VARIANTS OF THE GENE ENCODING THE BETA SUBUNIT OF PYROPHOSPHATE DEPENDENT PHOSPHOFRUCTOKINASE (PFP) AND THEIR TRANSCRIPTIONAL EXPRESSION IN SUGARCANE

SANUSHKA REDDY

Thesis presented in fulfilment of the requirements for the degree of
Master of Science at the University of Stellenbosch

Supervisor: Dr Barbara Huckett
Co-supervisor: Prof. Frikkie Botha

December 2001

DECLARATION

I the undersigned, hereby declare that the work carried out in this thesis is my own original work and that I have not previously in its entirety or in part submitted it at any university for a degree.

ABSTRACT

Sugarcane, a complex polyploid, may theoretically contain up to 12 alleles for a particular gene at a single locus. The number of alleles and the extent of their variation is of particular importance due to the potential for the exploitation of genetic variation through breeding. Also, allelic variation has implications for the manipulation of gene function via genetic engineering. Pyrophosphate dependent phosphofructokinase (PFP) is considered a key regulatory enzyme involved in sucrose biosynthesis, which may provide a target for genetic manipulation to increase sucrose yields in sugarcane. The enzyme is composed of a regulatory (alpha) and catalytic (beta) protein subunit encoded by the PFP-α and PFP-β genes respectively. The PFP-β gene, which has been shown to be a single locus gene in other plant species, was used in this study as a model for allelic variation in sugarcane. Two main areas of investigation involved genomic and expression analyses to further characterise the gene. Polymerase Chain Reaction (PCR) using specific primers previously designed from conserved regions of the PFP-β gene from different plant sources, were used to amplify across exons 10 to 12 of the sugarcane PFP-β gene. Two PCR products, designated PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ respectively, were obtained from commercial cultivars N19 and N21. Numerous clones of the fragments were obtained and sequenced. International database searches confirmed that both amplicons were identifiable as PFP-β. Comparative sequence analysis indicated that the PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ fragments were poorly homologous to each other, with higher regions of homology residing in the putative exon regions (77-78%) compared to the intron regions (34-56%). Although minor sequence variation was detected within the amplicon populations, it was evident that two major variants of the PFP-β gene are present in sugarcane. Southern hybridisation analysis revealed a simple banding pattern for PFP-β. Also, there are DNA polymorphisms for the genomic regions corresponding to the PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ fragments. Previous evidence indicates that both variants are also present in the ancestral sugarcane germplasm and maintain the same level of heterozygosity. The presence of both gene forms in the ancestral and commercial germplasm prompts speculation that the two variants may not segregate. This theory, together with the simple Southern hybridisation pattern

obtained for PFP-β, leads to the hypothesis that the two gene forms are at separate loci in the sugarcane genome, which may be closely linked on the same chromosome. The expression of the variants was investigated during different stages of sucrose accumulation in the sugarcane culm using a Reverse Transcription (RT)-PCR approach. A single, identical transcription product was isolated from these and other selected tissues of the plant. In addition, the same transcript was obtained from the ancestral species representatives of modern sugarcane, *Saccharum officinarum* and *Saccharum spontaneum*. Sequence comparison of the transcribed product and the derived exon regions of the two variants implies that the PFP-β gene represented by the PFP-B1/B8$_{1250bp}$ variant is being expressed in sugarcane while the gene form characterised by the PFP-B1/B8$_{1100bp}$ amplicon is silent. Northern hybridisation analysis indicates that PFP-β is differentially expressed at different stages of sucrose accumulation. PFP-β expression is higher in the immature culm tissue of sugarcane and low in the mature culm, which suggests that PFP-β is highly regulated during maturation. It is hypothesised that the PFP-β gene underwent duplication and that one gene form was subject to accumulative mutations evolving into a pseudogene. On the basis of present results, it can be suggested that future genetic manipulation of PFP-β should involve the gene variant characterised by the PFP-B1/B8$_{1250bp}$ fragment.

OPSOMMING

Suikerriet, 'n komplekse poliploïede organisme, kan teoreties tot 12 allele vir 'n spesifieke geen by 'n enkele lokus bevat. Die aantal allele en hul mate van variasie is veral van belang, weens die potensiaal vir die benutting van hierdie genetiese variasie deur teëling. Alleelvariasie het ook implikasies vir die manipulering van geenfunksie via genetiese ingineurswese. Pirofosfaat-afhanklike fosfofruktokinase (PFP) is 'n belangrike regulatoriese ensiem vir sukrose biosintese en kan geteiken word vir genetiese manipulasie met die oog op verhoogde sukroseproduksie in suikerriet. Die ensiem bestaan uit 'n regulatoriese (alfa) en katalitiese (beta) proteïen subeenheid wat deur die PFP-α en PFP-β gene respektiewelik gekodeer word. Die PFP-β geen, wat in ander plantspesies as 'n enkellokusgeen aangedui is, is in hierdie studie gebruik as 'n model vir alleelvariasie in suikerriet. Die twee hoofroetes van ondersoek wat gevolg is, behels genoom- en uitdrukkingsanalises om die geen verder te karakteriseer. Eksons 10 tot 12 van die suikerriet PFP-β geen is geamplifiseer met behulp van die Polimerase Ketting Reaksie (PKR), bemiddel deur die gebruik van spesifieke voorvoerders wat voorheen ontwerp was vanaf gekonserveerde areas van die PFP-β geen uit verskillende plantbronne. Twee PKR produkte, genoem PFP-B1/B8$_{1250bp}$ en PFP-B1/B8$_{1100bp}$ respektiewelik, is deur die kommersiële kultivars N19 en N21 opgelewer. Verskeie fragmentklone is gekonstrueer en hul DNA basisvolgorde is bepaal. Soektogte in internasionale databasisse het bevestig dat beide amplikons PFP-β was. Vergelykende DNA basisvolgorde analise het aangedui dat PFP-B1/B8$_{1250bp}$ en PFP-B1/B8$_{1100bp}$ fragmente swak homologie toon, terwyl 'n hoër mate van homologie in die oënskynlike ekson areas (77-78%), vergeleke met die intronareas (34-56%) gevind is. Alhoewel klein basisvolgordeverskille opgemerk is binne die amplikonpopulasies, was dit duidelik dat twee hoofvariante van die PFP-β geen in suikerriet teenwoordig is. Southern hibridisasie analisie het 'n eenvoudige bandpatroon vir PFP-β onthul. Daar is ook DNA polimorfismes vir die genoomstreke wat met die PFP-B1/B8$_{1250bp}$ en PFP-B1/B8$_{1100bp}$ fragmente ooreenstem. Vorige bewyse het aangetoon dat beide variante ook in die voorvaderlike suikerriet kiemplasma voorkom en dat dieselfde vlak van heterosigositeit gehandhaaf word. Die voorkoms van beide geenvorme in die voorvaderlike asook kommersiële kiemplasmas stel voor dat hierdie twee

variante miskien nie segregeer nie. Hierdie teorie, tesame met die eenvoudige Southern hibridisasiepatroon vir PFP-β, lei tot die hipotese dat die twee geen vorme by verskillende lokusse in die suikerrietgenoom voorkom, en dat hierdie lokusse nou gekoppel mag wees op dieselfde chromosoom. Die uitdrukking van die variante is ondersoek gedurende verskillende stadia van sukrose akkumulasie in die suikerrietstingel deur van Trutranskripsie (TT)-PKR gebruik te maak. 'n Enkele, identiese transkripsie produk is hieruit en uit ander geselekteerde plantweefsels geïsoleer. Dieselfde transkrip is ook verkry vanaf die voorouerlike spesieverteenwoordigers van hedendaagse suikerriet, *Saccharum officinarum* en *Saccharum spontaneum*. 'n DNA basisvolgordevergelyking tussen die getranskribeerde produk en die ekson-areas van die twee variante impliseer dat die PFP-β geen verteenwoordig deur die PFP-B1/B8$_{1250bp}$ variant uitgedruk word in suikerriet, terwyl die geen verteenwoordig deur die PFP-B1/B8$_{1100bp}$ amplikon stil is. Northern hibridisasie analise toon aan dat PFP-β verskillend uitgedruk word gedurende verskillende stadia van sukrose akkumulasie. PFP-β uitdrukking is hoër in die onvolwasse stamweefsel van suikerriet en laag in die volwasse stam, wat voorstel dat PFP-β hoogs gereguleer word gedurende maturasie. Daar word gehipotetiseer dat die PFP-β geen duplikasie ondergaan het en dat een geenvorm onderworpe was aan akkumulerende mutasies wat deur evolusie tot 'n pseudogeen gelei het. Dit word voorgestel, gebaseer op die huidige resultate, dat toekomstige genetiese manipulasie van die PFP-β geen, die geenvariant gekarakteriseer deur die PFP- B1/B8$_{1250bp}$ fragment moet behels.

ACKNOWLEDGEMENTS

- Firstly, I give thanks to God who has shaped my life and showered me with blessings.

- My sincere appreciation is extended to my Mum and Dad for making so many sacrifices and giving me opportunities that they themselves did not have. Thank you to my wonderful fiancé, Tyrrell Naidoo, for his patience, encouragement, enthusiasm and love that has motivated me during these past two years. I am also grateful to my brother, Sharendra Reddy, and sister, Nerushka Reddy, for their never-failing support.

- A special thank you to my supervisor, Dr Barbara Huckett, for all her encouragement, patience and assistance during the course of this study. Thank you Barbara, your approach to molecular biology has led me to a greater appreciation of the subject. Thanks also to my co-supervisor, Prof. Frikkie Botha, for his guidance and valuable input.

- I am deeply indebted to my colleagues at SASEX who have been instrumental in developing my skills in molecular biology. Thank you Debbie Carson, Stuart Rutherford, Derek Watt, Sandy Snyman, Julie Barnes, Aletta Bester, Gwethlyn Meyer, Lucy Thokoane, Bernadette Goodman and Mike Butterfield for your valuable input in my experiments and refreshing ideas. Thanks also to Busi Monosi, Natalie Lacko, Natalie Williams, Natalie Cheetam, Timothy Pepper, Diksha Lalbahadhur, and Zama Mshlali for your friendship and support.

- In addition, I thank Julie Barnes for kindly accepting to proof read this thesis.

- I am grateful to Sarita and Hennie Groenewald as well as to Anne Whittaker for their spontaneous assistance with aspects of my work.

