African Journal of Biotechnology Vol. 8 (19), pp. 4765-4773, 5 October, 2009
Available online at http://www.academicjournals.org/AJB
ISSN 1684–5315 © 2009 Academic Journals

Full Length Research Paper

Random amplification of genomic ends (RAGE) as an efficient method for isolation and cloning of promoters and uncloned genomic regions

Boney Kuriakose[1,3]*, V. Ganesan[1,4], George Thomas[2], Arun Viswanathan[1,5] and Narayanaswamy Anand[1]

[1]Centre for Advanced Studies in Botany, University of Madras, Guindy Campus, Chennai -600 025, India.
[2]Interfield Laboratories, Karuvelipady, Cochin - 680 005, India.
[3]Floriculture lab, Phytotron D, University of Pretoria Experimental farms, South Street, Hatfield, Pretoria, 0002, South Africa.
[4]Acme ProGen Biotech (India) Private Limited, 260G, Ram Square, Advaitha Ashram Road, Salem -636 004, Tamil Nadu, India.
[5]Department of Biotechnology, Sri Ramachandra University, No.1, Ramachandra Nagar, Porur, Chennai - 600 116, India.

Accepted 23 June, 2009

Isolation of complete coding sequences and regulatory regions is critical for the complete characterization of a gene. Efficient methods to obtain complete genomic or regulatory is important in the process of isolation. The utility of the available genome walking methods are influenced by factors like the size of the genome and the length of the desired sequence. This study utilizes a genome walking method - random amplification of genomic ends (RAGE) efficiently to obtain the 5' – regulatory sequence of a rice stress inducible gene *OsAsr1* and to obtain the full length sequence and promoter of the *HetR* gene of *Cylindrospermum stagnale* (*Cylindrospermum* sp. A1345). We demonstrate that this technique can be used for cloning of full length gene and promoters in organisms where whole genome data is unavailable utilising very little sequence information. Our studies show that RAGE can be a strong tool in functional genomics especially in the study of promoters

Key words: RAGE, genome walking, promoter, *Oryza rufipogon*, *Oryza malampuzhaensis*, *Anabaena*, *Cylindrospermum*, *Asr1*, *HetR.*

INTRODUCTION

The isolation of coding and regulatory genomic sequences is an important part in the study of any gene. The genomic regions flanking the coding sequence contain important information about the character of the gene and influence its expression. The 5' sequence contains important regulatory elements and structurally is composed of the promoter region. The eukaryotic genes contain untranslated (UTR) regions at both the 5' and 3' region which are important for their expression. Isolation of promoter regions and flanking regions of genes in humans, plants like rice and *Arabidopsis* and some other organisms has been made easier by the availability of the sequenced genome. While complete genomes are available for many commercially important organisms and a number of model organisms, they are still not available for most other organisms like cyanobacteria. In these organisms, the isolation of genes and promoters are still a tedious and sometimes complex process. The use of heterologous primers to screen organisms for similar sequences is one approach to clone these sequences. Many of these genes can have similar coding sequence but entirely different regulatory regions. For example, the polyubiquitin is a family of genes each having different lengths of coding sequences depending on the number of repeats of the ubiquitin coding sequence (Callis et al.,

*Corresponding author. Email:boneykuriakose@gmail.com. Tel: +27 792709391. Fax: +27 12 844 0276.

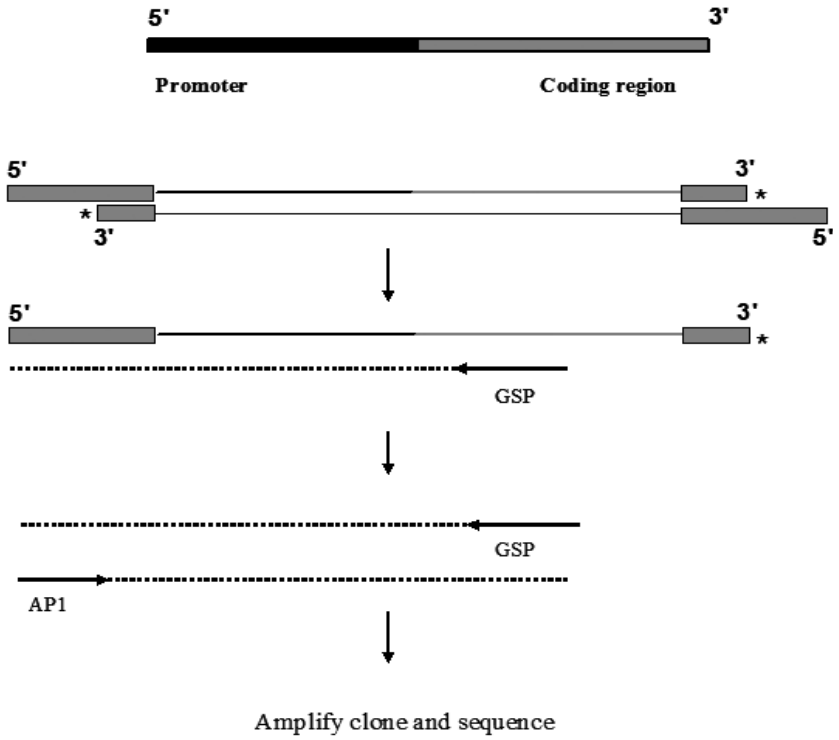

Figure 1. A Schematic representation of the use of RAGE for isolation of 5' genomic region and promoters. GSP is designed to bind to top strand. GSP = Gene specific primer, AP = adapter specific primer, * = amino blocking.

