A simplified method for the preparation of fungal genomic DNA for PCR and PAPD amplification. BioTechniques 1994;16:48-50.
14. Hellens RP, Edwards EA, Leyland NR, Bean S, Mullineaux PM. pGreen: a versatile and flexible binary Ti vector for *Agrobacterium*-mediated plant transformation. Plant Mol Biol Rep 2000;5:387-405.
15. Jefferson RA. Assaying chimeric genes in plants: The GUS gene fusion system. Plant Mol Biol Rep 1987;5:387-405.
16. Chui WL, Niwa Y, Zeng W, Hirano T, Kobayashi H, Sheen J. Engineered GFP as a vital reporter in plants. Curr Biol 1996;6:325-330.
17. Kazan K, Schenk PM, Wilson I, Manners JM. DNA microarrays: new tools in the analysis of plant defence responses. Mol Plant Pathology 2001;
18. Solano R, Stepanova A, Chao Q, Ecker JR. Nuclear events in ethylene signalling: a transcriptional cascade mediated by ETHYLENE-INSENSITIVE 3 and ETHYLENE RESPONSE-FACTOR 1. Genes Dev 1998;12:3703-3741.