- I also acknowledge the gift of the potato PFP-β cDNA clone from DT Dennis, which is greatly appreciated.

- Many thanks to the SA Sugar Association for the study bursary and the provision of accommodation at Singles' Residence.

- Finally, I would like to thank Prof. Glen Campbell for advising me to take Biology as a major at university seven years ago!

CONTENTS

LIST OF ABBREVIATIONS

A_{260}	absorbance at 260 nanometers
A_{280}	absorbance at 280 nanometers
ABA	abscisic acid
Acetyl-CoA	acetyl-coenzyme A
ADH	alcohol dehydrogenase
ADP	adenosine 5'-diphosphate
ATA	aurintricarboxylic acid
ATP	adenosine 5'-triphosphate
BLAST	basic local alignment search tool
BSA	bovine serum albumin
cDNA	complementary deoxyribonucleic acid
CODEHOP	consensus-degenerate hybrid oligonucleotide primers
CTAB	cetyltrimethylammonium bromide
dATP	deoxyadenosine 5'-triphosphate
dCTP	deoxycytidine 5'-triphosphate
DEPC	diethyl pyrocarbonate
dGTP	deoxyguanosine 5'-triphosphate
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxy nucleoside 5'-triphosphate
DTT	1,4-dithiothrietol
dTTP	deoxythymidine 5'-triphosphate
EDTA	ethylenediaminetetraacetic acid
EST	expressed sequence tag
FBPase	fructose-1,6-bisphosphatase (d-fructose-1,6-bisphosphate 1-phosphohydrolase), EC 3.1.3.11
Fru-1,6-P_2	D-fructose-1,6-bisphosphate
Fru-2,6-P_2	D-fructose-2,6-bisphosphate
Fru-6-P	D-fructose-6-phosphate
GUS	β-glucoronidase
hr	hour

IPTG	isopropyl-β-D-thiogalactopyranoside
min	minutes
mRNA	messenger ribonucleic acid
PAS	p-aminosalicyclic acid
PCR	polymerase chain reaction
PEG	polyethylene glycol
PFK	6-phosphofructokinase (ATP: D-fructose-6-phosphate 1-phosphotransferase), EC 2.7.1.11
PFP	pyrophosphate dependent phosphofructokinase (pyrophosphate: D-fructose-6-phosphate 1-phosphotransferase), EC 2.7.1.90
pfu	plaque forming units
Pi	inorganic phosphate
POP-6	performance optimised polymer-6
pp.	pages
PPi	inorganic pyrophosphate
PR	pathogenesis related
PVP	polyvinylpyrrolidone
RAPD	random amplified polymorphic deoxyribonucleic acid
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
RNase	ribonulease
rpm	rotations per minute
RT-PCR	reverse transcription polymerase chain reaction
sec	seconds
SDS	sodium dodecyl sulphate
TLC	thin layer chromatography
Tris	2-amino-2-(hydroxymethyl)-1,3-propanediol
UDP	uridine 5'-diphosphate
UDPGPPase	uridine 5'-diphosphate-glucose pyrophosphorylase (UTP α-D-glucose-1-phosphate uridylyl transferase), EC 2.7.7.9
UTP	uridine 5'-triphosphate
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Sucrose is a valuable source of revenue for South Africa, being the country's third most important agricultural export product (Anonymous, 1996). Approximately 2.3 million tons of sucrose is produced per season, half of which is exported, yielding a net income of approximately 1.28 billion rands (Anonymous, 2000). To remain internationally competitive, new sugarcane varieties combining high yield and high sucrose content with pest and disease resistance are being continually developed.

Crop improvement relies on the exploitation of genetic variation between parents. This variation may be due to genes at different loci or at one locus. Genes at one locus that differ from each other at one or more mutational sites are alleles (Kahl, 1995). Due to the mutation, one allele may confer a more favourable trait in relation to the other. For example, alleles associated with good breadmaking quality were identified in the wild relative of wheat (Saponaro et al., 1995); and in the wild relative of barley, alleles linked to high β-amylase activity in the grain were identified (Erkillä et al., 1998). The latter allele is potentially useful to increase levels of grain β-amylase activity in malting barley (Erkillä et al., 1998). Thus, the study of allelic variation is important for crop improvement and may be identified at the genetic level by determining the degree of homology between individual gene copies in crop plants. In sugarcane, allelic variation is not well characterised despite the large genetic base still available in hybrid, commercial sugarcane varieties (Roach and Daniels, 1987).

One alternative to conventional plant breeding is the use of genetic manipulation. By changing the level of enzymes involved in carbon allocation by genetic manipulation, the relative yields of the various storage products can be altered (Blakeley and Dennis, 1993). In this way, sucrose levels have been increased in transgenic plants (Worrell et al., 1991; Muller-Rober et al., 1992). One key regulatory enzyme involved in sucrose biosynthesis is pyrophosphate dependent phosphofructokinase (PFP), which catalyses the reversible

conversion of fructose-6-phosphate to fructose-1,6-bisphosphate (EC 2.7.1.90) (Dancer and ap Rees, 1989). PFP is absent from higher animals but has been characterised in plant cells and in some protozoans, for example *Amoeba* (Reeves *et al.*, 1974) and *Giardia lamblia* (Rozario *et al.*, 1995). In most plants, the enzyme is a heterotetramer consisting of two alpha and two beta protein subunits. Kinetic and chemical data have suggested that the alpha is the regulatory subunit and beta the catalytic subunit (Yan and Tao, 1984; Botha and Botha, 1991) These subunits are encoded by the alpha (PFP-α) and beta (PFP-β) genes respectively. The kinetic properties of PFP from several plant species have been studied. However, the physiological role of the enzyme remains speculative. In sugarcane, PFP activity has been shown to be negatively correlated with sucrose content (Whittaker and Botha, 1999). Sucrose content increases in the older internodes of the sugarcane stalk (Botha *et al.*, 1996), where PFP activity was also shown to be controlled by the expression of the β-subunit of the enzyme (Whittaker and Botha, 1999). Due to its relationship to sucrose, PFP has been recognised as a target for genetic manipulation to increase sucrose yields in sugarcane.

Commercial sugarcane varieties are interspecific hybrids originating from several species, particularly the ancestral species *Saccharum officinarum* and *Saccharum spontaneum* (Roach and Daniels, 1987). Due to its ancestry, modern sugarcane varieties are complex polyploids (Al-Janabi, 1998), a genetic characteristic that contributes to an increase in DNA content and genetic variation. During evolution, genes duplicated by polyploidy may be subjected to different fates. The gene may retain its original or similar function, attain a different function or be differentially regulated (for example in the multigene family), or one or more copy of the gene may become silenced (Wendel, 2000). Since sugarcane has chromosome numbers between 100–130 (Lu *et al.*, 1994) and suspected ancestral base chromosome numbers of 8 (*S. spontaneum*) (Bremer, 1961) or 10 (*S. officinarum*) (Sreenivasan *et al.*, 1987), at one locus there could be up to 12 alleles for a particular gene. The possibility that two or more alleles exist could have important implications in genetic manipulation, where attempts to down-regulate the gene may be negated by the presence of another allele.

The aim of this study was to investigate allelic variation in the complex genetic environment of sugarcane. The PFP-β gene was selected as a model for allelic variation since it has been shown to be at a single locus in other species, for example, castor bean (Blakeley *et al.*, 1992). In addition, PFP has been shown to be an important enzyme in sucrose biosynthesis, and can therefore be used in genetic manipulation. Preliminary work has suggested the presence of two variants of the PFP-β gene in both ancestral and commercial sugarcane genotypes (Harvey *et al.*, 1999).

The first specific objective of this study was to perform genomic analysis on part of the PFP-β genic region in order to confirm the presence of variants of the gene, by assessing sequence differences between the variants, and to gain insight as to whether they are alleles. Details of this genomic investigation, which was performed on commercial hybrid sugarcane varieties, are given in Chapter 3. The second main objective of this study was to identify which of the variants are expressed in relation to sucrose accumulation in sugarcane and to determine patterns of expression in both commercial and ancestral genotypes of sugarcane. Thus, expression of the PFP-β gene was investigated in a range of tissues, including the developing stalk of a commercial hybrid variety and in selected tissues from the ancestral varieties. Chapter 4 deals with the latter expression studies.

CHAPTER 2

LITERATURE REVIEW

2.1 Plant Genomes

2.1.1 The organisation of plant genomes

The term genome describes the total complement of DNA in a particular organelle (Paterson, 1996). In plants, there exists the mitochondrial genome, the plastid genome and the nuclear genome (Walbot and Cullis, 1985). The nuclear genome is organised into chromosomes composed of segments of genomic DNA which are either genic or intergenic sequences. Within these sequences are repetitive sequences, regulatory sequences and genes (Flavell and Moore, 1996).

2.1.2 Plant DNA content

There is a great deal of variation in the size of the nuclear genomes of higher plants. Bennett and Smith (1991) estimate this variation to be over 1500-fold. Angiosperms with small nuclear genomes, e.g. *Arabidopsis*, have a 2C nuclear DNA content of 0.30pg (145Mbp/ 1C) (where the C value is the size of the unreplicated haploid nuclear genome), while some polyploid genomes may have 11.92pg (5751 Mbp/ 1C) of DNA (Arumaganathan and Earle, 1991). Plant species with small genomes exist as fully functioning flowering plants; thus it is deduced that plant genomes are relatively large in comparison to the amount of DNA required to specify known functions. For small genomes, e.g. cotton, it is estimated that 5-10% of the genome encodes specific function while it is thought that this amount is even less in plants with large genomes, e.g. rye, where the value is 5% (Flavell, 1980).

2.1.3 Repetitive DNA

DNA-reassociation kinetic studies have revealed that the extra DNA content in plant genomes is due to the accumulation of repetitive sequences found within the intergenic region (Nagl, 1976). Flavell (1980) estimates that the proportions of repetitive DNA may be

as little as 10% in plants with small genomes (for example *Arabidopsis*) and may be as high as 95% in very large genomes. Reassociation kinetic studies have also indicated the interspersion patterns of repetitive sequences in relation to low or single copy sequences in plant genomes (Dean and Schmidt, 1995). In plants with large genomes, shorter stretches of low-copy sequences are present (<2kb) while in plants with small genomes, low-copy sequences are present in much longer stretches (for example 120kb in *Arabidopsis*).