1995). Each of these different genes within a family can have totally different promoters (Wei et al., 2003). In such cases, the use of heterologous primers for amplification of the full length sequence including its regulatory regions is ruled out using conventional PCR. There are also instances where isolation is approached by primers designed at the conserved regions and in these cases full length sequence of the gene(s) is not always obtained. Even in prokaryotes with relatively simple genomes and gene sequences, the isolation of genes could be difficult with the above hindrances. It would be thus ideal to have efficient tools to isolate genes with regulatory sequences if partial genomic sequence or cDNA sequence is known.

Several techniques have experimentally evolved for the purpose of cloning flanking sequences and walking on genomic DNA. In the inverse PCR the restricted DNA molecules are self ligated prior to PCR (Ochman et al., 1988). In the ligation mediated PCR, the restricted DNA is ligated to a DNA cassette and then PCR is carried out. These techniques were limiting in their specificities and so several modified procedures were developed. In another method, after ligation, a single strand is extended with a single biotin labeled primer, streptavidin coated magnetic beads were used for separating the specific fragments before PCR amplification was carried out (Rosenthal and Jones, 1990). Zhang and Chiang (1996) used a single gene specific primer to extend the desired product. A known oligonucleotide was ligated to this single stranded product using RNA ligase and this was subsequently amplified. The TAIL PCR (thermal asymmetric interlaced PCR) method used a long specific primer and small non-specific degenerate primers for amplification. The cycling program was designed with a low stringency cycle followed by two high stringency cycles (Liu and Whittier, 1995; Liu et al., 1995). This was used for cloning the *Pal* and *Pgi* promoters of yam (Terauchi and Kahl, 2000). The suppression of non-specific products in a ligation mediated PCR can be achieved by the careful design of the ligated DNA (Siebert et al., 1995). Kits such as the Genome Walker Kit (Clontech, Palo Alto, Calif.) have been used earlier in the cloning of the 5' regulatory sequences of genes of human and plant genes (Siebert et al., 1995; Filichkin et al., 2004; Gong et al., 2005; Hong and Hwang, 2006).

Here, we describe the extension of this technique for isolation of two stress inducible promoters from *Oryza rufipogon* (Accession no. DQ885913), *Oryza malampuzhaensis* (Accession no. DQ885914), wild relatives of rice using sequence information from its homologous gene *OsAsr1* in rice. We also used the technique in a prokaryotic system for cloning the full length *HetR* gene from a partial genomic clone of a *Cylindrospermum* species (Accession no. DQ439538). See Figures 1 and 2 for schematic description of the technique.

We also used relatively simple assay procedures to test the function of the plant promoter using transient

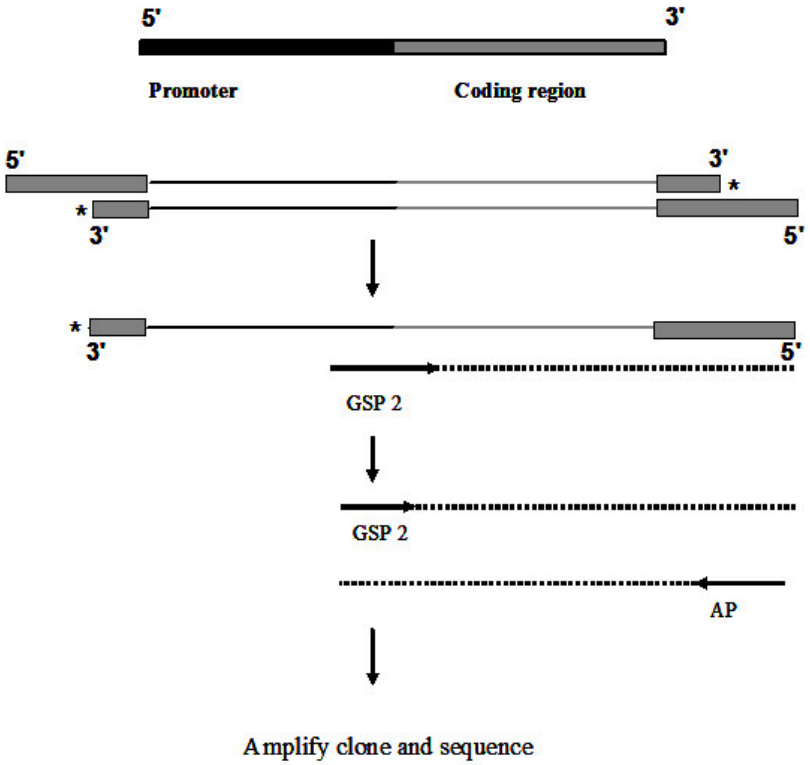

Figure 2. A Schematic representation of the use of RAGE for isolation of 3' ends of genes. Here GSP2 binds to the bottom strand. GSP = Gene specific primer, AP = adapter specific primer, * = amino blocking.

assays. A transient expression study does not give an accurate and conclusive picture of the expression pattern in the plant finally, but it could be used as a quick tool for assessing the functionality of an isolated promoter sequence, especially when the plant analyzed takes several weeks or months for growth and development (Dekeyser et al., 1990). Earlier, either electroporation (Fromm et al., 1985) or PEG treatment (Shillito et al., 1985) were used for the transformation of the DNA to plants. The presence of these DNA molecules in the cell as extra chromosomal molecules reduces the interaction of the sequences with chromosomal elements (Werr and Lorz, 1986). This reduces the so called "position effects" and "gene silencing" in these experiments in comparison to stable transformants. Later biolistic and *Agrobacterium* mediated transient assays were also developed for the purpose. Transient assays have been used extensively for promoter analysis in model plant systems like tobacco (Yang et al., 2000) and in other plants (Wroblewski et al., 2005).

