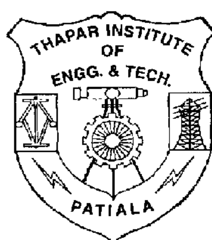
A RAPID MOLECULAR TECHNIQUE FOR CLONING OF 5' FLANKING REGION OF A POTATO GENE

A

THESIS REPORT



Submitted in partial fulfillment of the requirement for the award of the degree of Master of Science in Biotechnology

Under the guidance of:

Dr. N. Das
Assistant Professor

Submitted by:

Ms. Urvashi Jaiswal
Roll No. 3030127

DEPARTMENT OF BIOTECHNOLOGY AND ENVIRONMENTAL SCIENCES
THAPAR INSTITUTE OF ENGINEERING AND TECHNOLOGY
(DEEMED UNIVERSITY)
PATIALA – 147004

2005

CERTIFICATE

This is to certify that the thesis entitled, "A RAPID MOLECULAR TECHNIQUE FOR CLONING OF 5' FLANKING REGION OF A POTATO GENE" submitted by Urvashi Jaiswal in partial fulfillment of the requirement for the award of the degree of Master of Science in Biotechnology, to Thapar Institute of Engineering and Technology (Deemed University), Patiala is a record of student's own work carried out by her under my supervision and guidance. The report has not been submitted for the award of any other degree or certificate in this or any other university or institute.

(Dr. N. Das)	(Dr. N. Das)
Supervisor	Head
Deptt. Of Biotech. & Env. Sciences	Deptt. Of Biotech. & Env. Sciences
Thapar Institute of Engg. & Technology	Thapar Institute of Engg. & Technology
Patiala	Patiala

Dean, Academic Affairs
Thapar Institute of Engg. And Technology
Patiala

CANDIDATE'S DECLARATION

I, hereby declare that the work presented in the thesis entitled, "**A RAPID MOLECULAR TECHNIQUE FOR CLONING OF 5' FLANKING REGION OF A POTATO GENE**" in partial fulfillment of the requirement for the award of the degree of Master of Science in Biotechnology, Department of Biotechnology and Environmental Sciences, Thapar Institute of Engineering and Technology, Patiala, is an authentic record of my own work during the period of six months from January 2005 to June 2005, under the guidance of **Dr. N. Das**, Assistant Professor, Thapar Institute of Engineering and Technology, Patiala. I have not submitted the matter embodied in this thesis for the award of any other degree or diploma.

DATE: URVASHI JAISWAL

PLACE:

ACKNOWLEDGEMENT

This is a golden opportunity for me to convey my sincere regards for all those people who enabled me to accomplish my thesis work successfully.

*It is my privilege to express my sincere gratitude to **Dr.N Das**, Assistant Professor and Head of Department of Biotechnology and Environmental Sciences, for his guidance, valuable advice and constant patience through this work was able to take the shape in which it has been presented. It is his confidence imbibing attitude and splendid discussions and endless endeavors through which I have gained a lot and learned a lot building up my future and personality. And my special thanks to **Dr. Sunil Khanna**, former Head of the Department for his favor and concern.*

A special thanks to all faculty members for their constant encouragement and support throughout the project work.

I am thankful for the help rendered by the Ph.D scholars throughout my project period without which it would have been difficult for me to achieve my goals. My special thanks to Mr.Anshu Bansal, Mrs. Vijay Rana, Ms Sunita Bansal. I also thankful to Mr. Phulchand and Mr. Iqbal our lab staff. I feel lacunae of words to express my most heartfelt and cordial thanks to my friend Mr.Vijendra Mishra who has always stood by my side during all the tough times.

The whole credit of my achievements during the project work goes to my parents and my guide. It was their unshakeable faith in me that has always helped me to proceed further.

DATE: (URVASHI JAISWAL)

PLACE: Patiala

ABSTRACT

By conventional way, the promoter region of any gene can be obtained by screening of a huge number of overlapping recombinant clones of a genomic library with the help of a suitable nucleic acid probe. Usually, bacteriophage-based vectors are used for construction of eukaryotic genomic library. But this process of screening is laborious, time-consuming and cost-intensive.

On the other hand, a number of PCR-based molecular techniques are also used to amplify DNA outside a region of known sequence. The present thesis work mainly focused on a rapid PCR-based molecular technique namely Single Specific Primer PCR (SSP-PCR) for cloning of 5' flanking region of any gene of interest. It is one of the faster molecular techniques that can be employed for cloning of promoter region without resorting to conventional cloning procedures. As a case study we followed a cDNA sequence encoding soluble acid invertase from potato (*Solanum tuberosum* L.) as available in the database. For SSP-PCR, two oligonucleotide primers were designed based on this cDNA sequence where one served as single specific primer. Similarly pUC19 vector-based forward and reverse primers were also designed. Two primer pairs were employed keeping the single specific primer as common along with one vector-specific primer. Partially digested (using *Eco*RI enzyme) overlapping potato DNA fragments ligated into pUC19 vector were used as template. The size of the amplicon was approx. 1.0 kb. This amplicon was further processed & cloned into the *Sma*I site of pUC19 vector. A few recombinant plasmids were obtained which were characterized further through restriction analysis as well as PCR technique using internal gene-specific primer. The genomic inserts correspond to the 5' flanking region of the soluble acid invertase gene as supported by the above results that can be confirmed further by sequencing only. Similarly, *Sau*3AI digested potato DNA fragments were used in the SSP-PCR as mentioned above. Few clones were also obtained that are yet to be characterized. The molecular technique SSP-PCR appears to be quite promising as experienced from the present study.

1. INTRODUCTION

The study of gene structure and function has been greatly facilitated by recombinant DNA technology and it is now fundamental to the advance of modern biological sciences. The products of recombinant DNA technology range from recombinant proteins to engineered organisms. Large amount of commercially useful proteins, enzymes, various metabolites, hormones (human insulin, erythropoietin), anticoagulant (tissue plasminogen), monoclonal antibodies, vaccines, growth factors (somatostatin) and engineered organisms can be produced by employing these techniques. Production of recombinant proteins is a global trend due to certain advantages. It has a host peptide at its N-terminus to stabilize the molecule & prevent it from being degraded by the host cells. These also act as signal peptides, responsible for directing the recombinant protein to its correct position in the cell. It simplifies the problem of purification of the recombinant protein from the culture.

Of course, in practice the production of these is not as easy as it sounds. Besides these, the production of heterologous proteins is a challenge in biotechnology because it is difficult to identify expression systems that are appropriate or capable of expressing proteins of interest. For this special types of cloning vectors (expression vectors) are needed. Expression vector provides the DNA sequences required for transcription, translation & regulation of cloned genes. They allow the production of large amounts of recombinant proteins for research & commercial purposes, but satisfactory yields of these are often difficult to obtain due to the promoter, which is critical component of an expression vector. In fact the promoter controls the very first stage of gene expression (attachment of an RNA polymerase enzyme to the DNA) & determines the rate of which mRNA is synthesized. The amount of these obtained therefore depends to a great extent on the nature of the promoter carried by the expression vector. Thus an expression vector should be carrying a suitable promoter to transcribe the cloned gene at the highest possible rate. Thus, due to the significance of efficient promoter its functional characterization is prerequisite for expression vector.

Besides these, promoters are also regarded as molecular biological tools crucial for the regulation of the expression of genes of interest. A gene is a nucleotide sequence in a DNA molecule that acts as a functional unit for the production of a protein, a structural RNA, or a catalytic RNA molecule. The most important function of

DNA is to carry genes, the information that specifies all the proteins that make up an organism-including information about when, in what types of cells, and in what quantity each protein is to be made.

Regulation of gene expression in eukaryotes is complicated process. In eukaryotes at a given point of time, in a particular cell type only around15 % genes are expressed. Protein-coding genes are usually composed of a string of alternating introns and exons, but only a small percentage of this DNA codes for proteins or structural and catalytic RNAs. RNA polymerase II binds to specific sequence in DNA, known as **promoter** to regulate gene expression. TBP (TATA Box binding proteins) recognizes TATA Box and bind on it to facilitate the entry of RNA polymerase II. These together form basal transcription apparatus or transcription initiation complex (TIC), which initiate transcription, but it carries out transcription at very slow rate. To enhance the rate some other transcription factors bind to corresponding *cis*–acting elements through its **DNA binding domain** whereas through its **transactivation domain** it also interacts with RNA polymerase II. This interaction could be inducing some conformational change, which frequently speed up the transcription.