Repeated sequences have been characterised as tandem repeats, which include variable-number tandem repeats, minisatellite DNA, microsatellite DNA and telomeres; and dispersed repeated sequences, which include transposable and retrotransposable elements (Hulbert *et al.*, 1990). Repetitive sequences undergo rapid evolutionary changes compared to coding or single copy sequences (Flavell, 1986; Hulbert *et al.*, 1990). These regions become involved in unequal crossing over, transposition and gene conversion that may lead to changes in sequence and copy number. Sequence scrambling may occur which results in the insertion of short pieces of DNA into new sites in the genome. DNA segments of one tandem array may become part of other tandem arrays, and parts of one transposable element may become part of another (Lucas *et al.*, 1992; Moore *et al.*, 1993). Therefore plant species with a high proportion of repetitive DNA are predisposed to undergoing changes in genome organisation (Flavell, 1986).

2.1.4 Genic sequences

Genic sequences, which are unique and usually single copy (Paterson, 1996), contain the protein coding genes which are units of a few kilobases (Walbot and Cullis, 1985). Genes consist of coding regions (exons) and several intervening, non-coding regions (introns) and regulatory DNA sequences such as promoters upstream and termination regions downstream of the gene. The exon regions encode a polypeptide while the intron regions are spliced out during transcription (Futuyama, 1986). The general organisation of plant genomes is diagrammatically represented in Figure 2.1.
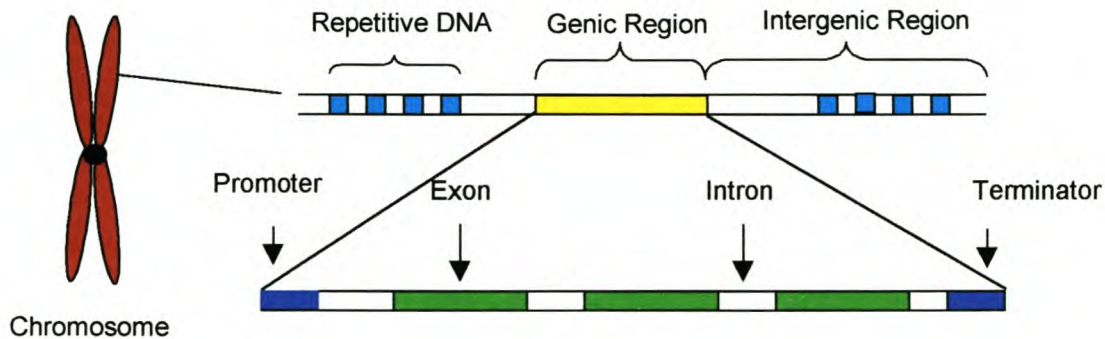
Figure 2.1 Diagram representing the organisation of plant genomes. At the sub-chromosome level there are segments of DNA which are either genic or intergenic sequences. The intergenic sequences contain repeated sequences that contribute to the large size of some plant genomes. At the genic level there are genes consisting of coding (exon) and non-coding (intron) regions, promoter and terminator elements. (adapted from Futuyama, 1986).

2.1.5 Pseudogenes

In addition to functional genes, there are pseudogenes in plant genomes (Brown, 1989). These genes have sufficient homology and the same exon-intron arrangements to indicate that they have arisen by the duplication of functional genes (Futuyama, 1986). However, they contain frame-shift mutations that prevent them from encoding a functional peptide. The pseudogenes diverge rapidly in sequence, so much so that they may no longer be recognised in relation to the functional gene (Smith, 1989). This relates to the "selfish" genes described by Doolite and Sapienza (1980). Selfish genes, as the name implies, do not encode any instruction which can be recognised, but remain in the genome since they retain the ability to replicate.

The most detailed account of the formation of a pseudogene for a duplicated gene is for PgiC2 in *Clarkia mildrediae* (Gottlieb and Ford, 1997). Of the 23 exons, 18 were sequenced and nine exons showed indels (insertions or deletions) causing frameshifts and a stop codon. Some deletions resulted in the loss of exon-intron splice junctions, and

insertions were also present. Introns were also degenerate, containing a total of 21 insertions, one of which (857bp with a stem-loop structure and direct terminal repeats) was indicative of a transposable element.

Pseudogenes have also been characterised in the lectin gene family of *Ricinus communis* (castor bean) (Treagar and Roberts, 1992). The presence of two and three frameshift mutations in two isolated clones respectively, and the lack of transcripts homologous to the putative lectin gene in the third clone, are indicative of pseudogenes. Other authors have suggested the presence of pseudogenes based on the observation of an absence of cDNA transcripts despite the appearance of genomic sequences for a particular gene (observed by Southern hybridisation). This has been described in the α-glucosidase gene from barley (Tibbot and Skadsen, 1996) and the gene for the catalytic subunit of serine/threonine protein phosphatases of *Nicotiana tabacum* (Suh *et al.*, 1998).

Numerous pseudogenes have been characterised in the uridylate rich small nuclear RNA (UsnRNA) genes in plants. Kiss and Solymosy (1993) described the presence of four tandem repeat elements in the 3' flanking region of the LeU3 gene flanked by five direct repeats. Kiss *et al.* (1989), in screening a tomato genomic library for LeU1 genes, encountered a pseudogene which was truncated at its 3' end and lacked the sequences necessary for transcription at the 5' flanking region. In potato, three UsnRNA pseudogenes with a defective promoter or coding region were also isolated (Vaux *et al.*, 1992).

2.1.6 Multigene families

Genes may be spaced out randomly along a DNA molecule or grouped into distinct clusters. In certain instances, the clusters are made up of genes that contain related units of biological information, i.e. the multigene family. It is estimated that there are 20 000 – 30 000 different gene families in higher plant genomes (Flavell and Moore, 1996). Members of multigene families share high sequence homology with each other and may be scattered on one or more chromosomes. It is hypothesised that the members of the gene family arose from a common ancestral gene by a series of duplications, and possibly relocation, followed by events that may have altered the sequences of the genes resulting

in specialisation of the genes (Raven *et al.*, 1986; Brown, 1989). Multigene families may be simple multigene families (all the genes are exactly the same), complex multigene families (the members are all different but expressed at the same time) or complex differentially expressed multigene families (the members are different and expressed at different times) (Brown, 1989).

Multi-gene families can encode genes with related functions, but in cases where the function is the same, differential regulation ensures that distinct genes are activated in response to different environmental and other stimuli. Chernys and Zeevaart (2000) characterised a 9-cis-epoxycarotenoid dioxygenase gene family in avocado, involved in the regulation of abscisic acid (ABA) biosynthesis. They found that two of the genes were strongly induced as the fruit ripened while the other gene was constitutively expressed in the fruit during fruit ripening as well as in leaves. The three genes were differentially expressed under conditions of water stress. Individual members of the *Arabidopsis thaliana* ribulose-1,5-bisphosphate carboxylase/ oxygenase small subunit gene family were differentially regulated by light of different qualities, which suggested that individual members of the family may be subjected to a variety of differing regulatory mechanisms (Dedonar *et al.*, 1993). The maize invertase gene family was found to group into two classes with contrasting sugar responses. The two classes showed differential organ and developmental expression (Xu *et al.*, 1996). The alternative oxidase of the soybean inner mitochondrial membrane is encoded by a multigene family with three known members, the genes of which were also differentially expressed in a tissue-dependent manner (Finnegan *et al.*, 1997).

There are varying degrees of homology between members of multigene families, which may indicate the likely role of the gene in the plant. For example, in the three genes isolated by Chernys and Zeevaart (2000), Pa NCED3 is 67% identical to Pa NCED1 while Pa NCED2 is 30% identical to Pa NCED1 at the deduced amino acid level. One of the genes, Pa NCED2, was found to be unlikely to be involved in ABA biosynthesis. In the 14-3-3 protein gene family in *Arabidopsis thaliana*, the 10 members of the family maintain their

distinctiveness at the N- and C- terminal ends, which suggests that they may encode for different functions (Wu *et al.*, 1997).

2.1.7 Alleles

Alleles are different forms of the same gene occupying the same locus on a particular chromosome (Kahl, 1995). Alleles arise by abrupt and usually stably inherited mutation (Fincham, 1983). Mutations may develop due to insertions, deletions, substitutions, transpositions, translocations, duplications and inversions at one or few nucleotides of the gene sequence (Raven *et al.*, 1986). Thus, variations within plant genomes may evolve over generations if the mutation is not detrimental or lethal.

The use of allelic variation in crop improvement

From the earliest food growers to modern plant breeders, the variation within wild plant species has been exploited in order to improve crop yields, and increase resistance to disease and adverse conditions (Simmonds, 1979; Nilan, 1981). In plants such as barley and peas, as well as maize and wheat, loci controlling phenotypes with much variability have been revealed. Under modern breeding, most crop plants have become genetically uniform. However, genetic variation is particularly important in an inbreeding species for the production of new varieties (van der Have, 1979). Therefore, allelic variation within species is potentially useful for the improvement of crops. For example, good bread making quality was found to be associated with three alleles at the Gli-A1^m locus of *Triticum monococcum,* a wild relative of modern wheat (Saponaro *et al.*, 1995). This locus can be introduced into modern wheat cultivars by crossing. The β-amy1 allele from the wild ancestor of cultivated barley was linked to high β-amylase activity in the grain. Use of this allele in the modern barley breeding program is potentially useful to increase levels of grain β-amylase activity in malting barley (Erkillä *et al.*, 1998). Law *et al.* (1998) detected allelic variation in the genes of group 1 chromosome for ear emergence time in Chinese Spring wheat. The authors suggest that within spring barleys there are subgroups of alleles with different degrees of vernalisation sensitivity controlling ear emergence time, and giving rise to some of the genetic variation, which they observed. Powell *et al.* (1985) recognised that allelic variation at the Gp *ert* locus in spring barley makes a large contribution to the

variation for thousand grain weight, main stem height and single plant height. Allelic variation at these loci is invaluable to the improvement of barley crops. Alleles that express a graminicide-tolerant acetyl-CoA carboxylase gene for tolerance to the sethoxydim herbicide have been described in maize (Gengenbach *et al.*, 1999). Mutants expressing high levels of graminicide-tolerant acetyl-CoA carboxylase have been identified. Thus these alleles may be used to improve the tolerance of maize to the sethoxydim herbicide.