Our studies of using the same tool on two distantly related organisms- a higher plant *O. rufipogon* and a cyanobacterium *Cylindrospermum stagnale* demonstrate that RAGE is an efficient tool for the isolation of promoters. It can also be used for cloning full length genomic sequences where information is available for only a part of the gene.

MATERIALS AND METHODS

Isolation of genomic DNA

Genomic DNA was isolated from *O. rufipogon* and *O. malampuzhaensis* using the CTAB method by Rogers and Bendich (1994). Genomic DNA from *Cylindrospermum stagnale* cells was isolated using the method by Chen and Kuo (1993). The DNA was electrophoresed on 0.8% agarose gels in 0.5 X TBE buffer to check intactness and was quantified by UV spectrophotometer at 260 nm.

Digestion of genomic DNA

Genomic DNA from *O. rufipogon* and *O. malampuzhaensis* (5 µg) and *Cylindrospermum* (2 µg) was digested with different blunt cutting restriction enzymes *Dra* I, *EcoR* V, *SnaB* I, and *Ssp* I (all enzymes were from MBI Fermentas). The reactions were carried out in 100 µl volumes with 50 units of enzyme in the appropriate buffers and incubated overnight at 37°C. The digested DNA was analyzed on 1% agarose gel in 0.5X TBE buffer.

Ligation of adaptors

The digested DNA was purified by extraction with a mixture of phenol : chloroform (1:1 ratio) and precipitated with 95% ethanol. After washing with 70% ethanol, the DNA was air dried and dissolved in sterile deionised water. Adapters were synthesized commercially (Metabion GmbH, Germany), based on the method by Siebert et al., 1995. The individual adapter strands that were synthesized were as follows. Long strand 5'- CTAATACGACTCACAT

AGGGCTCGAGCGGCCGCCCGGGCAGGT-3' and short strand - 5' ACCTGCCC -NH_2-3'. The individual strands were dissolved individually and annealed to obtain the adapter solution. The ligation reaction contained digested DNA (5 µg), adapter 5 µM, 1X ligase buffer (MBI Fermentas) and ligase 10 units (MBI Fermentas). The reaction mixture was incubated at 22°C overnight.

RAGE- PCR reaction

The RAGE-PCR for wild rice was carried out using gene specific ASRGSP1 (5'-GAATGCGGCCGCCGTAGGAGTCTACTCCGGTG G-3') (custom synthesized at Sigma Genosys, Bangalore, India) and an Adaptor Specific ADAPrimer1 (5'- GGATCCTAATACG ACTCACTATAGGGC-3') as the primers. The ligation mix was diluted in 1:10 ratio with sterile deionised water. 1 µl of the diluted mix was used as template in a 50 µl reaction containing the following components: 0.4 µM of the primers (Metabion GmbH, Germany), 200 µM dNTPs (Amersham Biosciences), 1X XT-20 buffer (Bangalore Genei, India) and 2 units of XT-20 PCR Polymerase (Bangalore Genei, India). The reactions were carried out under the following conditions: initial denaturation 94°C for 4 min, denaturation at 94°C for 1 min followed by 35 cycles of a single step annealing and extension of 68°C for 3 min and final extension of 68°C for 7 min. The PCR products were analyzed on 1% agarose gel in 0.5X TBE buffer.

The PCR product was diluted 1:50 and 1 µl was used as template for a 25 µl reaction using nested primers ASRGSP2 (5'–CCTTCTTGTGGTGGAACAG-3') and (ADAPrimer2 (5'- AATAGG GCTCGAGCGGC–3') (Metabion GmbH, Germany). The reaction contained 0.2 µM of the primers, 200 µM dNTPs (Amersham Biosciences), 1X Taq polymerase buffer (Bangalore Genei, India) and 1 unit of Taq DNA polymerase (Bangalore Genei, India). The reactions were carried out in the following conditions. Initial denaturation 94°C for 4 min, 25 cycles of denaturation 94°C for 1min, annealing 60°C for 30 sec and extension 68°C for 1 min 30 s. Final extension of 68°C for 5 min was added. The PCR products were analyzed on 1% agarose gel in 0.5X TBE buffer.

The RAGE-PCR for isolating *C. stagnale HetR* promoter and 5' region was carried out using the primers HET GSP1A (5'-GATCTGATCCATCGCACTGGGGYCAAGAC-3') (Sigma Genosys, Bangalore, India) and ADAPrimer1. The ligation mix was diluted in 1:10 ratio with sterile deionized water. 11 µl of the diluted mix was used as template in a 50 µl reaction containing the following components: 0.4 µM of the primers, 200 µM dNTPs (Amersham Biosciences), 1X XT-20 buffer (Bangalore Genei, India) and 2 units of XT-20 PCR polymerase (Bangalore Genei, India). PCR conditions were as follows: initial denaturation 94°C for 4 min, 30 cycles of 94°C for 1 min, annealing 65°C for 1 min and extension 68°C for 3 min. Final extension of 68°C for 10 min was added. The PCR products were analyzed on 1% agarose gel in 0.5X TBE buffer.