Although all cells of one organism contain more or less the same genetic information, but gene regulation as one of the major sources of diversity in the phenotypes seen in nature. Genes are turned on and others are turned off at different locations and times during the life cycle of an organism (spatio-temporal nature of gene expression). This turning on/off is very critical. In fact, genes are present everywhere but due to *trans*-acting factors, expressions are different in each cell type. If a particular cell has preexisted inactive transcription factor, then after signal transduction it gets modified by acetylation, methylation, and phosphorylation reactions and become active.

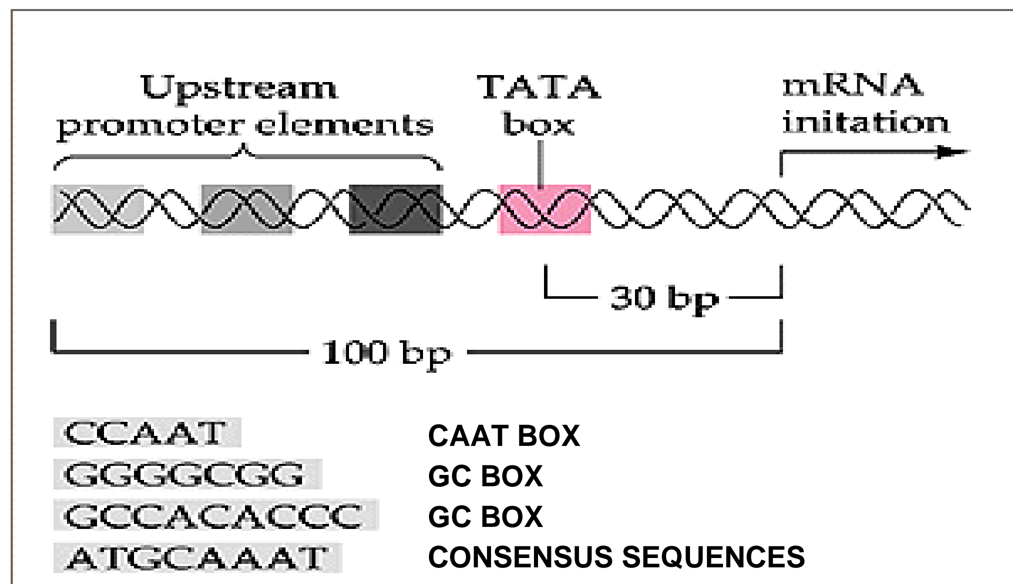
c*is*–acting sequences are short conserved tetrameric/hexameric DNA sequences. It lies upstream of the start point and usually spread out over a region of more than 200 bp such as: light responsive elements (LREs), Abscisic acid–response elements (ABREs). ABRE has ACGT tetra nucleotide core where bZIP (basic leucine zipper type of transcription factor) binds whereas *trans*-acting elements are proteins that drive the expression of gene upon binding to corresponding *cis*–acting elements, but which is not itself part of RNA polymerase such as: HLH (helix turn helix), bZIP (basic leucine zipper). Thus, the study of 5' flanking region of a gene is also weighted on to isolation and characterization of *cis*–acting elements and *trans*–acting factors.

Efficient promoter could be employed to drive the expression of any gene of interest either in constitutive or cell-type specific manner. Whether promoter is strong or weak could be authenticated by the expression of reporter gene such as GUS (β glucuronidase), GFP (green fluorescent protein) etc.

For study of gene regulation and the production of heterologous proteins we need to isolate promoters, but the screening of huge number of overlapping recombinant clones of a genomic library to obtain a promoter sequence when only the cDNA sequence is available can be an arduous task. It is also laborious, time consuming and cost-intensive. Developments in PCR-based approach have circumvented the laborious procedures of producing and screening genomic libraries. These PCR-based approaches are referred to as advance and faster molecular techniques for cloning of 5' flanking region of any gene of interest without need of genomic libraries. These include Inverse Polymerase Chain Reaction (IPCR), Rapid Amplification of Genomic DNA Ends Polymerase Chain Reaction (RAGE-PCR), Targeted Gene Walking Polymerase Chain Reaction (TGW-PCR) and Single Specific Primer Polymerase Chain Reaction (SSP-PCR).

1.1 PRIMARY INFORMATION ABOUT EUKARYOTIC PROMOTERS

Over the last two decades, the concept of a promoter has changed from that of a string of a few functional elements to that of a very complex region of interacting structural & functional elements. The promoter region is the key *cis*-acting regulatory eukaryotic module that responds to the RNA polymerase II & controls the transcription of adjacent coding region(s) into messenger ribonucleic acid (mRNA), which is then directly translated into proteins.



Diagrammatic view of 5' flanking region of eukaryotic genes

Transcription is a controlled process, where multiple DNA regions are involved and the promoter is the main determinant for the initiation of transcription and modulation of levels and timing of gene expression. Promoters in eukaryotic organisms e.g. plants, animals - comprise multiple elements, some of which are found in nearly all promoters. These include:

CAAT box: A consensus sequence lies 80 bp upstream of the start point (+1). It plays an important role in promoter efficiency. This box is replaced in plants by a consensus sequence called the **AGGA box**.

TATA box: A sequence usually located around 25 bp upstream of the start point. The TATA box tends to be surrounded by GC rich sequences. The TATA box binds RNA polymerase II and a series of transcription factors (TFIIX, being X a letter that identifies an individual transcription factor) to form an initiation complex.

GC box: A sequence rich in guanidine (G) and cytidine (C) nucleotides is usually found in multiple copies in the promoter region, normally surrounding the TATA box.

Conserved eukaryotic promoter elements	Consensus sequence
CAAT box	GGCCAATCT
TATA box	TATAA
GC box	GGGCGG

RNA polymerase II is the enzyme that transcribes a gene into RNA. It works in conjunction with other transcription factors that recognize signals embodied in the promoter region.

1.2 TYPES OF PROMOTERS USED TO REGULATE GENE EXPRESSION

Promoters are categorized according to the type or degree of control of gene expression: control in all or virtually all tissues or control depending on the tissue and the developmental stage of the plant. Additionally, promoters may operate in response to external and, in some cases, controllable stimuli. They can be generally divided as:

Constitutive promoters: These promoters direct expression in virtually all tissues and are largely, if not entirely, independent of environmental and developmental factors. As their expression is normally not conditioned by endogenous factors, constitutive promoters are usually active across species and even across kingdoms.

Tissue-specific promoters: These direct the expression of a gene in specific tissue(s) or at certain stages of development. They may be induced by endogenous and exogenous factors, so they may be also classified as inducible promoters.

Inducible promoters: The activity of these promoters is induced by the presence or absence of biotic and abiotic factors. Inducible promoters are a very powerful tool in genetic engineering because the expression of genes linked to them can be regulated to function at certain stages of development of an organism or a particular tissue. These promoters are quite popular nowadays because their performance is not conditioned to endogenous factors but external ones that ideally can be artificially controlled. Within this group, there are promoters modulated by abiotic factors such as light, oxygen levels, heat, cold and wounding etc. Promoters induced by environmental factors such as water or salt stress, anaerobiosis, temperature, illumination and wounding have potential for use in the development of plants resistant to various stress

conditions. These promoters contain regulatory elements that respond to such environmental stimuli.

Since some of these factors are difficult to control outside an experimental setting, promoters that respond to chemical compounds (Chemically regulated promoters), which are not found naturally in the organism of interest, are of particular interest. Along these, promoters that respond to antibiotics, copper, alcohol, steroids, and herbicides, among other compounds, have been adapted and refined to allow the induction of gene activity at will and independently of biotic or abiotic factors.

1.3 HOW DO WE OBTAIN PROMOTER SEQUENCES?

Promoters play a critical role in recombinant DNA technology, as they are the key regulatory eukaryotic modules in expression of gene of interest either in constitutive or cell type specific manner.

A. PROMOTER ISOLATION BY CONVENTIONAL STRATEGY: Molecular cloning of 5' flanking sequences of any gene of interest through conventional strategies requires screening of huge number of overlapping recombinant clones (usually in bacteriophage–based vectors) of a genomic library. A good genomic library is a random collection of various overlapping DNA fragments that represent the whole genome of an organism.

Isolation of a good quality genomic DNA (free from inhibitory materials) is the first step to construct a genomic library. Then the restriction digestion of source genomic DNA is carried out only to a partial extent by tetracutter/hexacutter restriction endonucleases. DNA fragments should be generated in overlapping manner to minimize the loss of a particular gene fragment and clone random DNA fragments of a large size in appropriate vectors such as bacteriophage λ vector (approximately 20 kb DNA fragment can be cloned in bacteriophage λ vector), cosmid, YAC (yeast artificial chromosome). Furthermore, the recombinant clones should overlap with one another, giving an opportunity for **'Chromosomal walking**" studies. Because of the larger size of each cloned DNA, fewer clones are required for a complete or nearly complete library.