A wide genetic base is available for commercial sugarcane in the form of wild related species. In Taiwan, a related sugarcane species was crossed with commercial sugarcane and it was shown that the hybrid produced high yields and retained resistance to disease (Lo and Shong, 1968). However, attempts to increase the genetic diversity of breeding lines have been of little consequence since there is still sufficient genetic variation in existing commercial sugarcane breeding pools for crop improvement (Roach and Daniels, 1987). This genetic variation, which has been based on phenotypic differences between different sugarcane varieties, has been exploited by plant breeding to produce new sugarcane varieties which are resistant to diseases (downy mildew, Fiji, leaf blight, leaf scald, mosaic virus, ratoon stunting, rust, smut and yellow spot) (Lew, 1987) and pests (rats and borers) (Lew, 1987), stress (drainage, drought, fertility, herbicide, salinity and temperature) (Lew, 1987), as well as varieties which have improved growth and high sucrose yields (Lew, 1987).

2.1.8 Polyploidy

Polyploids are a type of heteroploid containing three or more genomes (Da Silva and Sobral, 1995). Autopolyploidy involves the doubling of a single genome, while allopolyploidisation involves the merger of two, often highly differentiated genomes into a common nucleus (Wendel, 2000). Thus polyploids have multiple copies of chromosomes containing the same genetic information, i.e. each gene is present in multiple copies. In this way, polyploidy contributes to a large part of the genetic variation in plant DNA content (Da Silva and Sobral, 1995). Approximately 47-52% of all angiosperm species and 70% of grass species are polyploid (Grant, 1981; Paterson, 1996). Most of the economically important crops are polyploid and contain two or more distantly related genomes that have

co-evolved. Examples of these crops are cotton, wheat, soybean, tobacco and rape (Paterson, 1993).

Chromosome pairing between homologous chromosomes during meiosis can lead to many chromosome combinations when genes are present in multiple copies. In addition to homologous pairing, homoeologous pairing between genetically equivalent chromosomes from distinct species may occur in plant species derived from the hybridization of different species (Da Silva and Sobral, 1995).

Soltis *et al*. (1992) and Al-Janabi *et al*. (1994) suggest that the evolution and speciation of polyploids may be difficult to ascertain since the interspecific boundaries of some polyploid plant species are unclear due to the possibility of hybridisation and chromosome doubling. Soltis and Soltis (1993) describe the evolution of polyploids as an ongoing and dynamic process in plants. According to Wendel (2000), some plant species are not obvious polyploids, since over time a polyploid genome could evolve rapidly so that the number of genes expressed could be reduced to a level that is representative of the diploid ancestor (Grant, 1981). After polyploidisation, genomes may continue to evolve and sequences may be subjected to elimination or modification (Wendel, 2000).

The evolutionary fate of duplicated sequences varies widely among taxa and among the various components of the genome (Wendel, 2000). Duplicate genes may have long-term maintenance of similar or identical functions. Genes derived by polyploidy may persist since mutations in one gene copy may lead to negative interactions with the products of other essential genes (Wendel, 2000).

According to Wendel (2000), persistence of duplicate genes does not necessarily imply the retention of identical protein function or expression rather, functional diversification may simply be difficult to recognise. Duplication leads to relaxation of selection on one gene copy, allowing divergence between the duplicated genes and the acquisition of new function. These changes in function may confer quite significant adaptations. Such functional divergence is apparent in members of plant multigene families.

The other possible outcome of gene duplication is gene silencing, which is the loss or inactivation of the gene. Mutations in coding and regulatory regions may lead to differences in protein expression or function (diversification) as well as pseudogene formation (permanent silencing) (Wendel, 2000). Transposable elements may be one of the causative agents of gene silencing. For example, non-functional nitrate reductase genes in tobacco were shown to be caused by insertion of a retrotransposon (Grandbastien, 1992). Marshall *et al.* (1994) propose that over evolutionary time, genes that have been silenced as a consequence of point mutations in coding and regulatory sequences may be resurrected, that is, their functions restored by back mutations or gene conversion with a functional homologue.

Interactions between duplicated sequences and protein products may occur. In addition to changes operating at the level of gene expression, the genes themselves might interact via recombination, gene conversion or other mechanisms. In polyploids it is often observed that expression levels are positively correlated with gene copy number (Roose and Gottlieb, 1980), whereas in some cases 'dosage compensation' is observed in that overall expression levels are maintained despite copy number increases (Birchler and Newton, 1981; Guo *et al.*, 1996).

Thus there is genetic variability within polyploids which may be partially responsible for their evolutionary success (Soltis *et al.*, 1992). However, their associated genomic complexity often confounds genetic studies.

2.1.9 Isozymes

Isozymes are multiple forms of an enzyme which catalyse the same reaction (Kahl, 1995). During evolution, changes in the coding regions of the genome may occur that either result in the loss of the protein product, or allow the function of the protein product to be maintained (Tanksley and Pichersky, 1988). If selection favours the maintenance of the protein product, isozymes may result. In some cases, the isozymes may have very similar kinetic and physical properties, for example phosphoglyceromutase, recognised by ion exchange chromatography, in the developing endosperm *of Ricinus communis* (Botha and

Dennis, 1986). However, isozymes differ in their physico-chemical properties and regulation (Kahl, 1995).

In maize endosperm, the two isozymes of sucrose synthase were found to be critical: one plays the dominant role in providing the substrates for cellulose biosynthesis while the other isozyme is required for generating precursors for starch biosynthesis (Chourey et al., 1998). Isozymes may be expressed in a tissue-specific manner. Biles and Martyn (1993) showed that peroxidase, polyphenoloxidase and shikimate dehydrogenase displayed tissue-specific isozymes in watermelon seedlings. In horseradish, the genes for isozymes of peroxidase were found to be subject to organ-specific expression and were functional despite the 10-44% difference in nucleotide homology between the genes (Kawaoka et al., 1992).

Maize Alcohol Dehydrogenase (ADH) is an isozyme that is encoded by two genes, which display a high degree of homology to each other at the amino acid and nucleotide level (82% and 87% respectively) (Dennis et al., 1984, 1985). Most plants have duplicate ADH genes which suggests that both genes are required for ADH activity, however in *Arabidopsis*, one active gene is sufficient to maintain ADH activity.

Isozymes may be encoded by gene variants at one locus (alleles) or by genes found at different loci. The phosphoenolpyruvate carboxylase isozyme involved in C4 photosynthesis in maize is allelic (Hudspeth and Grula, 1989). Comparison of the alleles showed 26 nucleotide differences, three of which produced amino acid substitutions, 22 were silent substitutions in the coding region, and one occurred in the 3' untranslated region.

2.1.10 Conserved gene sequences
Most organisms share very significant similarities in DNA since genes which encode instructions essential to the survival of an organism are likely to remain similar over evolutionary history (Paterson, 1996). Mutations or changes in such basic instructions may

be fatal to the organism. This conservation of genes across species affords the opportunity to use sequence information from one organism and apply it to another.

2.2 Sugarcane

2.2.1 The origin of sugarcane

Sugarcane, a domesticated crop, has been cultivated due to its high sucrose content. Present day varieties are interspecific hybrids of the genus *Saccharum*. This genus, characterised by a high degree of polyploidy, belongs to the family Gramineae, tribe Andropogoneae (Stevenson, 1965). *Saccharum,* together with the closely related interbreeding genera *Erianthus* (section *Ripidium*), *Miscanthus* (section *Diantra*), *Sclerostachya* and *Narenga* make up the *Saccharum* complex (Daniels and Roach, 1987). This complex has been implicated in the origin of sugarcane; however, the actual contribution of these genera to sugarcane is uncertain due to their high and variable ploidy levels (D' Hont *et al.*, 1996).

The history of sugarcane, as reviewed by Roach and Daniels (1987), elucidates the genetic constitution of the commercial crop. Until the end of the 19th century, the principally cultivated canes were *Saccharum officinarum*, together with *S. barberi* and *S. sinense* (Bremer, 1961). *S. sinense* and *S. barberi* are two groups thought to be derived from the natural hybridisations between *S. spontaneum* and *S. officinarum* (Bremer, 1961). *S. officinarum* is characterised by its large size and good sucrose quality in comparison with other *Saccharum* species. The outbreak of disease prompted hybridisation involving *S. officinarum* and the wild *S. spontaneum* (Stevenson, 1965). Although its sugar content was low, *S. spontaneum* was a favourable species due to its vegetative vigor and a high degree of resistance to abiotic stress, disease and parasites (Sreenivasan *et al.*, 1987). The progenies from this hybridisation were repeatedly backcrossed with *S. officinarum* (noble cane), a process referred to as nobilisation, to reduce the effect of the wild parent (Stevenson, 1965). According to Bremer (1961), nobilisation is characterised by asymmetric chromosome transmission. Lu *et al.* (1994) have identified that in hybridization between *S. officinarum* (2n=80) and *S. spontaneum* clones (2n=40-128), the F1 hybrid

obtains 2n=80 chromosomes from the noble female parent and n=20-64 from the wild male parent. Backcrosses with *S. officinarum* result in 2n+n transmission of chromosomes but from the second backcross, chromosomal transmission becomes normal. A consequence of this is that modern cultivated varieties have chromosome numbers between 100 and 130. Five to ten percent of this chromosome number comes from *S. spontaneum*. However, it has been suggested that the major part of the diversity between cultivars is due to the *S. spontaneum* fraction of the genome (Lu *et al.*, 1994).

Both *S. spontaneum* and *S. officinarum* are thought to have autopolyploid origin (D' Hont *et al.*, 1996). From an evolutionary perspective, autopolyploidy is synonymous with polysomic inheritance (Da Silva and Sobral, 1995). Soltis and Soltis (1993) regard polysomy as a potential source of genetic attributes such as enzyme multiplicity, increased heterozygosity, and increased allelic diversity.

2.2.2 The genetic complexity of sugarcane

Sugarcane cultivars have ploidy levels that may be difficult to ascertain since the number of chromosomes per cell of the same plant may sometimes vary (referred to as chromosomal mosaicism) (Burner and Legendre, 1994). In addition, sugarcane genomes have extensive chromosome duplication (Grivet *et al.*, 1996) and a large quantity of DNA per haploid complement (Dean and Schmidt, 1995). The implication of the polyploid nature of sugarcane is that there are multiple copies of genes, although these genes are not necessarily identical and a variable number of copies of the gene could be expressed. This genetic complexity makes sugarcane a difficult system in which to perform genetic studies and complicates attempts to improve the crop systematically.