The PCR product was diluted 1:50 and 1 µl was used as template for a 25 µl reaction using nested primers HET GSP1B (5'–GCACTGGGGCCAAGAC-3') (Sigma Genosys, Bangalore, India) and ADAPrimer 2. The reaction contained 0.2 µM of the primers, 200 µM dNTPs (Amersham Biosciences), 1X Taq polymerase buffer (Bangalore Genei, India) and 1 unit of Taq DNA polymerase (Bangalore Genei, India). The reactions conditions were denaturation at 94°C for 4 min, followed by 30 cycles of denaturation at 94°C for 30 s, annealing 45°C for 30 s and extension 68°C for 45 s. Final extension of 68°C for 5 min was added. The PCR products were analyzed on 1% agarose gel in 0.5X TBE buffer.

The RAGE-PCR for isolating the *C. stagnale HetR* 3' region was carried out under the same conditions except for the primers used. The gene specific primer used for the primary RAGE PCR reaction was HET GSP2A (5'-CGTACCTGGGCAGATAAATATCACCAAG-3') and the nested primer used was HET GSP2B (5'-GATAAA TATCACCAAG-3') (both custom synthesized by Sigma Genosys, Bangalore, India).

Cloning of the PCR products

The amplified products obtained from the secondary nested PCR reaction were purified from the gel using Perfectprep® Gel Cleanup kit (Eppendorf AG, Germany). About 50 ng of the purified product was used for ligation in a 10 µl reaction using InsT/Aclone™ PCR Product Cloning Kit (MBI Fermentas). The ligation reaction was performed at 22°C overnight. The reaction mixture was transformed into competent *Escherichia coli* DH5α cells. The colonies obtained were screened by plasmid mini prep extraction (Sambrook et al., 1989).

Sequencing of the clones and sequence analysis

The clones were sequenced commercially at the Department of Biotechnology, National Sequencing Facility, University of Delhi South Campus, New Delhi, India using M13 primers. Alignments of the forward and reverse sequences were carried out using Bioedit software (Hall TA, 1999) and full length sequence was reconstructed. The plant promoter analysis was carried out online using PLACE tool (Higo et al. 1999). Restriction analysis of large DNA sequences which were used for *in silico* analysis was carried out using the online tool NEB Cutter v 2.0 (http://tools.neb.com/NEBcutter2).

***Agrobacterium* mediated transformation of tobacco leaf discs**

The putative promoter was cloned into a plant binary vector pCAMBIA 1305.1 (CAMBIA, Canberra, Australia), by replacing the CaMV 35S promoter driving the GUS gene. The plasmid pCAMBIA 1305.1 was digested with *NcoI* and SmaI to release the CaMV 35S vector. The RAGE-PCR product of putative *Asr1* promoter from *O. rufipogon* was digested with *NcoI* and ligated to the vector to obtain the construct pSBKA 301. This construct was used mobilized into *Agrobacterium tumefaciens* LBA 4404. Tobacco leaf discs were transformed according to Horsch et al. (1985).

Histochemical assays for the GUS activity

Promoter activity was analyzed by staining for GUS enzyme activity according to Jefferson et al. (1987). Tissues were first vacuum-infiltrated in a staining solution consisting of 2 m*M* 5-bromo-4-chloro-3-indolyl-β- glucuronic acid (X-Gluc), 50 m*M* sodium phosphate (pH 7.0), 10 m*M* EDTA, 1 m*M* potassium ferricyanide and 1 m*M* potassium ferrocyanide. The tissues were destained using methanol and were observed under light microscope and photographed.

RESULTS AND DISCUSSION

Isolation of the *OrAsr1* promoter from wild rice *O. rufipogon*

The rice *OsAsr1* (Accession No. AF039573) gene encodes a protein which is induced under abscisic acid, salinity and water stress (Vaidyanathan et al., 1999) and was isolated in our group from a salt tolerant coastal rice cultivar Pokkali. The tomato homologue of this gene is well studied and has been shown to be expressed during senescence and ripening (Iucern et al., 1993, Amitai-

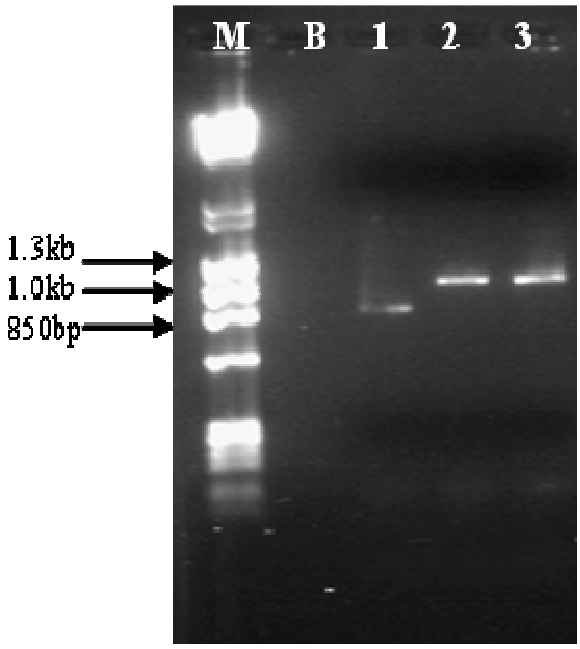

Figure 3. Os Asr1 RAGE PCR gel using primer pair GSP1 and AP1. Lane M, DNA-Hind III digest and øX174 DNA-Hae III digest (Finnzymes); lane B, Blank; lane 1, *O. rufipogon*; lane 2, *O. malampuzhaensis*; and lane 3-*O. sativa* cv. Pokkali.