During screening, appropriate radiolabelled probes are usually used. The probe may be single-stranded oligonucleotide or corresponding cDNA clone or a segment of gene of interest.

However, construction of a good quality genomic library is a tedious job. It is laborious, time-consuming and cost-intensive. At present, many PCR-based molecular approaches permit the rapid amplification of regions of unknown sequence flanking a known DNA sequence without utilizing conventional cloning procedures.

B. PROMOTER ISOLATION BY CONVENTIONAL POLYMERASE CHAIN REACTION: The conventional PCR involves two oligonucleotide primers, which flank the DNA sequence that is to be amplified. Based on the known gene promoter sequences oligonucleotide primers can be designed. Conventional PCR makes possible to obtain promoter sequences of a corresponding gene only from related organisms.

C. PROMOTER ISOLATION BY OTHER ADVANCED POLYMERASE CHAIN REACTION: A limitation of conventional PCR is the requirement for the two primers, which define the region to be amplified. It cannot be employed to amplify sequences that lie beyond a region for which a primer pair can be designed. A number of other PCR-based techniques can be used for walking into uncloned genomic DNA such as Inverse PCR, Single Specific Primer PCR, Rapid amplification of genomic DNA ends PCR and Targeted gene walking PCR.

Inverse PCR allows the amplification of DNA flanking region of a known sequence (Ochman *et al*, 1988 and Triglia *et al*, 1988). The method is based upon cutting DNA with a restriction endonuclease, and ligating the fragments intramolecularly to form circular molecules. Primers designed to extend outwardly from a known core sequence can then be used to amplify a linear fragment, which comprises sequences from a specific circular molecule. The amplified sequences are those that flank the core sequence in the genome, their lengths depending upon the positions of the restriction sites on each side whereas Single Specific Primer PCR is also a faster molecular technique for amplification of 5’ flanking region of any gene of interest using one vector specific primer and second gene specific primer. The chromosomal DNA is digested with a restriction endonuclease and ligated into linearized plasmid DNA. This used as template for SSP-PCR, by using one vector specific and other gene specific primer, 5’ flanking region of any gene of interest could be amplified.

2. LITERATURE REVIEW

Obtaining the promoter sequence for a gene when only the cDNA sequence is available can be an arduous task (Triglia *et al*, 1988). It requires screening of huge number of overlapping recombinant clones of genomic library by using corresponding cDNA as probe. Developments in PCR in the last several years have circumvented many experimental obstacles by *in vitro* amplification of loci that are otherwise difficult to clone or study *in vivo* (Novak *et al*, 1997).

Thus the polymerase chain reaction (PCR) is a powerful molecular technique allowing the enzymatic amplification of specific regions of DNA without utilizing conventional cloning procedures. But there is a major limitation of PCR is that it enables the amplification only of the region of DNA situated between two convergent primers. However, using the conventional PCR procedure, DNA sequences that lie immediately outside the primers are apparently inaccessible because oligonucleotides that prime DNA synthesis into flanking regions, rather than included regions, allow only a linear increase in the number of copies. The linear increase occurs because, for each primer, there is no priming of DNA synthesis in the reverse direction (Ochman *et al*, 1988).

At present a number of faster molecular techniques could be used for walking in uncloned genomic DNA without resorting to conventional cloning procedures. Amplification of 5' flanking region of any gene of interest can be easier by applying these rapid molecular techniques when only its cDNA sequence is available such as Inverse PCR (IPCR), Single Specific Primer PCR (SSP-PCR), Rapid Amplification of Genomic DNA Ends (RAGE) PCR and Targeted Gene Walking (TGW) PCR.

IPCR leads to the amplification of previously unknown sequence since the primers, which initially face away from each other on the linear template, can be made to face each other as in normal PCR following circularization of the template (Triglia *et al*, 1988). Inverse PCR circumvents the laborious procedures of producing and screening genomic libraries. It has successfully been used to isolate the seed lipoxygenase promoter from pea as well as a wound–inducible promoter from asparagus (Forster *et al*, 1994).

Any method for the *in vitro* amplification of the DNA sequences that flank a known segment of DNA such as IPCR would have many useful applications in genetics. Examples of applications include relatively easy identification of the consensus sequences for insertion of transposable elements and determination of the DNA sequences left behind after partial excision (Ochman *et al*., 1988).

Genes identified as encoding anti-microbial proteins can now be rapidly & precisely introduced into elite germplasm, creating novel pathogen resistant lines (Shah 1997). The polygalacturonase-inhibiting proteins (PGIPs) are key members of a class of anti-fungal proteins that potently inhibit the activity of cell wall-degrading fungal enzymes (De Lorenzo and Cervone, 1997). A polygalacturonase-inhibiting protein (*pgip*) gene from *Malus domestica* cv Granny Smith apple plants was cloned by degenerate oligo-primed polymerase chain reaction (PCR) and Inverse PCR (Melanie *et al*., 1999). An alignment of the pear and bean PGIP sequences was used to design degenerate PCR primers in highly conserved regions. Degenerate PCR allowed the amplification of a 351 bp internal fragment of the *pgip* gene, termed *ipgip*. The DNA sequence of *ipgip* was used to design IPCR primers. A southern blot of apple genomic DNA probed with the *ipgip* fragment was used to identify restriction enzyme sites for IPCR. IPCR enabled cloning of the remainder of the gene, from which a composite *pgip* gene sequence was constructed. A new set of PCR primers were designed to the 5' and 3' ends of the gene, which allowed amplification of the full-length gene from apple genomic DNA. This method has broad application to isolation of homologues of any gene for which some sequence information is known (Melanie *et al*., 1999).

Single Specific Primer PCR is also an example of faster molecular technique. It designed for DNA amplification when only one gene specific primer is available, has been shown to amplify fragments of upto 1.8 kb. By optimizing this technique, it was possible to amplify significantly longer DNA fragments (Novak *et al*., 1997).

SSP-PCR has been adopted to solve a variety of problems, including cloning of DNA contiguous to the site of insertion of the Tn*916* transposon in S*treptococci.* Chromosomal DNA is isolated from a strain harboring a Transposon (Tn*916*) within the locus of interest. The chromosomal DNA is digested with a restriction endonuclease & ligated into linearized plasmid DNA. The resulting population of ligated molecules is comprised of a complex mixture of chromosomal DNA & chromosomal–plasmid DNA hybrids. The plasmid-derived region of the hybrid molecules provides the downstream priming site (the anchor) for PCR amplification. The upstream primer may be specific for

the Transposon or a gene specific primer. The SSP-PCR protocol was successfully used with a Transposon–specific primer to amplify a 0.85 kb fragment adjacent to a Tn*916* insertion in S*treptococcus mutans*. No product was observed with the parental strain lacking a Tn*916* insertion. It also applied for cloning of 5'-regions of genes encoding three different streptococcal proteins and cloning a locus encoding structural and biosynthetic genes for biosynthesis of mutacin II, a novel antimicrobial peptide from mutans streptococci (Novak *et al*, 1997).

Optimization of the concentrations of magnesium, nucleotides, template DNA, primers, annealing temperatures, number of cycles and other interdependent variables are important for yield, length of product and specificity in PCR. Asymmetric ratios of gene-specific and vector-specific primers reduce the background of nonspecific amplification. However, the annealing temperature and time of extension are the two most important factors (Novak *et al*, 1997).

Besides these, several other methods have been developed for the amplification of DNA sequences that flank regions of known sequences. These include TGW-PCR (Targeted Gene Walking PCR) (Parker *et al*, 1991) and RAGE PCR (Cormack *et al*, 1997).

Targeted gene walking PCR can be used for the amplification of unknown DNA sequences adjacent to a short stretch of known sequence by using the combination of a single, targeted sequence specific PCR primer with a second, nonspecific 'walking' primer. This technique can replace conventional cloning and screening methods with a single step PCR protocol to greatly expedite the isolation of sequences either upstream or downstream from a known sequence. A number of potential applications are discussed, including its utility as an alternative to cloning and screening for new genes or cDNAs, as a method for searching for polymorphic sites, restriction endonuclease or regulatory regions and its adaptation to rapidly sequence DNA of lengthy unknown regions that are contiguous to known genes (Parker *et al*, 1991).