Genetic engineering, which is an alternative approach to conventional plant breeding to improve crops, may be difficult in sugarcane. Attempts to increase expression of a particular gene may be confounded by the presence of more than one copy of the gene and unexpected gene silencing may occur. In addition, attempts to eliminate gene expression may be negated by the expression of another, non-identical copy of the same

gene. All the gene copies would have to be identified and characterised before gene expression could be eliminated.

In commercial sugarcane cultivars, there could be up to 12 alleles for a particular gene at one locus since sugarcane has chromosome numbers between 100–130 (Lu *et al.*, 1994) and ancestral base chromosome numbers of 8 (*S. spontaneum*) or 10 (*S. officinarum*) (Bremer, 1961; Sreenivasan *et al.*, 1987). Allelic variation is best established by targeting a single or low copy gene. The PFP-β gene, which is related to sucrose metabolism and has been shown to be a single copy gene in castor bean (Blakeley *et al.*, 1992), may be at a single locus in sugarcane. Therefore, PFP-β is a good candidate gene for the study of allelic variation in sugarcane.

2.3 Pyrophosphate dependent phosphofructokinase (PFP)

PFP (EC 2.7.1.90) is an enzyme, with no known isozymes, that catalyses the reversible conversion of fructose-6-phosphate (Fru-6-P) to fructose-1,6-bisphosphate (Fru-1,6-P_2) using inorganic phosphate (PPi) as the phosphoryl donor. Phosphofructokinase (PFK) also catalyses this reaction, however ATP is the phosphoryl donor and the reaction is irreversible (Blakeley *et al.*, 1992). These reactions are represented in figure 2.2.
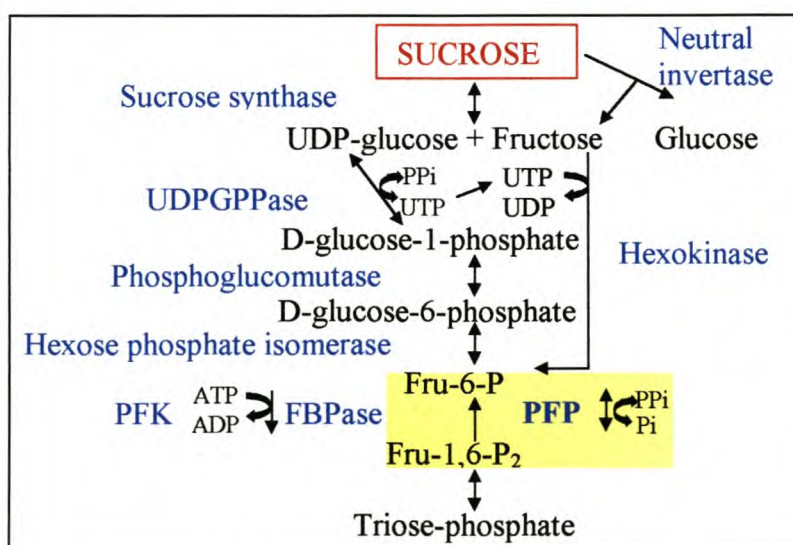
Figure 2.2 Diagram representing part of the sucrolytic sequence of reactions in the cytosol (modified from Plaxton, 1996). The reaction catalysed by PFP is highlighted.

PFP was first isolated in the eukaryote *Entamoeba histolytica* (Reeves *et al.*, 1974) and subsequently in the prokaryote *Propionibacterium shermanii* by 'O Brien *et al.* (1975). PFP is absent from animal cells, but high activity has been discovered in plant cells and since its discovery in pineapple leaves (Carnal and Black, 1979), the enzyme has been investigated and characterised in a variety of plant tissues. Carnal and Black (1983) showed that PFP is present in most photosynthetic organisms by investigating the activity of the enzyme in a photosynthetic bacterium (*Rhodospirillum rubrum*), in the algae (Charophytes), and among both lower and higher plant species. Enzyme kinetic studies reveal that in many plant tissues, PFP activity is either equal to or exceeds that of PFK (Carnal and Black, 1979; Kowalczyk *et al.*, 1984; Macdonald and Preiss, 1986; Stitt, 1990).

2.3.1 Molecular forms of PFP

In most plants, PFP is a heterotetramer composed of alpha (α) and beta (β) subunits with molecular masses of 65–67kDa and 60kDa, respectively. In wheat seedlings, Yan and Tao (1984) have isolated two forms of PFP. The larger form appears to be a tetramer composed of two subunits of 67kDa (α) and 60kDa (β) with a molecular structure consisting of $\alpha_2\beta_2$. The smaller form of PFP is a dimer of 60kDa. The composition of PFP in potato tuber is $\alpha_2\beta_2$ (Kruger and Dennis, 1987). According to Botha *et al.* (1988), the PFP proteins from bean, cucumber and potato contain both the PFP protein subunits associated with plant PFP. Cheng and Tao (1990) detected a 60kDa PFP protein from mung bean which contains both the α and β-like proteins. Similarly, PFP purified from carrot roots have a novel subunit composition in that only the β subunit is present (Wong *et al.*, 1988). It has also been shown that only the β subunit is present in some tissues of *Citrillus lanatus* (Botha and Botha, 1991). However in the cotyledons of this plant, PFP is present in the different isoforms which include the α and β subunits. Three forms of PFP have been purified from green and red tomato fruit (Wong *et al.*, 1990). The first form contains the typical α and β subunits, the second form a β- doublet subunit, while the third appears to contain a β-singlet subunit.

Carlisle *et al.* (1990) have shown that PFP α and β subunits are related, with approximately 40% of the amino acid residues being identical. However, the two subunits have been

shown to be immunologically distinct in potato tuber (Kruger and Dennis, 1987), wheat germ and lettuce leaves (Cheng and Tao, 1990).

Various lines of evidence exist to support the theory that the α and β subunits are the regulatory and catalytic subunits, respectively. Based on kinetic data, Yan and Tao (1984) proposed that the smaller form of PFP in wheat seedlings, which contains only the β subunit, is the catalytic subunit of the larger form. Carlisle et al. (1990) compared the nucleotide and derived nucleotide sequences of cDNA clones of the α and β subunits of PFP from potato tubers, with that of the known PFK enzyme from *E. coli*. The regions of homology between both types of clones and the PFK sequence from *E. coli* were residues that are involved in the binding of Fru-6-P/ Fru-1,6-P_2 in the *E. coli* PFK. The conservation of the catalytic residue (aspartic acid -175) in the PFP-β subunit with that of the residues involved in the catalytic activity of *E. coli* PFK implies that PFP-β retains the catalytic activity and is therefore the catalytic subunit. Due to the substitution of the aspartic acid residue in the α subunit with an asparagine, this subunit has been proposed to play a regulatory rather than a catalytic role. Cheng and Tao (1990) showed that the α subunit of PFP was more susceptible to trypsin digestion than the β and that PFP cleaved by trypsin retained most of its stimulated activity. Based on these results the α subunit is considered to function as a regulatory protein. Further evidence was provided by the study by Botha and Botha (1993), who demonstrated that the application of the growth substances kinetin and etherel to isolated cotyledons of *C. lanatus* caused an increase in total PFP activity by an induction of the β subunit. Levels of the α subunit were unaffected by etherel but decreased by kinetin. The correlation provided evidence that the β subunit is the catalytic moiety of the enzyme.

2.3.2 Proposed roles of PFP

In most plant tissues PFP is highly stimulated by Fru-2,6-P_2. Examples of this relationship have been shown in mung bean seedlings (Sabularse and Anderson, 1981), potato tubers (Van Schafingten et al., 1982), spinach leaves (Cseke et al., 1982), pea seeds (Wu et al., 1983), castor bean endosperm (Kombrink et al.,1984), wheat seedlings (Yan and Tao, 1984), black gram seeds (Ashihara and Stupavaska, 1984), pineapple (Carnal and Black,

1979), cucumber seeds (Botha *et al.*, 1986) and soybean cells (Macdonald and Preiss, 1986). PFP of green tomato fruit, as in other starch storing tissues, shows a marked increase in affinity for Fru-6-P substrate on activation by Fru-2,6-P$_2$, however the enzyme from sugar storing tissues (red tomato fruit) shows either little change or a decrease in affinity for the Fru-6-P substrate (Wong *et al.*, 1990). The authors suggest that the kinetic properties of the enzyme enable the tissue to store sucrose.

The role of PFP in plant tissues is uncertain (Botha and Botha, 1993). However, since PFP catalyses a reaction that is readily reversible and observed to be close in equilibrium *in vivo* (Stitt, 1990), it could theoretically function in the glycolytic or gluconeogenic direction. Stitt (1989) proposes that the properties of PFP may provide a way of modifying fluxes in response to external factors or when large amounts of carbon are being moved between starch, sucrose and organic acids. Based on observations that PFP activity is high in tissues engaged in biosynthetic activity, Dennis and Greyson (1987) propose that PFP plays a role of a biosynthetic enzyme to equilibrate pools of hexose and pentose phosphate. Stitt (1989) proposes that the properties of PFP enable the enzyme to operate as a PPi-stat, which is engaged by rising Fru-2,6-P$_2$ and/ or falling Pi. Dancer and ap Rees (1989) hold the view that PFP maintains the cytosolic concentration of PPi according to the supply and demand of PPi. Dennis and Greyson (1987) propose that PFP provides carbon skeletons during periods of high biosynthetic activity. The function of PFP has also been viewed as an adenylate bypass for glycolysis when phosphate levels are low (Duff *et al.*, 1989).

2.3.3 Sugarcane PFP

Whittaker (1997) reported differences of PFP activity within the developing stalk and between a selection of varieties, which implies that PFP in sugarcane is in part modulated by coarse regulation of transcriptional and/ or translational factors. Whittaker and Botha (1997) showed that in the internodes of sugarcane culm, there is an increase in PPi and a decrease in Fru-2,6-P$_2$ levels, so much so that it could lead to a very significant decline in the activity of PFP. From the range of calculated values for the PFP catalysed reaction in the glycolytic direction, it was concluded that the PFP catalysed reaction is in equilibrium at

all stages in the sugarcane culm. Botha *et al.* (1996) noted that sucrose increased from the young to the older internodes in sugarcane, and Whittaker (1997) observed that high levels of PFP activity remained in sugarcane internodes after growth had ceased.