Zeigerson et al., 1995). The gene product is a Zn^{2+} and sequence dependent DNA binding protein and is localized in both cytoplasm and nucleus (Kalifa et al., 2004). The isolation of the *Asr1* promoter was intended at studying in detail about the expression and the signals which drive the expression of the gene. We screened stress tolerant wild relatives of rice *O. rufipogon, O. officinalis, O. malampuzhaensis* and *O. nivara* along with the salinity tolerant cultivar Pokkali for the presence of the gene using primers designed based on Vaidyanathan et al. (1999). The gene was also present in other wild relatives of rice *O. officinalis* and *O.nivara* (data not shown). *O. rufipogon* is a highly stress tolerant wild relative of rice found distributed from South Asia to South East Asia. The common cultivated rice, *O. sativa* is thought to have evolved from the wild perennial *O. rufipogon* (Khush, 1997). This information suggested that the gene is probably conserved across these species. Using a gene specific primer designed for *OsAsr1* (ASR GSP1), a RAGE- PCR was carried out on the adapter ligated library constructed by digesting the genomic DNA with *SspI*. We obtained an 850 bp product from the library constructed using *O. rufipogon* DNA and a 1.1 Kb fragment in the *O. malampuzhaensis* DNA The Pokkali DNA which was used as a control also amplified a 1 kb product (Figure 3). The promoters were named as *OrAsrP* and *OmAsrP*. The 1.1 Kb DNA fragment contained part of the gene sequence and the upstream promoter region (Figure 3).

In silico sequence analysis of the *OrAsrP* and O*mAsrP* promoters

We analyzed the sequences using the online PLACE search tool to identify the *cis* elements on the promoter. The study revealed the presence of elements that are normal to all abscisic acid and stress inducible genes like the abscisic acid response element (ABRE) and drought response element (DRE) (Table 1). A number of *cis* elements have been defined based on *in silico* and functional analysis data on a large number of promoters and is available in the database. The tools like PLACE help in identifying these sequences based on homology searches and help to predict function of a promoter. When a promoter contains a *cis* element like the ABRE or DRE, it implies that it is expressed during stress. A combination of any of these elements with other factors would lead us to the possible nature of gene expression, which then can be ascertained by functional studies.

In case of the *OrAsrP* and *OmAsrP*, we carried out homology with the rice genome sequence and the putative promoter sequence was deposited in the Genbank. (Accession numbers DQ885913 and DQ885914).

Transient expression of the *Asr1* promoter in tobacco leaf discs

The final sequence isolated had to be functionally analyzed for confirming the promoter activity. We used an easy but effective transient expression system for this purpose. The function of the promoter was studied in transient assays using tobacco leaf discs using the construct pSBKA 301. The GUS staining assays showed the presence of promoter activity of the isolated sequence (Figure 4). The promoter after being completely characterized could be used for expression of genes which might confer salinity tolerance or other stress responses in plants.

Cloning of the promoter and 3' region of *Cylindrospermum stagnale HetR* gene

We used four different restriction enzymes (*Sca* I, *Ssp* I, *Dra* I and *SnaB* I) to digest the *C. stagnale* genomic DNA. The RAGE-PCR gave different sized fragments for each digestion (Figure 5). We cloned and sequenced the 650 bp fragment obtained with *Dra* I enzyme. The sequence was reconstructed carrying the genomic clone of *C. stagnale HetR* gene and its 5' upstream region.

Most of the previous studies with this technique were for isolation of the 5' end of the genes including promoters. We wanted to ascertain whether we could use it for cloning a missing 3' end of a gene sequence. We had previously isolated a partial *C. stagnale HetR* clone using heterologous primers designed for an *Anabaena variabilis* sequence. In order to demonstrate that the RAGE technique could also be used to clone a full length gene

Table 1. *cis* elements and motifs on the *Asr1* promoter.

Motif Name/ Factor	Position from ATG	Sequence	Predicted Function
ABRELATERD1	-611 -605 -256	ACGTG	Transcriptional regulation of ABI3- and ABA-responsive genes (Simpson et al., 2003)
ACGTATERD1	-611 -605 -587 -256 -611 -605 -587 -256	ACGT	Induction by dehydration stress and dark-induced senescence (Simpson et al., 2003)
DPBFCOREDCDC3	-705 -195	ACACNNG	bZIP transcription factors; induced by ABA (Kim et al., 1997)
DRE2COREZMRAB17	-652	ACCGAC	Transcription activators that function in drought-, high-salt- and cold-responsive gene expression (Kizis and Pages, 2002)
DRECRTCOREAT	-709 -652	RCCGAC	Dehydration-responsive element/C-repeat; drought; high-light; cold induced (Dubouzet et al., 2003)
MYCATERD1	-687	CATGTG	Dehydration response elements (Tran et al., 2004)
MYCCONSENSUSAT	-714 -687 -507 -714 -687 -507	CANNTG	MYC recognition site found in the promoters of the dehydration-responsive gene (Abe et al., 2003)
NTBBF1ARROLB	-147 -277	ACTTTA	Tissue-specific and auxin-regulated expression (Baumann et al., 1999)
PREATPRODH	-717	ACTCAT	Proline- and hypoosmolarity-responsive expression (Satoh et al., 2002)
SEBFCONSSTPR10A	-64	YTGTCWC	Similar to the auxin response element (Boyle and Brisson, 2001)
TAAAGSTKST1	-277 -264 -810 -163 -146 -138	TAAAG	Guard cell-specific gene expression (Plesch et al., 2001)
WBOXNTERF3	-926	TGACY	Wound induced expression (Nishiuchi et al., 2004)
WRECSAA01	-578	AAWGTATCA	Wound induced expression (Palm et al., 1990)
WRKY71OS	-925 -62 -36	TGAC	Binding of WRKY Transcription factors; positive and negative regulators of abscisic acid signaling (Xie et al., 2005)

when only partial sequence is available, we digested the *C. stagnale* genomic DNA with *Ssp* I, *SnaB* I, *Dra* I and *EcoR* V. After ligating the adapters, RAGE-PCR was carried out using specific primer (GSP2A) aimed at amplifying the 3' region of the gene. Of the various fragments obtained, we cloned and sequenced the 700 bp