Whereas in RAGE PCR amplification of upstream genomic sequences based solely on downstream DNA information from a cDNA clone. In this novel and rapid technique, genomic DNA (gDNA) is first incubated with a restriction enzyme that recognizes a site within the 5' end of a gene, followed by denaturation and polyadenylation of its free 3' ends with terminal transferase. The modified gDNA is then used as template for PCR using a gene-specific primer complementary to a sequence in the 3' end of its cDNA and an anchored deoxyoligothymidine primer. A second round of

PCR is then performed with a second, nested gene-specific primer and the anchor sequence primer. The resulting PCR product is cloned and its sequence determined. Three independent plant genomic clones were isolated using this method that exhibited complete sequence identity to their cDNAs and to the primers used in the amplification (Cormack *et al*, 1997). The mouse growth hormone releasing hormone gene promoter was obtained by using the RAGE technique (Mizobuchi and Frohman, 1993). Upstream region of a salt–induced gene based upon a partial cDNA clone (RSC5-U) GenBank accession no. BG734522 obtained from sunflower (*Helianthus annuus* L.) has been cloned by RAGE technique (Baird *et al*, 2001).

AIM OF THE PRESENT STUDY

The main focus of this thesis work is to adopt a rapid molecular technique for cloning of 5' flanking region of any gene of interest without the need of tedious conventional screening procedure. As a case study, we followed the known cDNA sequence corresponding to soluble acid invertase from a potato cultivar as available in the database. The purpose was to isolate & characterize 5' flanking region of the corresponding gene.

The objectives were framed as follows:

- To adopt and perfect the SSP-PCR molecular technique
- To employ this technique for isolation of 5' flanking region of soluble acid invertase gene
- To clone the amplicons into plasmid vectors for further characterization

3. MATERIALS

3.1 PROCUREMENT OF POTATO GERMPLASM AND OTHER MATERIALS:

- The germplasm of potato cultivar name Kufri Chipsona-1 was procured from Central Potato Research Institute (CPRI), Shimla.
- Various enzymes used were purchased from Bangalore Genei Pvt. Ltd., Bangalore and Amershan Biosciences Ltd., Hongkong.
- The required chemicals were bought from Sisco Research Laboratories Pvt. Ltd., and HiMedia Pvt. Ltd., Mumbai.
- Primers were bought from Bangalore Genei Pvt. Ltd., Bangalore.

3.2 STRAINS AND PLASMIDS:

- *E.coli* DH5α: *sup*E44 ΔlacU169 (Φ80 *lac*ZΔM15) *hsdR17* *rec*A1 *end*A1 *gyr*A96 *thi*-1 *rel*A1
- pUC19 (2686 bp): Genbank accession no: X02514 (Yanisch-Perron *et al*.,1985)

The above bacterial strain was routinely maintained in the laboratory. *E.coli* DH5α strain was maintained on Luria agar medium, whereas, those transformed with pUC19 plasmid were maintained on LA-Ampicillin medium.

3.3 MEDIA USED

Luria Bertani (LB) Medium:

Yeast extract - 0.5 % (w/v)
Tryptone - 1.0 % (w/v)
NaCl - 1.0 % (w/v)
Agar - 1.5 % (w/v)

Volume was made up by single distilled water.

For preparing LA-Ampicillin medium, ampicillin was added to the LA medium at the working concentration of 50 μg/ml, after autoclaving.

3.4 BUFFERS USED

Gel Loading Buffer (5X):

Sucrose - 35 % (w/v)

EDTA - 50.0 mM (pH 8.0)

Bromophenol blue - 0.2 % (w/v)

Volume was made up by double distilled water.

STET BUFFER:

Sucrose - 8.0 % (w/v)

Triton X 100 - 0.5 % (w/v)

EDTA - 50.0 mM (pH 8.0)

Tris HCl - 10.0 mM (pH 8.0)

Volume was made up by double distilled water and autoclaved.

TBE BUFFER (5X):

Tris base - 54 g/l

Boric acid - 28 g/l

EDTA - 3.8 g/l

The pH of the buffer was set at 8.0

Volume was made up by double distilled water and autoclaved.

TE BUFFER (1X):

Tris HCl - 10.0 mM (pH 8.0)

EDTA - 1.0 mM (pH 8.0)

Volume was made up by double distilled water and autoclaved.

TAE BUFFER (5X)

Tris base - 24.2 g/l

Glacial acetic acid - 5.7 ml/l

EDTA (pH 8.0) 0.5M - 10 ml/l

Volume was made up by double distilled water and autoclaved.

STE BUFFER

NaCl - 150 mM

Tris HCl (pH 8.0) - 25 mM

EDTA (pH 8.0) - 2.5 mM

3.5 ENZYMES USED

3.5.1 Restriction Enzymes:

Various hexacutter and tetracutter restriction enzymes such as *Eco*RI, *Bam*HI, *Sau*3AI, *Sma*I, *Hind*III and *Hinc*II were used in this study. Restriction digestion was carried out in buffer supplied by manufacturer. Depending on specific enzyme, reaction was carried out at appropriate temperature and BSA added as required.

3.5.2 Other Enzymes:

Ribonuclease A

Stock solution - 10 mg/ml

Working solution - 10-15 μg/ml

DNase free Ribonuclease A was prepared in a buffer containing 10 mM Tris (pH 8.0) and 15 mM NaCl. To prepare DNase free RNase, the solution was boiled for 10 minutes, followed by the slow cooling, after which it was dispensed into aliquots and then stored at -20°C for subsequent use.

Lysozyme

Stock solution - 10 mg/ml

Working solution - 300-400 μg/ml

Volume was made up by double distilled water.

Freshly prepared lysozyme was used in regular work.

T4 DNA Ligase

Stock conc. - 5 U/ μl

Working conc. – 5 U/50 μl of the reaction volume

This was diluted using dilution buffer as provided by the manufacturer.

Klenow fragment of DNA polymerase I

Stock solution - 5 U/ µl

Working solution - 2.5 U/50 µl of the reaction volume

3.6 OTHERS

X-gal (5-Bromo-4-chloro-3-indolyl-β-D-galactoside)

Stock solution -20 mg/ml

Working solution -20 µg/ml

It was prepared by dissolving the required amount in N, N-dimethyl formamide.

IPTG (Isopropyl thio-β-D-galactoside)

Stock solution - 100 mg/ml

Working solution - 100 µg/ml

It was prepared in fresh and sterile water.

4. METHODOLOGY

4.1 SETTING UP RESTRICTION DIGESTION REACTION:

Restriction endonucleases are the enzymes that recognize short DNA sequences and cleave the double stranded DNA at specific sites within or adjacent to the recognition sequences. The main components of a restriction digestion reaction are-

DNA - 0.5-1 μg

Restriction enzyme - 1-10 units (as supplied by the manufacturer)

Buffer - used at a conc. of 1X (as supplied by the manufacturer)

BSA - used at a conc. of 1X (as supplied by the manufacturer)

Sterile distilled water - To make up the desired volume

Usually restriction digestion was carried out in a reaction volume of 15 μl for 2-3 hrs at temperatures depending upon restriction enzymes used.

4.2 AGAROSE GEL ELECTROPHORESIS:

Agarose gel electrophoresis was performed using standard methods (Sambrook *et al*, 1989). 0.8 % agarose gel was made in 0.5X TBE buffer to which ethidium bromide dye was added (working conc. 0.5-1.0 μg/ml). The DNA samples were loaded after mixing well with the gel loading buffer and electrophoresis was carried out at 2-5 V/cm till the tracking dye covered two-third of the gel length. Finally, the DNA bands were visualized under UV light.

4.3 SETTING UP LIGATION REACTION:

A ligation reaction was set up in order to ligate the insert into vector (pUC19), using the enzyme T4 DNA ligase. It catalyzes the formation of phosphodiester bond between the juxtaposed 5'-phosphate and 3'-OH termini in the duplex DNA. It can join blunt as well as the cohesive end termini. The main components of a ligation reaction are as follows-

Vector ~ 0.3 μg

Insert ~ 0.3 to 0.6 μg

T4 DNA Ligase enzyme – 1-10 units (as supplied by the manufacturer)

T4 DNA Ligase buffer containing PEG 8000 was used in case of blunt end ligation. The reaction volume was made up to 15 μl and the reaction was set up ~25°C for cohesive end ligation & at 21°C for blunt end ligation. It was carried out for 2-3 hours.

4.4 PRIMER DESIGNING:

Primer designing checkpoints: During primer designing certain points should be considered to avoid non-specific amplification. Primer designing is always a critical and most important part of any PCR technique. For this primer length should be 17 to 25 nucleotides. Its GC content should be preferably around 50%. Sequences with long runs of any single nucleotide should be avoided. Primer with significant secondary structure is undesirable. There should be no complementarity between the primer pair.

Primers based on Acid invertase cDNA sequence

The following primers were designed based on the available cDNA sequence corresponding to soluble acid invertase of S. *tuberosum* **(Accession No: X70368)** in the database. A small stretch of the sequence at the 5' end of the above cDNA is given below.