Whittaker and Botha (1997) measured the activity pattern of PFP in the top 10 internodes of different sugarcane varieties in order to determine the relationship between PFP, sucrose accumulation and fibre content. PFP activity was shown to increase from internode 3 (maturing culm) reaching highest activity in internodes 6 and 7, and subsequently declining to internode 10 (mature culm). PFP activity was negatively correlated with sucrose content and positively correlated with fiber content across four commercial varieties and in 72 clones of an F1 population PFP was found to co-segregate with fibre and sucrose content. In a further study of PFP activity in sugarcane internodes with maximum rates of sucrose accumulation (between internodes 6 and 8) across varieties with known differences in sucrose content, PFP activity was shown to be inversely correlated with sucrose content and positively related to respiration (Whittaker and Botha, 1999). In addition, the latter study suggested that levels of PFP activity in sugarcane are controlled by the expression of the β subunit of the protein.

2.3.4 PFP genetic studies

PFP has been characterised by genetic analysis in relatively few plant species. In potato, cDNA clones for the α and β subunits were isolated by screening the potato tuber expression library with affinity purified antibodies (Carlisle *et al.*, 1990). Differential expression levels of the two subunit genes were observed in potato tuber, leaf and sprout tissues. The results revealed a single transcript of 2.3kb for the α subunit and 2.1kb for the β subunit. In each case, the size of the mRNA obtained was close to that predicted from the cDNA sequence and the deduced amino acid sequences of each subunit cDNA clone were predicted to produce peptides with molecular weights in close proximity to those determined by electrophoresis. This evidence suggested that the cDNA clones obtained for each subunit were full-length.

In a subsequent study by Blakeley *et al.* (1992), near full-length cDNA clones for each subunit were isolated and sequenced from a cDNA expression library derived from developing castor seed endosperm. The cDNA clones from potato tuber and castor bean were 90% identical in the derived amino acid residues for the respective subunits. The results of Blakeley *et al.* (1992) indicated a single mRNA of approximately 2kb for each subunit, in both developing and germinating castor seeds. As shown by Carlisle *et al.* (1990), it was observed that expression of the PFP genes is tissue specific. Blakeley *et al.* (1992) observed a discrepancy between the appearance of mRNA and polypeptide for the α and β subunits of PFP in germinating castor seed endosperm, and the presence of mRNAs in the absence of protein in cotyledon tissue. This implied that regulation of expression is post-transcriptional. In addition, there is transcriptional and tissue-specific regulation of expression of both subunits. Southern blot analysis of the castor genomic DNA suggests that there is a single gene for each PFP subunit. The results of Southern blot analysis of potato PFP indicates that the gene may be encoded by a multigene family or the results may be due to the polyploid nature of the plant (Blakeley *et al.*, 1992). In a subsequent study, Todd *et al.*, (1995) isolated a full length PFP-β cDNA clone from the castor bean cDNA library using part of the clone obtained previously.

Available PFP-β sequence data

Sequence data obtained from genetic studies have been entered into international databases, which can be accessed through the World Wide Web, for example GenBank. These provide a resource for discovering similarities between uncharacterised sequences with those that have known sequence identity. The full length PFP-α and β genes from *Ricinus communis* and from *Solanum tuberosum*; together with the PFP gene cDNA sequences from bacteria (*Thermoproteus tenax*, *Dicotyoglomus thermophilum*, *Amycolaptopsis methanolica*), and the full-length PFP-β gene sequence from citrus (*citrus x paradisi*) are available on the GenBank database. PFP Expressed Sequence Tags (ESTs) from *Arabidopsis*, *Oryza sativa* (rice), *Zea mays* (maize) and *Lycopersicon esculentum* (tomato) are also available on the database.

Groenewald (1998) obtained an incomplete sugarcane PFP-β cDNA clone from the mature culm library and used it to screen a sugarcane genomic library. In this study, the first five exons and 1 747bp of the 5' flanking region of the PFP-β gene from sugarcane were sequenced.

Variants of the PFP-β gene in sugarcane

Recent work on sugarcane PFP has indicated the presence of two possible variants of the PFP-β gene (Harvey, 1998). Polymerase Chain Reaction (PCR) analysis of genomic DNA and library screening of two cDNA libraries constructed from the leafroll and mature stalk showed the presence of two fragments (1250bp and 1100bp) in a commercial variety (N21) and representatives of the ancestral species (*Saccharum officinarum* and *Saccharum spontaneum*). Since the fragments were present in both the ancestral species and the commercial variety, it was concluded that the evolution of the two variants occurred before the evolution of the ancestral species. International database searches revealed that both the fragments were significantly homologous to the PFP-β gene, however sequence analysis indicated that the 1250bp and 1100bp fragment populations were 77-78% homologous in the exon regions and 48-59% homologous in the intron regions. Sequence comparison of the derived exon regions of the 1250bp fragment from N21 with the PFP-β cDNA clone isolated from the leafroll cDNA library, shows a 94% similarity. In contrast, sequence comparison of the cDNA clone with the putative exon regions of the 1100bp fragment shows a 67% similarity. From this evidence, Harvey (1998) inferred that the variant characterised by the 1250bp fragment is expressed in sugarcane leafroll.

If genes are present in multiple copies and these copies vary with respect to each other, attempts to remove the gene by genetic engineering would require specific antisense constructs of each gene copy.

2.4 Genetic manipulation of metabolic pathways

The relative quantity of carbohydrates, lipids, and proteins which are accumulated in a variety of plant storage organs, determines the utility of crop plants (Blakeley, 1997). By

manipulating this resource allocation, the nutritional or economic value of the plants can be enhanced. The glycolytic, gluconeogenic, pentose phosphate and shikimate pathways are involved in the allocation of carbon to the various biosynthetic pathways. Using genetic engineering and transformation, the levels of enzymes in these metabolic pathways can be manipulated and the relative yields of the various storage compounds may be altered (Blakeley and Dennis, 1993).

Genetic manipulation may be challenging and produce results contrary to those expected. For example, attempts to alter glycolytic flux through over-expression of pyruvate kinase resulted in complete inactivation of cytosolic pyruvate kinase in the leaves of tobacco plants (Gottlob-McHugh *et al.*, 1992). Other studies have been successful in manipulating the composition of carbon-containing storage compounds. Sucrose phosphate synthase, recognised as a site for the regulation of sucrose synthesis, has been introduced into tomato leaves where elevated levels of the enzyme resulted in the increase of sucrose levels (Worrell *et al.*, 1991). ADP-glucose phosphorylase, an enzyme important to starch biosynthesis, has also been targeted for manipulation. A bacterial ADP-glucose phosphorylase was expressed in potato tubers (Stark *et al.*, 1992). This resulted in a large increase in the starch content of these organs. In an alternate study, this enzyme activity was inhibited in transgenic potatoes. There was a reduction in starch synthesis and the tubers contained an elevated level of sucrose (Muller-Rober *et al.*, 1992).

2.4.1 Genetic manipulation of PFP

PFP has been recognised as a key regulatory enzyme in sucrose biosynthesis (ap Rees *et al.*, 1985) and is therefore an appropriate target for genetic manipulation. Hajirezaei *et al.* (1994) have genetically manipulated PFP in potato plants using antisense constructs to the genes encoding the α and β subunits of PFP. The virtual elimination of the activity of PFP (approximately 70-90% in growing plant tubers and 88-99% in stored tubers) had no effect on the growth or development of transgenic potatoes. Sonnewald (1992) introduced a bacterial pyrophosphatase into the cytosol of transgenic plants. The ability of PFP to operate in the glycolytic direction was reduced, glycolysis was no longer able to operate effectively and consequently sucrose accumulated in the leaves.

CHAPTER 3

ANALYSIS OF THE GENOME FOR PFP-β SEQUENCES

3.1 Introduction

Commercial sugarcane varieties are interspecific hybrids characterised by a high level of ploidy. Due to the high ploidy levels and an estimated chromosome number between 100 and 130 (Lu *et al.*, 1994), with a base chromosome number of 10 (D'Hont *et al.*, 1996), there may be multiple alleles of a given gene at any one locus. The enzyme PFP, which is involved in sucrose biosynthesis, is composed of the regulatory (α) and the catalytic (β) protein subunits which are encoded by the PFP-α and PFP-β genes respectively. The PFP-β gene has been shown to be a single copy gene in castor bean (Blakeley *et al.*, 1992) and may be a single copy gene in sugarcane. Since allelic variation is best established by targeting a single copy gene, the PFP-β genic region provides an appropriate model for the study of allelic variation in sugarcane.

Molecular approaches that have been used extensively to locate genes of interest include the polymerase chain reaction (PCR), molecular cloning and Southern hybridisation analysis. Since its development by Saiki *et al.* (1985), PCR has become a widespread technique used for an increasing number of different applications. The reaction involves two oligonucleotide primers, designed to flank the region of interest in the target DNA, which hybridise to opposite strands of a DNA molecule. Repetitive series of cycles of template denaturation, primer annealing and the extension of the annealed primers by DNA polymerase, results in the exponential accumulation of a specific fragment whose termini are defined by the 5' ends of the primers (Erlich, 1989). There is no single set of conditions that will be optimal for all possible reactions, thus for all primer/template combinations, different conditions have to be evaluated for the production of specific PCR fragments (Erlich, 1989). PCR offers a cheaper, easier and faster method of sequence identification in comparison to the alternative method of screening a library for a given sequence by

hybridisation (Charlieu, 1994). PCR fragments are resolved based on the size of the amplicons. However, within the amplicon population, variations within individual sequences may exist. The process of molecular cloning allows individual fragments from the amplicon population to be isolated. Then, based on sequence analysis, variations between the individual sequences may be resolved.

A large body of DNA sequence data is available on international databases. By aligning novel sequences with these previously characterised genes or proteins, insights into their common attributes and evolutionary origins may be obtained (Altschul, 1998). The Basic Local Alignment Search Tool (BLAST) (Altschul et $al.$, 1990) facilitates such an alignment. The BLAST algorithm, developed by the National Center for Biotechnology Information (NCBI) at the National Library of Medicine (USA), is a heuristic tool for finding ungapped, locally optimal sequence alignments (Altschul et $al.$, 1990). By using this algorithm, which is accessible on the NCBI website (http://www.ncbi.nlm.nih.gov/), a similarity search between a query sequence and a specific database can be made and sequence identity inferred. There are both peptide sequence databases and nucleotide sequence databases available and several programs that could be used to perform the search. Among these are the BLASTn program, which compares a nucleotide sequence query against a nucleotide sequence database, and BLASTx, which compares the six-frame conceptual translation products of a nucleotide query sequence (both strands) against a protein sequence database. The BLASTn program is frequently used to obtain identical DNA sequences and splicing patterns, while the BLASTx program, is commonly used to analyse new DNA to find genes or homologous proteins (Brenner, 1998).