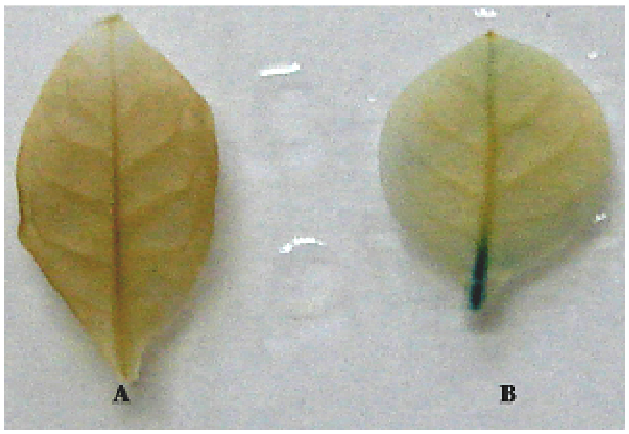

Figure 4. Transient GUS staining in tobacco plantlets. **A.** Control untransformed tobacco. **B.** Transformed leaf indicating the wounded site showing intense staining.

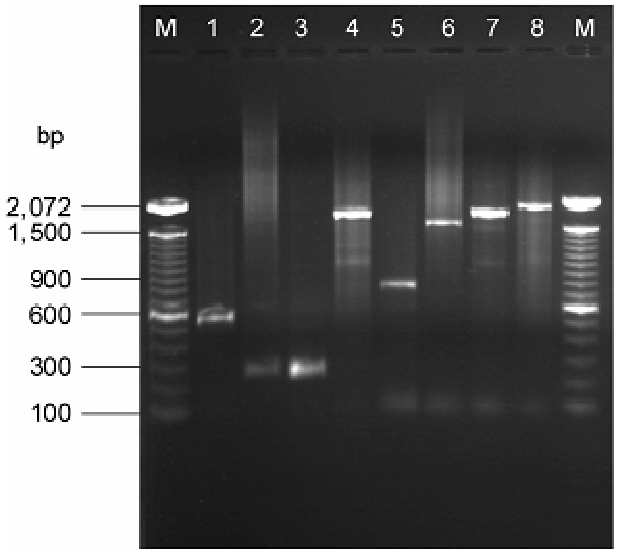

Figure 5. RAGE amplification of *Cylindrospermum stagnale* DNA digested with various restriction endonucleases. Lane M, Trackit 100 bp DNA ladder (Invitrogen); lanes 1 – 4, PCR with GSP1A (1 = DraI, 2 = EcoRV, 3 = SnaBI, 4 = SspI); lanes 5 - 8 PCR with GSP2A (5 = DraI, 6 = EcoRV, 7 = SnaBI, 8 = SspI).

product obtained in the *Dra* I digestion reaction (Lane 5, Figure 5). The clone was sequenced and the incomplete 3' region of the *C. stagnale HetR* gene was deciphered.

We were thus able to use the RAGE PCR to clone both the 5' and 3' flanking sequences of a partial genomic sequence from *C. stagnale*. The full length gene sequence was reconstructed using the RAGE PCR product sequence and the partial sequence previously isolated and deposited in GenBank (Accession no. DQ439538). The full length gene along with the 5' region of the gene could be cloned based on the sequence information available and was used for expression studies after cloning into appropriate expression systems (data not shown). This showed that RAGE used for obtaining sequences using a 3' genome walk similar to that carried out for the 5' genome walk for isolation of 5' UTRs and promoters.

Specificity of RAGE-PCR

The RAGE-PCR proves to be an efficient technique as the careful design of the gene specific primers confers high specificity to the reaction. The fool proof design of the adapter and adapter specific primer ensures that unless a specific template is available by the synthesis of a strand using the GSP, no product will be amplified. The vectorette and suppression features integrated into the design of the adapters are important for the specificity of the reaction (Siebert et al., 1995).

The ability of the RAGE-PCR to amplify products depends on the location of restriction enzyme sites on the genome. This was demonstrated efficiently in the experiment using *C. stagnale*, where a product of different length was obtained in all digestions. In experiments involving *O. rufipogon*, a higher plant, however, we were able to obtain good amplification only in two out of the five enzymes used. This may be due to the presence of too many sites for the enzymes in the genome and so the fragments obtained were too small to give a specific amplification. This was ruled out since we did not see any small sized products accumulating in the gel. Another reason could be that the sites are too few and so the effective size of the digested DNA fragment is too long to be amplified by the PCR. This reasoning could be valid since we observed non-specific streaks in some reactions (Data not shown).