AGTACCATTC CAGTTATGAC CCGGAAAACT CCGCCTCCCA TTACACATTC CTCCCGGATC AACCC
65

GATTC CGGCCACCGG AAGTCCCTTA AAATCATCTC CGGCATTTTC CTCTCCTCTT TCCTTTTGCT
130

TTCTGTAGCC TTCTTTCCGA TCCTCAACAA CCAGTCACCG GACTTGCAGA GTAACTCCCG TTCGC
195

CGCCG CCGTCAAGAG GTGTTTCTCA GGGAGTCTCC GATAAGACTT

AI-R181, the Reverse primer, is complementary to the bases 162-181 of the above cDNA clone, having sequence as follows:

5'- CTC TGC AAG TCC GGT GAC TG - 3'

In this study, AI-R181 primer served as a gene-specific primer. It was designed in order to get amplification of 5' flanking region of the corresponding gene.

Besides this primer, one nested or internal primer is also designed to further characterization of the amplified product.

K20-AI, the nested primer, corresponds to the bases 1-20 of the of the acid invertase cDNA clone (accession number. X70368) having sequence as follows:

5'- AGT ACC ATT CCA GTT ATG AC –3'

During Nested-PCR, the primer combination **AI-R181& K20-AI** was used.

Primers based on pUC19 vector

UP-F361, the Forward primer, corresponds to the bases 361-380 of the cloning vector pUC19 (accession number. X02514), having sequence as follows:

5'- TTT CCC AGT CAC GAC GTT GT-3'

UP-R486, the Reverse primer, is complementary to the bases 467-486 of the cloning vector pUC19 (accession number. X02514), having sequence as follows:

5'-TCA CAC AGG AAA CAG CTA TG-3'

Both the above primers were designed near the MCS (multiple cloning site) sequence of pUC19 vector. It could be mentioned here that the bases 396-452 correspond to the MCS of pUC19.

4.5 SINGLE SPECIFIC PRIMER POLYMERASE CHAIN REACTION:

Single Specific Primer PCR was carried out using genomic insert ligated into vector as template and the following primer pairs: -

- AI-R181 & UP-F361
- AI-R181 & UP-R486

The composition for a typical 50 μl SSP-PCR reaction was as follow:

PCR buffer	- used at a concentration of 1X
Template DNA	- 0.1 to 1.0 μg
Forward primer	- 20 pmoles
Reverse primer	- 20 pmoles
dNTPs	- 125 μM
Taq DNA polymerase	- 3 U/ μl

Finally volume made up to 50 μl with sterile water.

During SSP-PCR, the temperature cycling parameters were as given below:

Pre PCR for Initial denaturation at 94°C, 1 min 30 sec

PCR for 30 cycles

Denaturation	- 94°C, 1 min
Annealing	- 55°C, 2 min
Polymerization	- 72°C, 2 min

POST PCR - Final extension at 72°C, 5 min in order to ensure the completion of DNA synthesis at 3' recessed ends.

4.6 KLENOW TREATMENT:

Klenow fragment is a proteolytic product of *E.coli* DNA polymerase I, which retains polymerization and 3' to 5 ' exonuclease activity. Klenow treatment was carried out as a polishing step of the PCR-amplified products. Because *Taq* DNA polymerase has a tendency to add extra 'A' residue at the 3' ends. Apart from this, some 3' recessed termini also occur in PCR amplified products. Here Klenow treatment was carried out for 30 min at room temperature in presence of dNTPs.

4.7 PURIFICATION OF DESIRED DNA BANDS:

1. ELECTRO –ELUTION TECHNIQUE: In this technique the coveted DNA bands were recovered from the 0.8 % agarose gel prepared in 1X TAE buffer. Electrophoresis was carried out using 1-2 V/cm voltage gradient. After resolution of bands, with the help of a sterile blade a slice of agarose gel containing the desired band was excised. This agarose gel slice was transferred to a prepared dialysis bag and submerged in 0.6-0.7 ml 1X TAE buffer.

Electrophoresis has done at 70 volt, for 1:30 hours. Transfer the eluted DNA sample in clean microfuge tube. Then solvent extraction was done with saturated phenol twice followed by mixture of phenol and chloroform (1:1). Centrifugation was done at 8,000 rpm for 10 min. Then transfer the upper aqueous layer in clean microfuge tube and 1/10th volume of 3 M sodium acetate (CH_3COONa) was added followed by the addition of 2.5 volumes of ethanol. Mixed well and kept at -20°C. The DNA pellet was washed with 70 % ethanol, air-dried and finally dissolved in 13-15 µl of TE buffer and stored at -20 °C for the subsequent use.

2. LOW MELTING AGAROSE TECHNIQUE: The desired DNA bands were recovered from the gels using standard methods (Wieslander *et al*, 1979). 0.8 % low melting agarose gel was prepared in 0.5 X TAE buffer. Electrophoresis was carried out using 1-2 V/cm voltage gradient. After resolution of bands, with the help of a sterile blade a slice of agarose containing the desired band was excised and transferred to a clean microfuge

tube. Then incubated at 65°C for 5 min to ensure melting of agarose gel. Solvent extraction was done using equal volume of saturated phenol. Upper aqueous layer was transferred to fresh microfuge tube. To the previous microfuge tube STE buffer was added for re-extraction followed by centrifugation at 8,000 rpm for 8 min. Upper aqueous layer of later was added to microfuge tube having upper aqueous layer of first microfuge tube. Now total extraction was done with equal volume of phenol twice followed by once with chloroform. Precipitation was done by adding 1/10th volume of 3 M sodium acetate (CH_3COONa) followed by addition of 2 volumes of ethanol or equal volume of isopropanol. The DNA pellet was washed with 70 % ethanol, air-dried and finally dissolved in 15 µl of TE buffer and stored at -20°C for the subsequent use.

4.8 TRANSFORMATION OF *E.coli* DH5α STRAIN:

E.coli was transformed with the ligation mix using the standard $CaCl_2$ method. To prepare the competent cells, a single bacterial colony was inoculated in 25 ml of LB and incubated at 37°C at 120 rpm overnight. The overnight grown culture was reinoculated into 25 ml fresh LB. The culture was grown up to 0.4 OD at 590 nm.

Then cooled in ice water and the cell pellet was recovered by centrifuging at 6000 rpm for 10 min. The pellet was resuspended well in 15 ml of ice-cold 0.1 M $CaCl_2$, again recovered by centrifugation. The pellet was then resuspended in 1 ml of ice-cold 0.1 M $CaCl_2$ for at least two hrs at 0°C for the development of 'competence'. 100 µl of the competent cell suspension was dispended into sterile microfuge tubes at 0°C. 5-6 µl of the ligation mixture was added to the cell suspension, mixed well and then kept at 0°C for 30 minutes. Heat shock was given to all the tubes at 42°C for 2 minutes, followed by the addition of 1 ml LB and further incubation at 37°C for 1 hr. 100 µl of the above transformed cells were plated on LA-ampicillin medium containing X-gal and IPTG. The plates were then incubated at 37°C for 16-18 hrs. The transformants were selected based on α-complementation (blue/white selection).

4.8 ISOLATION OF PLASMID DNA FROM BACTERIAL TRANSFORMANTS BY BOILING PREP METHOD:

Plasmids were isolated by rapid boiling method (Holmes *et al*, 1981). Bacterial transformant colonies were inoculated aseptically in 4.5 ml LB containing ampicillin in

test tubes. The culture was incubated overnight at 37°C for 120 rpm. Cells were harvested from 1.5 ml overnight grown culture in microfuge tubes. The dried pellet was loosened by vortexing, followed by resuspension in 800 μl of STET buffer. 30 μl of lysozyme was added to the bacterial suspension and mixed well. Each microfuge tube containing cell suspension then kept in boiling water bath for 1.5 minutes. After cooling down to room temperature high speed centrifugation (12,000 rpm) was carried out for 15 minutes. After removing the pellet, 2.0 μl of RNase solution was added to the supernatant to remove the contaminating RNA. After incubation at 37°C for 45 minutes equal volume of phenol:chloroform was added, mixed for 5-7 minutes and centrifugation was performed at 10,000 rpm for 10 minutes. To the upper aqueous layer, $1/10^{th}$ volume of 3 M sodium acetate (CH_3COONa) and equal volume of isopropanol was added for precipitation. Solution was incubated at 4°C for 45 minutes. Then it was centrifuged at 10,000 rpm for 10 minutes. The DNA pellet was washed with chilled 70 % ethanol to ensure the removal of excess salts and other impurities. Finally, DNA pellet was air dried at room temperature and dissolved in 20-30 μl of TE buffer.

4.10 ANALYSIS OF RECOMBINANT PLASMIDS:

Various recombinant plasmid DNA samples in this study were digested with several restriction enzymes to check the number of fragments generated and the orientation of the genomic inserts. Moreover, the same plasmids were used as template to carry out PCR using the primer pairs as discussed earlier.