Southern hybridisation provides an insight into the organisation of particular sequences within the genome (Dvorak and Dubcovsky, 1996). The technique, which has become increasingly specialised, has been used for various applications which include the determination of the copy number of particular genes (Blakeley et $al.$, 1992; Harris et $al.$, 2000), investigating the diversification of genomic regions between species (Hesslop-Harrison and Schwarzacher, 1996), and detection of polymorphisms and identification of markers (Jones et $al.$, 1997; Hahn and Grifco, 1996). The technique is based on the

restriction of genomic DNA with a class of enzymes known as restriction endonucleases, which cleave the DNA at specific nucleotide sequences or recognition sites (Sambrook *et al.*, 1989). The restricted DNA is then separated according to size on an agarose gel by electrophoresis and transferred to a filter where the DNA is fixed. Heterologous or homologous radiolabeled DNA probes are hybridised to the filter and autoradiography is used to detect the position of bands complementary to the probe.

Previous work by Harvey *et al.* (1999) suggested that there are two variants of the PFP-β gene in sugarcane. This work showed the presence of the variants in both the commercial genotype N21 and in both the ancestral noble parent *S. officinarum* and the wildtype ancestral parent *S. spontaneum.* Consequently, the ancestral *Saccharum* species representatives Black Cheribon (*S. officinarum*) and Coimbatore (*S. spontaneum*), together with the commercial hybrid cultivar N21 were selected for the present study. Another commercial hybrid cultivar, N19, which is the predominantly cultivated variety in irrigated areas and which was readily available at SASEX, produces low levels of phenolics during DNA extractions. Therefore, this variety was also included as part of the study material.

The objectives of this study were to confirm the presence of variants of the PFP-β gene and to determine whether they represent allelic forms at a single locus or genetic entities at separate loci. The molecular approach to obtain these objectives involved two stages. Firstly, the polymerase chain reaction (PCR), cloning and subsequent sequence analysis were used to investigate sequence variation of a core portion of the PFP-β gene. Thereafter Southern hybridisation analysis was used to investigate the copy number of the PFP-β gene and thus provide some insight into the nature of the observed variations.

3.2. Materials and Methods

3.2.1 Sugarcane varieties

The sugarcane cultivars Coimbatore and Black Cheribon, representing the ancestral sugarcane species *Saccharum spontaneum* and *Saccharum officinarum* respectively, were used for this study. Commercial varieties N19 and N21 were collected from the SASEX

nursery, while ancestral varieties were supplied from the museum lines by the SASEX Plant Breeding department. The sugarcane plants were stripped of the outer leaves and the soft inner leafroll was used for DNA extractions within 24 hours of acquiring the plants.

3.2.2 Sugarcane genomic DNA library

Previously, at SASEX, a Lambda FIX®II Custom Genomic Library (Stratagene) was constructed from N19 and N12 genomic DNA mixed in equal proportions and prepared using Xho I. N19 and N21 are two of the widely used cultivars in the sugar industry and between them contain a relatively broad range of genetic characteristics. The library, with an estimated titre of 1.7×10^9 pfu/ml, was stored in aliquots of 1ml at -80 ∘C.

3.2.3 Oligonucleotide primers

The degenerate PFP-B1 primer was previously designed from a region of high homology between castor bean PFP (accession number Z32850 on the GenBank database) and potato PFP (accession number J05629) cDNA fragments (Groenewald, 1998). PFP-B8 was designed from the sugarcane PFP-β sequence (Groenewald, 1998). The universal primers SP6 (Butler and Chamberlain, 1982) and T7 (Davanloo et $al.$, 1984) were designed to amplify the regions flanking the multiple cloning site found on the pGEM®-T Easy Vector (Promega). Similarly, the primers T7 and T3 (Stratagene), complementary to the T3 and T7 promoters of the pCR- Skript™ SK(+) Vector (Stratagene) were designed to amplify the region flanking the multiple cloning site of the vector. A summary of the primers and their sequences are listed in Table 3.1. All primers were synthesized by Roche Diagnostics (Germany) and diluted with sterile water to working stocks of 6μM, then stored at -20∘C.

3.2.4 Cloning vector

The 3018bp pGEM®-T Easy Vector (Promega) represented in Figure 3.1, was used for cloning. This vector has been prepared for cloning by restriction with the Eco RV enzyme and a 3' terminal thymidine (T) has been added to both ends for 'sticky-end' or T-A cloning. Efficient ligation of PCR products relies on the non-template dependent activity of Taq DNA polymerase to add a single deoxyadenosine (A) to the 3' ends of amplified DNA, compatible with the 'T' overhang on the vector.

Table 3.1 Summary of previously designed primers used in various applications.

Primer	Sequence (5' to 3')	length in bases	Target site/ binding position	Expected amplicon size
PFP-B1 Forward (Groenewald, 1998)	ATI-GAT-TTC-ATI-CCI-GAG-GT*	20	Sugarcane PFP-β gene (exon 9)	Approximately 1 200bp
PFP-B8 Reverse (Groenewald, 1998)	CCA-TCA-GTG-ATG-TCA-ATG-CTG	21	Sugarcane PFP-β gene (exon 13)	
T7 Forward (Davanloo et al.,1984)	TAC-GAC-TCA-CTA-TAG-GGC-GA	18	73bp upstream of the multiple cloning site on the pGEM®-T Easy Vector 58bp downstream of the multiple cloning site of the pCR-Skript™ SK(+) Vector	Variable depending on the size of the insert
SP6 Reverse (Butler and Chamberlain, 1982)	ATT-TAG-GTG-ACA-CTA-TAG-AAT-AC	23	99bp downstream of the multiple cloning site on the pGEM®-T Easy Vector	
T3 Reverse	AAT-TAA-CCC-TCA-CTA-AAG-GG	20	67bp upstream of the multiple cloning site of the pCR-Skript™ SK(+) Vector	

* I =inosine

34

The multiple cloning site of the vector is found within the α-peptide coding region of the β–galactosidase enzyme (which elicits a blue colour on indicator plates). Thus positive transformants can be selected based on the insertional inactivation of the α -peptide.
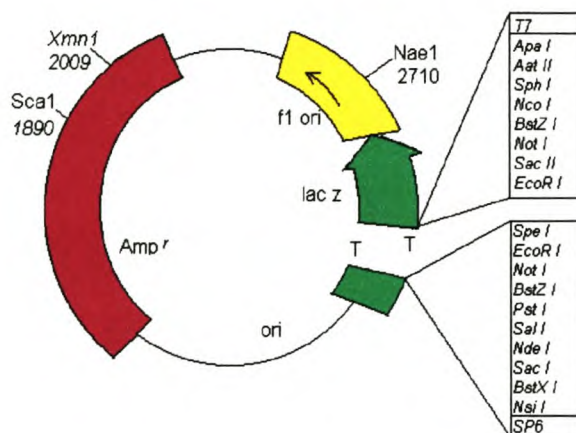
Figure 3.1 pGEM®-T Easy Vector circle map. Restriction enzyme sites flanking the multiple cloning site are shown. Blue/ white selection of recombinants is possible by the insertional inactivation of the lac-z gene. In the presence of an inducer [isopropyl-β-D-thiogalactopyranoside (IPTG)] and a chromogenic substrate [5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal)] the β-galactosidase gene product results in the formation of blue colonies. However, when a DNA fragment is inserted, the reading frame of the lac-z gene is disrupted. This interferes with expression of the gene and results in the formation of white colonies. Putative recombinants can be checked for inserts by PCR amplification with the SP6 and T7 primers. Amp^r is the ampicillin resistance gene, which is used for plasmid selection on ampicillin antibiotic medium.

3.2.5 Bacterial strain

The *Escherichia coli* strain JM109 (Promega) (Yanisch-Perron *et al.*, 1985) was used as host cell type for transformation. This strain is compatible with blue/ white screening and ampicillin selection. The strain is *recA*⁻ which prevents undesirable recombination events.

The JM109 cells were either made competent in-house (section 3.2.16) or supplied, in competent form, with the pGEM®-T Easy Vector system. The JM109 competent cells were stored at -80∘C and briefly thawed on ice prior to use in transformation.

3.2.6 Maintenance of plasmids and bacterial strains

Bacterial cultures were grown on Luria Bertani (LB) medium (1% (w/v) tryptone, 1% (w/v) NaCl, 0.5% (w/v) yeast extract powder, pH 7.5). Transformed bacterial cells were grown at 37∘C on solid LB medium (LB medium and 1.5% (w/v) agar) supplemented with 100μg/ml final concentration of the ampicillin antibiotic for plasmid selection. Working stocks, grown in petri dishes (streak plates), were sealed with Parafilm and stored at 4∘C for up to three weeks before cultures were re-streaked. Stab cultures were prepared by stabbing a needle or loop-full of cells into LB-agar (supplemented with 100μg/ml ampicillin) contained in 1.5ml Corning™ vials. The cells were allowed to grow overnight at 37∘C, with the lids of the vessels loose, before lids were tightened. The stab cultures were then maintained in the dark at room temperature for up to 6 months. Bacterial cultures grown in LB medium with 100μg/ml ampicillin, were mixed with 15% (v/v) glycerol, frozen in liquid nitrogen and stored at -80∘C for long term storage of cultures.

3.2.7 Full length PFP-β cDNA clones

Full-length PFP-β cDNA clones from potato (accession number JO5629, cloned in the pCR-Skript™ Vector) were provided as a gift.