We analyzed the genomes of rice and *Anabaena* in order to obtain more information. We used random stretches of sequence from Chromosome 10 and 11 from rice genome and carried out restriction analysis *in silico* using the five enzymes we had used for restriction digestion of the DNA in RAGE PCR (Tables 2 and 3). The frequency of restriction sites revealed that among the enzymes we used, *Dra* I is the most frequently cutting enzyme in the genome. *Dra* I however, failed to give rise to amplification in the reaction involving rice DNA in our experiments probably due to non-availability of a site in the close vicinity. *Ssp* I is the next most frequent cutter in the list, followed by *SnaB* I and *EcoR* V. We obtained amplification with *Ssp* I but *SnaB* I and *EcoR* V failed to amplify. The complete genome of any member of *Cylindrospermum* genus is not available and therefore we analyzed a part of the closely related organism *Anabaena* genome (*Anabaena variabilis* ATCC 29413 Accession No.-NC 007413) which is fully sequenced. In this case, frequencies of restriction sites were different for different enzymes (Table 3). We obtained amplification with all the four enzymes used for the RAGE-PCR. The *Dra* I gave the lowest size amplification in both the experiments, with *EcoR* V and *Ssp* I

Table 2. *In silico* analysis for determining the average no. of restriction sites on rice chromosome 10 and 11.

Enzyme	Chromosome 11			Chromosome 10				Average no. of sites/kb
	1-300 kb	300-600 kb	600-900 kb	1-300 kb	300-600 kb	600-900 kb	900-1200 kb	
Dra I	240	251	248	265	249	264	265	0.84
EcoRV	70	56	54	48	79	65	55	0.2
ScaI	89	89	96	95	73	110	89	0.3
SnaBI	43	53	45	65	43	59	47	0.16
SspI	254	250	221	259	248	252	232	0.81

Table 3. *In silico* analysis for determining the average no. of restriction sites on chromosome of *Anabaena variabilis* ATCC 29413.

Enzyme	*A. variabilis* chromosomal DNA			Average no. of sites/ kb
	1 - 300 KB	1 - 1.55 MB	5 - 5.5 MB	
Dra I	206	230	202	0.7
EcoRV	97	98	92	0.3
ScaI	83	54	49	0.2
SnaBI	38	36	34	0.12
SspI	368	273	274	1.0

giving larger sized fragments. The frequency of restriction sites of each enzyme on the genome seems to influence the size of products obtained in the reaction, although further studies with more organisms and with more number of enzymes will have to be carried out to substantiate the claim.

The frequency of digestions also explains why some enzymes readily give amplification products while some do not. Therefore, *in silico* studies with a number of restriction enzymes on related organisms could be helpful in designing RAGE experiments. For example, in large genomes a frequent cutter has to be used to obtain fragments that could be amplified and in organisms having smaller genome less frequent cutters could be used. This was also observed in our experiments with wild rices, where *Ssp* I digested genomic DNA library prepared with ligated adapters were able to amplify fragments in RAGE PCR whereas libraries using other enzyme digestions failed to amplify or amplified only small stretches. In another experiment using rice for isolation of an anther specific promoter, we were able to obtain amplified fragments only with *Ssp* I, although we tried other enzymes*; Eco*RV, *Hinc* II, *Pvu* II, *Stu* I and *Sma* I (Kuriakose et al., 2009). In *Cylindrospermum*, a cyanobacterium with smaller genome we were able to obtain products with all the enzymes we tried. Our studies suggest that frequent cutters like *Dra* I and *Ssp* I could be used when shorter fragment like the missing portion of a gene has to be cloned. In case of longer fragment like regulatory sequences or full length genes less frequent cutters could be used. Siebert et al. (1995) used adapter ligated libraries for different enzymes and screened them. We suggest a similar approach except that the screening of libraries may be based on the information generated by *in silico* restriction analysis of the same or related organisms and decide– whether to walk short distance.

We also observed that the digestion of DNA was also a very important factor. For this the quality of the extracted DNA has to be of good quality. The DNA from different plants can be of varying quality depending on the plant and the growth conditions. The best method available for that plant DNA extraction has to be used. Our experiments showed that the classical CTAB method (Rogers and Bendich, 1994) was the best for rice and related plants. However, for other plants the best method available for each species should be used.

Conclusion

Our studies demonstrate that the RAGE-PCR can be an easy and efficient tool for isolating the uncloned 5' or 3' regions of a known sequence. This can be a valuable tool for isolation of promoters of genes from organisms whose genomes have not been sequenced. This has been demonstrated in this study with the wild relatives of rice, where information from a heterologous sequence was used to isolate the *O. rufipogon* and *O. malampuzhaensis Asr1* promoter. However, the technique is limited by the complexity of the genomes especially the size. The presence of multiple copies of target gene or if target gene belongs to a family of homologous genes could also lead to complex results. In such cases the design of the gene specific primer has to be with utmost caution to avoid mispriming from related sequences. In relatively simple prokaryotic genomes like *Cylindrospermum*, these

may not be limiting factors as we obtained amplifications with all the four different enzymes that we used demonstrating the efficiency of the RAGE-PCR as a strong tool in genomics.

ACKNOWLEDGEMENTS

BK and GV contributed equally to this work. BK and AV received the Senior Research fellowship from Council of Scientific and Industrial Research (CSIR), Government of India and GV was a recipient of a research fellowship from Ministry of Environment and Forest (MoEF), Government of India under All India Coordinated project on Taxonomy of Algae (AICOPTAX). We acknowledge the Rockefeller Foundation, NY, USA for a training fellowship to B.K. and the Biotechnology Career Development Award to G.T. The Authors thank SPIC Science Foundation, Chennai, India and Director, Centre for Advanced Studies in Botany, University of Madras, India for kindly providing facilities and support.