5. RESULTS AND DISCUSSION

Single Specific Primer Polymerase Chain Reaction is a PCR-based rapid molecular technique for amplification of 5' flanking region of any gene of interest without resorting to conventional cloning procedures. As a case study, we focus on soluble acid invertase gene from potato cultivar as its corresponding cDNA clone available in the database.

This thesis work mainly dealt with the isolation and partial characterization of the 5' flanking region of soluble acid invertase gene from the potato cultivar Kufri Chipsona 1. This cultivar is suitable to our agro-climate conditions. The results are given in the following sections:

5.1 RESTRICTION ANALYSIS OF THE TOTAL POTATO DNA:

The total DNA isolated from the potato cultivar, Kufri Chipsona-1 was digested with restriction enzymes namely *Eco*RI, *Bam*HI and *Sau*3AI. *Eco*RI and *Bam*HI are hexacutters whereas *Sau*3AI is a tetracutter. Restriction pattern in each case clearly indicated that the total DNA sample was essentially devoid of inhibitory materials (Fig. 1). More extensive smear was obtained in case of the tetracutter *Sau*3AI as expected.

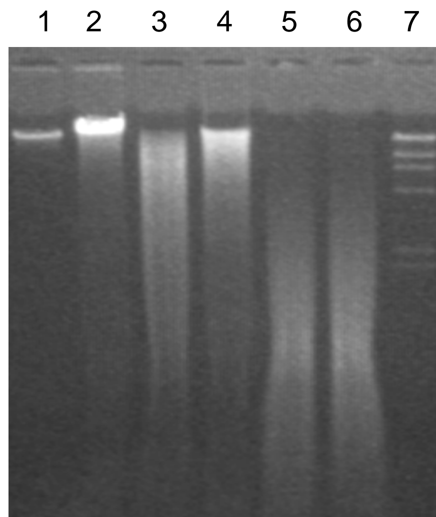


Fig.1. Restriction analysis of Kufri Chipsona1 total DNA

Lane 1, bacteriophage λ DNA; Lane 2, total DNA of Kufri Chipsona1; Lanes 3 & 4, total DNA digested with *Eco*RI & *Bam*HI respectively; Lanes 5 & 6 total DNA digested with *Sau*3AI for 20 min and 40 min respectively; Lane 7, λ DNA digested with *Hind*III

5.2 SINGLE SPECIFIC PRIMER PCR

In order to carry out SSP-PCR the total potato DNA was digested with *Eco*RI to generate small-sized overlapping genomic fragments. The digestion pattern was thoroughly analysed by agarose gel electrophoresis (data not shown). These random DNA fragments were ligated to the *Eco*RI site of pUC19 vector, which served as template in the SSP-PCR. The details of the oligonucleotide primers as used in the study are given in section 4.4 of Methodology. Out of these primers, AI-R181 served as 'single specific primer' with respect to soluble acid invertase gene whereas the primers UP-F361 & UP-R486 were pUC19 specific. It may be mentioned here that the primer K20-AI corresponding to the 5' end of the soluble acid invertase cDNA was only used for characterization purpose. In a SSP-PCR, the pair-wise primer combination (keeping AI-R181 common in both the cases) were as follows: -

Primer pair I: AI-R181& UP-F361
Primer pair II: AI-R181& UP-R486

These combinations were employed since the potato genomic DNA fragments could be ligated to vector in either orientation. During PCR, temperature cycling parameters were maintained as given in section 4.5 of Methodology.

The PCR amplified products were analyzed by 0.8% agarose gel electophoresis (Fig. 2). In the case of primer pair II only approx. 1 kb amplicon was obtained. The result suggests that insert ligated in reverse orientation at corresponding site of pUC19 vector as well as some portion of 5' flanking region of acid invertase is also amplified. The data appeared to be quite promising that led to further cloning & characterization as discussed in the following sections.

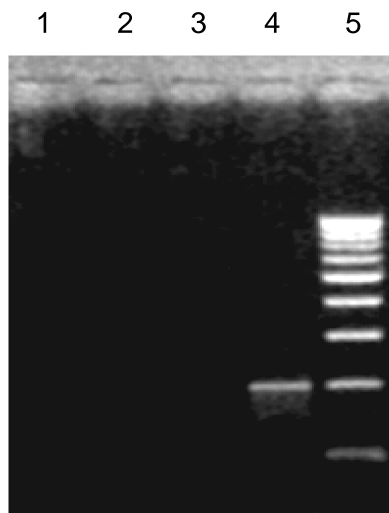


Fig.2. Analysis of SSP- PCR amplified products using the template, *Eco*RI digested genomic DNA fragments ligated into *Eco*RI site of pUC19

Lane 1, Control PCR with UP-F361& AI-R181 primers without template; Lane 2, Control PCR with UP-R486 & AI-R181 primers without template; Lane 3, the above template using UP-F361 & AI-R181 primers (primer pair I); Lane 4, the above template using UP-R486 & AI-R181 primers (primer pair II); Lane 5, 500 bp ladder

5.3 MOLECULAR CLONING OF THE SSP-PCR AMPLIFIED PRODUCT

KLENOW TREATMENT: As a polishing step, the above amplified product was treated with Klenow enzyme.

As the amplicon was obtained using pUC19 specific primer UP-R486, it was quite lightly that one end of the above amplicon contained entire MCS of pUC19 vector. Therefore prior to cloning the polished amplicon was further digested with *Sma*I to remove the MCS region as mentioned above.

LIGATION: The resulting DNA fragment was purified by electoelution (as described in section 4.7) and inserted at *Sma*I site of pUC19 vector for cloning. A specific blunt end ligation buffer containing PEG 8000 (poly ethylene glycol) was used to improve the efficiency of the ligation.

TRANSFORMATION AND SELECTION OF TRANSFORMANTS: A number of white transformant colonies were obtained on LA-ampicillin plates containing X-gal and IPTG corresponding to each set of ligation mixture as mentioned in the previous section. The

plasmid DNA isolated separately from each purified white colony in mini scale by boiling method (as mentioned in section 4.9). Only a few plasmid samples were found to contain inserts. The recombinant plasmids were designated as pUN-IP01, pUN-IP02 & pUN-IP03, which were pursued for further analyses.

5.4 FURTHER CHARACTERIZATION OF RECOMBINANT PLASMIDS

RESTRICTION ANALYSIS: Restriction analysis plays a significant role in partial characterization of recombinant clones. It confirms the intactness of the restriction sites in genomic clones & efficient ligation of the insert.

All the recombinant plasmids namely pUN-IP01, pUN-IP02, and pUN-IP03 were digested with *Eco*RI as shown in Fig. 3. pUN-IP01 recombinant clone gave rise to single band whereas ~1.0 kb and ~0.8 kb DNA fragments were released from pUN-IP02 & pUN-IP03 respectively.

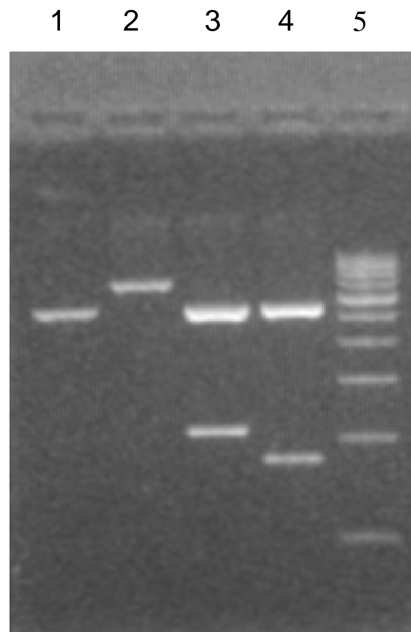


Fig.3. Restriction analysis of recombinant plasmids with *Eco*RI

Lane 1, pUC19 linearized with *Eco*RI; Lanes 2, 3 & 4 pUN-IP01, pUN-IP02 & pUN-IP03 clones linearized with *Eco*RI respectively; Lane 5, 500 bp ladder,

Moreover, the restriction pattern of the above recombinant plasmids with *Bam*HI is shown in fig. 4. All the recombinant plasmids gave rise to single band. It is apparent that there is no internal *Bam*HI site in the cloned genomic inserts. Similarly, Fig. 5 shows the restriction pattern of the above recombinant plasmids with *Hind*III. A DNA fragment of approx 0.7 kb was released in case of only pUN-IP02.

Here, each restriction digestion pattern was closely examined. The data suggest that the genomic inserts are placed in opposite orientation in cases of pUN-IP01 & pUN-IP03 that could be easily confirmed by further PCR & sequencing. The restriction pattern of pUN-IP02 appeared to be different. As soluble acid invertase gene encoded by multigene family, the insert in case of pUN-IP02 might correspond to the other member of the same gene family that can be easily substantiate by further sequencing.