REFERENCES

Amitai-Zeigerson H, Scolnik PA, Bar-Zvi D (1995). Tomato Asr1 mRNA and protein are transiently expressed following salt stress, osmotic stress and treatment with abscisic acid. Plant Sci. 110: 205-213

Callis J, Carpenter T, Sun CW, Vierstra RD (1995). Structure and evolution of genes encoding polyubiquitin and ubiquitin-like proteins in Arabidopsis thaliana ecotype Columbia. Genetics, 139: 921-39

Chen W, Kuo T (1993) A simple and rapid method for preparation of gram negative bacterial genomic DNA. Nucl. Acids Res. 21: p. 2260

Dekeyser RA, Claes B, De Rycke RMU, Habets ME, Van Montagu MC, Caplan AB (1990) Transient gene expression in intact and organized rice tissues. Plant Cell 2: 591-602

Filichkin SA, Leonard JM, Monteros A, Liu P, Nonogaki H (2004). A Novel Endo-β-Mannanase Gene in Tomato LeMAN5 is Associated with Anther and Pollen Development. Plant Physiol. 134: 1080-1087

Fromm M, Taylor LP, Walbot V (1985). Expression of Genes Transferred into Monocot and Dicot Plant Cells by Electroporation. Proc. Nat. Acad. Sci. USA, 82: 5824-5828.

Gong H, Hu W, Jiao Y, Pua E (2005). Molecular characterization of a Phi-class mustard (*Brassica juncea*) glutathione S-transferase gene in *Arabidopsis thaliana* by 5'-deletion analysis of its promoter. Plant Cell Rep. 24: 439-447.

Hall TA (1999). BioEdit: a user friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids Symp. Ser. 41: 95-98.

Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999). Plant *cis*-acting regulatory DNA elements (PLACE) database Nucl. Acids Res. 27: 297-300.

Hong JK, Hwang BK (2006). Promoter activation of pepper class II basic chitinase gene, CAChi2, and enhanced bacterial disease resistance and osmotic stress tolerance in the CAChi2-overexpressing *Arabidopsis*. Planta 223: 433-448

Horsch RB, Fry JE, Hoffman NL (1985). A simple and general method for transferring genes into plants. Science, 227: 1229-1231.

Iucern ND, Bartholomew DM, Hitz WD, Scolnik PA (1993). Tomato (*Lycopersicon esculentum*) transcript induced by water deficit and ripening. Plant Physiol. 102: 1353-1354

Jefferson RA, Kavanagh TA, Bevan MW (1987). GUS fusions: β-*glucuronidase* as a sensitive and versatile gene fusion marker in higher plants. EMBO J. 6: 3901-3907

Kalifa Y, Gilad A , Konrad Z, Zaccai M, Scolnik PA, Bar-Zvi D (2004). The water- and salt-stress-regulated *Asr1* (abscisic acid stress ripening) gene encodes a zinc-dependent DNA-binding protein. Biochem. J. 381: 373-378.

Khush GS (1997). Origin, dispersal, cultivation and variation of rice. Plant Mol. Biol. 35: 25-34.

Kuriakose B, Arun V, Gnanamanickam SS, Thomas G (2009). Tissue-specific expression in transgenic rice and Arabidopsis thaliana plants of GUS gene driven by the ' regulatory sequences of an anther specific rice gene YY2. Plant Sci. doi:10.1016/j.plantsci. 2009.06.016.

Liu YG, Whittier RF (1995). Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. Genomics, 25: 674-681

Ochman H, Gerber AS, Hartl DL (1988). Genetic applications of an inverse polymerase chain reaction. Genetics, 120: 621-623.

Rogers SO, Bendich AJ (1994). Extraction of DNA from plant, fungal and algal tissues. In: Gelvin SB, Schilperoort RA (eds) Plant Molecular Biology Manual. Boston, MA: Kluwer Academic Publishers, D 1: 1-8.

Rosenthal A, Jones DSC (1990) Genomic walking and sequencing by oligo-cassette mediated polymerase chain reaction. Nucleic Acids Res 18: 3095-3096

Sambrook J, Fritsch EF, Maniatis T (1989). Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Shillito RD, Saul MW, Pasrkowski J, Müller M, Potrykus,I (1985). High efficiency direct gene transfer to plants. Nat. Biotech. 3: 1099-1103.

Siebert PD, Chenchik A, Kellogg DE, Lukyanov KA, Lukyanov SA (1995). An improved PCR method for walking in uncloned genomic DNA. Nucl. Acids Res. 23: 1087-1088.

Terauchi R, Kahl G (2000). Rapid isolation of promoter sequences by TAIL-PCR: the 5'-flanking regions of *Pal* and *Pgi* genes from yams (*Dioscorea*). Mol. Gen. Genet. 263: 554-560.

Vaidyanathan R, Kuruvilla S, Thomas G (1999). Characterization and expression pattern of an abscisic acid and osmotic stress responsive gene from rice. Plant Sci. 140: 21-30.

Wei H, Wang ML, Moore PH, Albert HH (2003). Comparative expression analysis of two sugarcane polyubiquitin promoters and flanking sequences in transgenic plants. J. Plant Physiol. 160: 1241-51

Werr W, Lorz H (1986). Transient gene expression in a Gramineae cell line. A rapid procedure for studying plant promoters. Mol. Gen. Genet. 202: 471-475

Wroblewski T, Tomczak A, Michelmore R (2005). Optimization of Agrobacterium-mediated transient assays of gene expression in lettuce, tomato and Arabidopsis. Plant Biotechnol. J. 3: 259-273

Yang Y, Li R, Qi M (2000). In vivo analysis of plant promoters and transcription factors by agroinfiltration of tobacco leaves. Plant J. 22: 543-551

Zhang XH, Chiang VL (1996). Single-stranded DNA ligation by T4 RNA ligase for PCR cloning of 5'-noncoding fragments and coding sequence of a specific gene. Nucl. Acids Res. 24: 990-991.