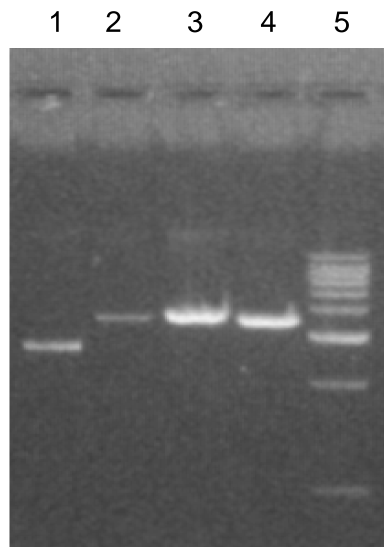


Fig.4. Restriction analysis of recombinant plasmids with *Bam*HI

Lane 1, pUC19 linearized with *Bam*HI; Lanes 2, 3 & 4 pUN-IP01, pUN-IP02 & pUN-IP03 clones linearized with *Bam*HI respectively; Lane 5, 1kb ladder

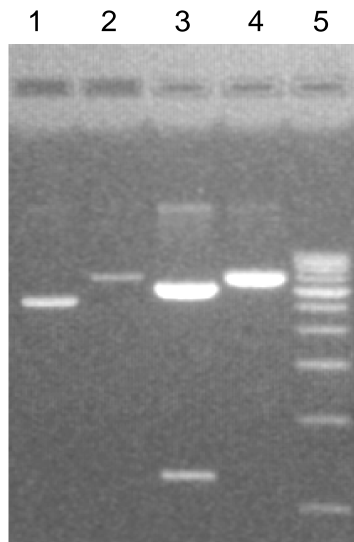


Fig.5. Restriction analysis of recombinant plasmids with *Hind*III

Lane 1, pUC19 linearized with *Hind*III; Lanes 2, 3 & 4 pUN-IP01, pUN-IP02 & pUN-IP03 clones linearized with *Hind*III respectively; Lane 5, 500 bp ladder

PCR-BASED ANALYSIS: For further characterization, each recombinant plasmid was used as template separately to carry out PCR using the same combination of primer pairs (i.e. primer pair I & primer pair II) as used in original PCR. This was done to check whether the genomic inserts in the recombinant plasmids correspond to the DNA fragment that was amplified in original PCR. The result is shown in Fig. 6.

The primer pair I (UP-F361& AI-R181) worked well for pUN-IP01 whereas primer pair II (UP-R486 & AI-R181) worked well for pUN-IP02 & pUN-IP03. The result weighted that the genomic inserts are placed in opposite orientation in cases of pUN-IP01 & pUN-IP03. It also supported the view that genomic inserts in all the three clones (pUN-IP01, pUN-IP02 & pUN-IP03) may correspond to the same soluble acid invertase gene family.

Further, out of these three recombinant clones, pUN-IP02 was taken for further characterization. For this purpose one internal primer namely K20-AI (corresponding to the 1-20 base of soluble acid invertase cDNA) was employed.

First of all, pUN-IP02 recombinant plasmid was digested with restriction enzyme *Eco*RI. It generated two distinct bands, of size 2.7 kb and ~0.8 kb (as shown in Fig. 2) which were purified by low melting agarose method (as mentioned in the section 4.7 of Methodology). Both of these were used as template separately for Nested-PCR using

the primer pair combination K20-AI & AI-R181 (where AI-R181 is single specific primer). PCR result showed consistency with primer designing thus indicating that the insert corresponds to the soluble acid invertase gene. Similarly, approx 1.0 kb SSP-PCR amplified product (as shown in Lane 4, Fig. 1) was also used as template using the same primer combination K20-AI & AI-R181.

The primer pair worked well for smaller fragment of recombinant clone pUN-IP02 (~0.8 kb) as well as for SSP-PCR amplified product (as shown in fig. 7). 181 bp amplicon was obtained in both the cases. The result ensures that the genomic insert in the recombinant clone namely pUN-IP02 corresponds to acid invertase gene.

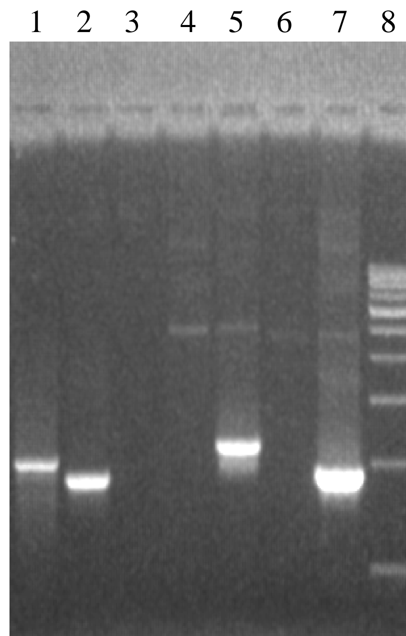


Fig.6. Characterization of cloned fragments by PCR

Lane 1, Approx. 1.0 kb SSP-PCR amplified product using UP-R486 & AI-R181 primers (The same is shown in Lane 4, Fig. 1); Lane 2, pUN-IP01 clone as template using UP-F361& AI-R181 primers; Lane 3, pUN-IP01 clone as template using UP-R 486 & AI-R181 primers; Lane 4, pUN-IP02 clone as template using UP-F361& AI-R181 primers; Lane 5, pUN-IP02 clone as template using UP-R486 & AI-R181 primers; Lane 6, pUN-IP03 clone as template using UP-F361& AI-R181 primers; Lane 7, pUN-IP03 clone as template using UP-R486 & AI-R181 primers ; Lane 8, 500 bp ladder

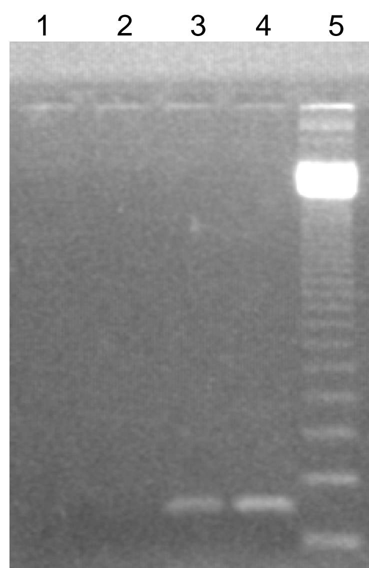


Fig.7. Partial characterization of cloned fragment by Nested PCR

Lane 1, pUC19 vector as template using K20-AI & AI-R181 primers; Lane 2, larger band of pUN-IP02 as template using K20-AI & AI-R181 primers; Lane 3, smaller band of pUN-IP02 as template using K20-AI & AI-R181 primers; Lane 4, approx. 1.0 kb SSP-PCR amplified product (The same is shown in Lane 4, Fig. 1) using K20-AI & AI-R181 primers; Lane 5, 123 bp ladder

Similarly, SSP-PCR was also carried out using the template; *Sau*3AI digested genomic DNA ligated into the *Bam*HI site of pUC19 vector (Fig. 8). All the PCR conditions were kept same (as mentioned in the section 4.5 of Methodology). An amplicon of approx.1 kb was obtained in the case of primer pair II. This suggests that genomic insert ligated in reverse orientation.

The above-amplified product was treated with Klenow enzyme as a polishing step. Further, *Hinc*II digestion was carried out to remove extra vector specific MCS region. The resulting DNA fragments were ligated into the *Sma*I site of pUC19 vector to be transformed in *E.coli* DH5α.

A number of white transformants were selected on LA-ampicillin plates containing X-gal and IPTG corresponding to each set of ligation mixture as mentioned in the previous section. The plasmid DNA was isolated separately from each purified white colony in mini scale by boiling method. Only a few plasmid samples were found to contain inserts. The recombinant plasmids were designated as pUN-IP04 & pUN-IP05, which were pursued for further analyses.

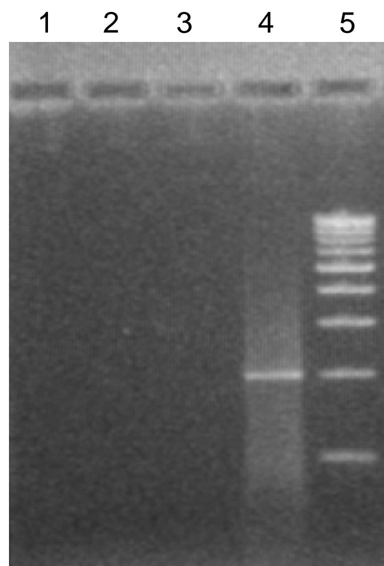


Fig.8. Analysis of SSP- PCR amplified product using the template, *Sau*3AI digested genomic DNA fragments ligated into the *Bam*HI site of pUC19

Lane 1, Control PCR with UP-F361& AI-R181 primers without template; Lane 2, Control PCR with UP-R486 & AI-R181 primers without template; Lane 3, the above template using UP-F361 & AI-R181 primers; Lane 4, the above template using UP-R486 & AI-R181 primers; Lane 5, 500 bp ladder

The restriction banding pattern of the recombinant clones pUN-IP04 & pUN-IP05 with *Eco*RI, *Bam*HI & *Hind*III respectively (as shown in Fig. 9) appeared to be different. On digestion of recombinant clones pUN-IP04 & pUN-IP05 with *Eco*RI, DNA fragments of sizes ~8 kb and ~3 kb respectively, were obtained. Similarily, in *Hind*III digestion approx 0.8 kb and ~0.6 kb DNA fragments were released in the case of pUN-IP04 & pUN-IP05 respectively.

The apparent variation in restriction analysis suggest that in these recombinant clones namely, pUN-IP04 & pUN-IP05, the genomic inserts might correspond to the other members of the same soluble acid invertase gene family. These clones need to be further characterized. **Further study of these clones is under process.**

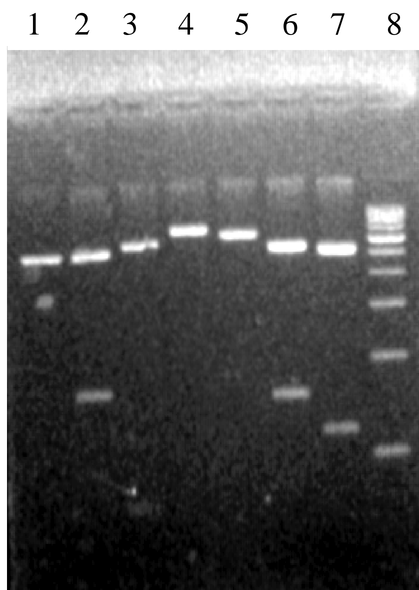


Fig.9. Restriction analysis of recombinant plasmids with *Eco*RI, *Bam*HI and *Hind*III

Lane 1, pUC19 linearized with *Eco*RI; Lanes 2 & 3, pUN-IP04 & pUN-IP05 clones digested with *Eco*RI respectively; Lanes 4 & 5, pUN-IP04 & pUN-IP05 clones digested with *Bam*HI respectively; Lanes 6 & 7 pUN-IP04 & pUN-IP05 clones digested with *Hind*III respectively; Lane 8, 500 bp ladder

In this study, SSP-PCR appeared to be a simple, efficient & rapid molecular technique for cloning unknown DNA fragments by walking from a known region. This method eliminates the need to construct & screen genomic libraries for obtaining the promoter sequence of any gene of interest. Here, this technique helps successfully in cloning the 5’ flanking region of the soluble acid invertase gene from our own potato cultivar. It may be mentioned here that the genomic clone corresponding to the available cDNA sequence as mentioned in this study is not yet available in the database. However, the generated recombinant clones have only a part of 5’ flanking region of gene. Further sequencing will help to delineate the promoter & *cis*–acting element positions in 5’ flanking region of acid invertase gene. This type of promoter region would help to study the molecular mechanism involved in the expression of cold inducible genes.

SUMMARY

The various experimental steps that were adopted during this thesis work could be summarized in following points: -

- Primer designing: It is always a critical and most important part of any PCR-based technique. AI-R181 & K20-AI are primers (20-mer each) that correspond to soluble acid invertase cDNA (accession no. X70368). AI-R181 is a gene specific primer, designed for the amplification of 5' flanking region of the corresponding gene whereas K20-AI served as nested primer for characterization of amplicon. UP-F361 & UP-R486 are pUC19 vector (accession no. X02514) specific primers (20-mer each).

- Restriction digestion: In order to carry out SSP-PCR, the total potato DNA (from the potato cultivar Kufri Chipsona-1) was digested partially with *Eco*RI restriction enzyme to generate overlapping genomic DNA fragments. At the same time, pUC19 vector was linearized with *Eco*RI.

- Ligation into pUC19 vector: The above random potato DNA fragments were ligated into the *Eco*RI site of pUC19 vector, which served as template.

- SSP-PCR: It was carried out by using the above template along with primer pair I (AI-R181& UP-F361) & primer pair II (AI-R181 & UP-R486) accordingly. In case of primer pair II approx. 1.0 kb amplicon was obtained. The result was quite promising for subsequent cloning purpose.

- Molecular cloning: Prior to cloning into the plasmid vector, the above amplicon was treated with Klenow enzyme as a polishing step followed by digestion with the restriction enzyme *Sma*I to remove extra vector specific MCS sequences as associated with the amplicon during SSP-PCR. The resulting DNA fragments were ligated into *Sma*I site of pUC19 vector and then transformed into *E.coli* DH5α. Selection of positive transformant was based on α-complementation (blue/white colony selection). Only a few plasmids were

found to contain inserts. These putative recombinant plasmids were designated as pUN-IP01, pUN-IP02 & pUN-IP03.

- Partial characterization of recombinant clones: Further characterization of these recombinant clones was done through restriction analysis with various restriction enzymes such as *Eco*RI, *Bam*HI & *Hind*III. The result of restriction analysis suggests that the genomic inserts were placed in opposite orientation in cases of pUN-IP01 & pUN-IP03 that need to be further confirmed by sequencing. However, the restriction pattern of pUN-IP02 appeared to be different which suggests that the insert in this recombinant clone might correspond to different member of same soluble acid invertase gene family. However, acid invertase cDNA specific primers (K20-AI &AI-R181) worked well to amplify 181 bp DNA using the insert of pUN-IP02 as template. This was adopted as a step towards partial characterization.

- Similarly, all the above experimental steps were also followed for *Sau*3AI digested overlapping genomic fragments ligated into *Bam*HI site of pUC19 vector used as template for SSP-PCR under the same PCR conditions. Here also, in case of primer pair II approx.1 kb amplicon was obtained. Prior to cloning, this amplicon was polished with Klenow enzyme followed by digested with *Hinc*II to remove extra vector specific MCS sequence. The resulting DNA fragments were ligated into *Sma*I site of pUC19 vector to be transformed into *E.coli* DH5α. Transformants are selection on basis the of α-complementation (blue/white colony selection). Only a few plasmids were found to contain inserts. These recombinant plasmids were designated as pUN-IP04 & pUN-IP05 that require further characterization.

REFERENCES

1. Cormack, R. S. and Somssich, I. E. (1997) Rapid amplification of genomic ends (RAGE) as a simple method to clone flanking genomic DNA. Gene, 194(2): 276-6.
2. Holmes, D.S. and Quigley, M. (1981) A rapid boiling method for preparation of bacterial plasmids. Analytical Biochemistry, 114: 193.
3. Liu, X. and Baird, W. V. (2001) Rapid amplification of genomic DNA ends by *Nla*III partial digestion and polynucleotide tailing. Plant Molecular Biology Reporter, 19: 261-267.
4. Mandel, M. and Higa, A. (1970) Calcium-dependent bacteriophage DNA infection. Journal of Molecular Biology, **53**: 159-162.
5. Melanie, S. A., Ian, A. D. and David, K. B. (1999) Isolation by PCR-based methods of a plant antifungal polygalacturonase-inhibiting protein gene. Molecular Biology and Genetics, Vol. **2** No 3.
6. Novak, J. and Novak, L. (1997) Amplification of Flanking Regions: New Applications and Performance Optimization of Single Specific Primer PCR and T-Vector Cloning. *Promega Notes Magazine* Number **61**, p.26.
7. Ochman, H., Gerber, A.S. and Hartl, D.L. (1988) Genetic applications of an inverse polymerase chain reaction, *Genetics* **120**, 621-625.
8. Parker, J.D., Rabinovitch, P.S. and Burmer, G.C. (1991) Targeted gene walking polymerase chain reaction. *Nucl. Acids Res.* **19**, 3055-3060.
9. Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular cloning-A laboratory manual, second edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
10. Triglia, T., Peterson, M.G. and Kemp, D.J. (1988) A procedure for *in vitro* amplification of DNA segments that lie outside the boundaries of known sequences. *Nucl.Acids Res.***16**, 8186.
11. Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene, **33**: 103-119.
12. Zrenner, R., Schuler, K. and Sonnewald, U. (1996) Soluble acid invertase determines the hexose–to-sucrose ratio in cold–stored potato tubers. Planta **198** (2), 246-252.