3.2.8 Isolation of DNA from sugarcane

Modified Dellaporta method

DNA was routinely extracted from sugarcane using a modified Dellaporta method (Dellaporta *et al.*, 1983). Five stalks of sugarcane, collected from different individuals of a single variety, were stripped of their outer leaves and the soft inner leafroll exposed. The leafroll tissue was sliced into discs approximately 2mm thick. Six grams of pooled tissue were weighed out, frozen in liquid nitrogen and ground to a fine powder in a mortar and pestle. The tissue was transferred to 35ml extraction buffer (100mM Tris-HCl, 500mM

NaCl, 50mM EDTA and 1% (v/v) β-mercaptoethanol, pH 8.0), vortexed and 3.5ml 20% (w/v) sodium dodecyl sulphate (SDS) added. Samples were incubated at 70∘C for 1.5hr and 7ml of 5M potassium acetate added. The tubes were agitated gently and incubated on ice for 20min. The samples were centrifuged for 15min at 10 322xg in a Beckman Avanti™ J-251 centrifuge and the supernatant filtered through four layers of gauze cloth into a new tube. Isopropanol (3/4 volume) was added to each tube and mixed gently. DNA was spooled out using sterile glass hooks (prepared by flaming the end of a Pasteur pipette) and dissolved in 1ml TE buffer (10mM Tris-HCl, 1mM EDTA, pH 8.0) at 37∘C overnight. Protein was removed by adding an equal volume of chloroform: isoamylalcohol [24:1(v/v)], vortexing and centrifuging at 4 032xg for 10min. The DNA-containing aqueous phase was retained and the chloroform: isoamylalcohol clean-up repeated if any protein remained in the sample. The DNA was stored in 1.5ml Corning™ vials at -20∘C.

Alternative DNA extraction methods included the modified Honeycutt method (Honeycutt *et al.*,1992), a quick DNA extraction method and use of the Nucleon™ PhytoPure™ Kit (Amersham).

Honeycutt method

Six grams of leafroll tissue, suspended in 40ml of homogenisation buffer [50mM Tris-HCl (pH 8.0), 25mM EDTA (pH 8.0), 0.5mM spermidine, 1% (w/v) polyethylene glycol (PEG) 8000, 0.1% (v/v) β- mercaptoethanol and 0.35M sucrose], was homogenised for 2min using the Ultra-Turrax T25 homogeniser. The homogenate was filtered through 2 layers of sterile cloth and centrifuged for 20min at 5 000xg at 4∘C. The pellet was re-suspended in 10ml wash buffer [50mM Tris-HCl (pH 8.0), 0.5mM EDTA, 0.5mM spermidine, 0.1% (v/v) β-mercaptoethanol and 0.35M sucrose]. Two milliliters 5M NaCl, 1ml 10% (w/v) SDS and 1.3ml 10% (w/v) cetyltrimethylammonium bromide (CTAB) were added to the tube on ice (with stirring after the addition of each reagent). The samples were incubated at 60∘C for 30min, cooled at room temperature for 15min and proteins extracted by the addition of an equal volume of chloroform: isoamylalcohol [24:1(v/v)]. Tubes were inverted until the two phases emulsified and subsequently centrifuged at 3 500xg for 10min at 4∘C. The aqueous phase was removed and the chloroform extraction repeated. DNA was extracted from the

aqueous phase by the addition of an equal volume of isopropanol. The tubes were gently inverted and the DNA spooled out using sterile glass hooks. The DNA was re-suspended in 1ml TE buffer (pH 8.0).

Quick DNA extraction method

Rapid DNA extraction was achieved by incubating 50mg of the grindate in 1ml of the extraction buffer [50mM Tris (pH 8.0), 5mM EDTA (pH 8.0), 1% (w/v) PEG 8000, 0.5mM spermidine, 0.1% (v/v) β-mercaptoethanol, 0.5% (w/v) SDS and 1M NaCl] at 70°C for 1hr. The samples were centrifuged for 10min and the supernatant removed. A 3/4 volume of isopropanol was added to the aqueous phase and the DNA pelleted by centrifugation at 14 000rpm in a microcentrifuge at 4°C. The pellet was washed once with 70% ethanol and re-suspended in 40μl sterile water.

Nucleon™ PhytoPure™ Kit (Amersham)

One gram of leafroll tissue was frozen and ground under liquid nitrogen in a mortar and pestle. The grindate was transferred to a 14ml Falcon® tube and lysis was achieved by adding 4.6ml reagent 1, 1.5ml reagent 2, followed by incubation at 65°C for 10min and on ice for 20 min. DNA was extracted by the addition of 2ml chloroform (pre-equilibrated to _20°C) and 200μl resin suspension, followed by incubation at room temperature for 10min (with regular manual agitation). The sample was centrifuged for 10min at 1 300xg and the DNA was recovered from the upper phase by precipitation with an equal volume of cold isopropanol. The DNA was pelleted for 5min at 4 000xg, washed with cold 70% ethanol, air-dried and re-suspended in 50μl sterile water.

3.2.9 New primer design

Degenerate primers, PFP-B SR forward and PFP-B SR reverse, were designed to amplify across the putative exon 10, 11 and 12 region of the sugarcane PFP-β gene using the Consensus-Degenerate Hybrid Oligonucleotide Primers (CODEHOP) software available on the World Wide Web (Rose *et al.*,1998). The amino acid sequences of the PFP-β gene from several species: Castor Bean (GI number 2499489 on the GenBank database), Potato (GI number 2507174), Citrus (GI number 3790100) and a sugarcane cDNA clone

sequence (Groenewald, 1998) were aligned using Clustal X (1.8) multiple sequence alignment. This aligned data was imported into CODEHOP where the Blockmaker program selected blocks of highly conserved homologous protein sequences. The software generated possible forward and reverse primers with a short 3' degenerate core region and a longer 5' consensus clamp region. The PFP-B SR forward and PFP-B SR reverse primer pair, selected from putative exon 9 and 13 of the sugarcane PFP-β gene respectively, was synthesized by Roche Diagnostics (Table 3.2). The level of degeneracy is given as the product of the number of possible bases in each position. The lower and upper case letters indicate the degenerate core and the non-degenerate consensus clamp regions of the primer respectively. Degeneracies are defined below the table.

Table 3.2 Selected degenerate primers and their sequences.

Primer	Sequence 5' -3'	Degeneracy
PFP-B SRfor	CGGAGGGCCTGATCgayttyathcc*	12
PFP-B SRrev	TCCATCAGGGCGGTCarngcngtncc*	128

* y =C/T	h =A/C/T	r =A/G	n =A/C/G/T

The selected primer sequences and their relative position to the PFP-B1/ PFP-B8 primers on the castor bean exon regions are shown in Figure 3.2.
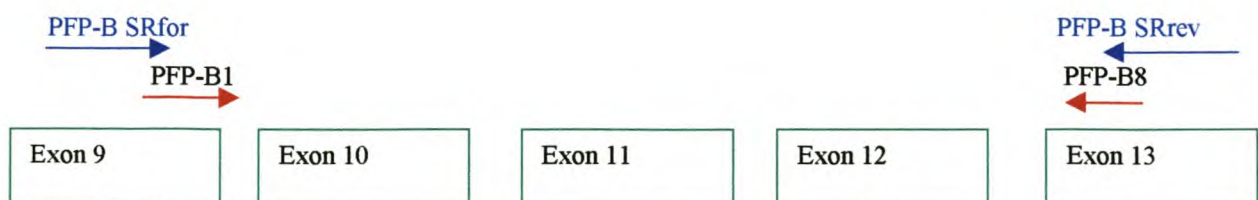
Figure 3.2 Diagram representing the positions of the PFP primers in relation to the castor bean (*Ricinus communis*) cDNA sequence. The primers amplify across exons 10, 11 and 12, and four intron regions. The predicted size of the PFP-B SR forward/reverse product is 12bp larger than the PFP-B1/B8 amplicon.

3.2.10 PCR amplification of part of the PFP-β gene

Amplification using the PFP-B1/ PFP-B8 primers

PCR was performed in 50μl reaction volumes on a Perkin Elmer GeneAmp® PCR System 9700 thermocycler. One PCR reaction contained: 1x *Taq* DNA Polymerase magnesium free buffer (Promega), 1.5mM $MgCl_2$, 0.33mM deoxy-nucleoside triphosphates (dNTPs), 0.2μM each primer, 1 unit *Taq* DNA polymerase (Promega) and 75ng template. DNA was amplified using the following thermal profile: 1 cycle of 94°C for 90sec; 10 cycles of 94°C for 45sec, 45°C for 30sec, 72°C for 2min; 30 cycles of 94°C for 30sec, 40°C for 30sec, 72°C for 90sec and two holds of 72°C for 2min and 35°C for 5min.

Amplification using the PFP-B SRfor /PFP-B SRrev primers

Amplification with the new degenerate primers was accomplished using a touchdown PCR. One PCR reaction contained 1x GeneAmp® PCR buffer II (10mM Tris-HCl, pH 8.3, 50mM KCl), 1.5mM $MgCl_2$, 0.1mM dNTPs, 0.2μM each primer, 2.5 unit Ampli*Taq* Gold™ (Perkin Elmer) and 150ng template in a volume of 30μl. The reaction was overlaid with 50μl mineral oil (Sigma) and cycled in a Hybaid OmniGene thermocycler. Cycling included a single heat activation cycle of 95°C for 12min, followed by 1 cycle of 95°C for 1min, 50°C for 1min, 72°C for 2min; 1 cycle of 95°C for 1min, 49°C for 1min, 72°C for 2min; 1 cycle of 95°C for 1min, 48°C for 1min, 72°C for 2min; 1 cycle of 95°C for 1min, 47°C for 1min, 72°C for 2min; 1 cycle of 95°C for 1min, 46°C for 1min, 72°C for 2min; 35 cycles of 95°C for 1min, 45°C for 1min, 72°C for 2min, and two holds of 72°C for 5min and 35°C for 5min.

3.2.11 Gel electrophoresis and visualisation of DNA

One percent (w/v) agarose gels were prepared using TBE buffer (45mM Tris-borate, 1mM EDTA, pH 8.0) and agarose. Loading dye (20% (w/v) ficoll, 0.15% (w/v) bromophenol blue) was added to the samples, which were loaded on to the gel alongside molecular weight marker III or a 100bp DNA step ladder (Promega) (Figure 3.3). Samples were electrophoresed at 6V/cm until the dye front had migrated to the end of the gel. Gels were stained in 1μg/ml final concentration ethidium bromide solution for 30min and de-stained in

1mM $MgSO_4$ for 30min. DNA fragments were visualised on a Hoefer Mighty Bright UV transilluminator at 300nm and photographed using a Vilber Lourmat photographic system. Pictures were printed on Sony high density, light sensitive paper using a Sony video graphic printer.
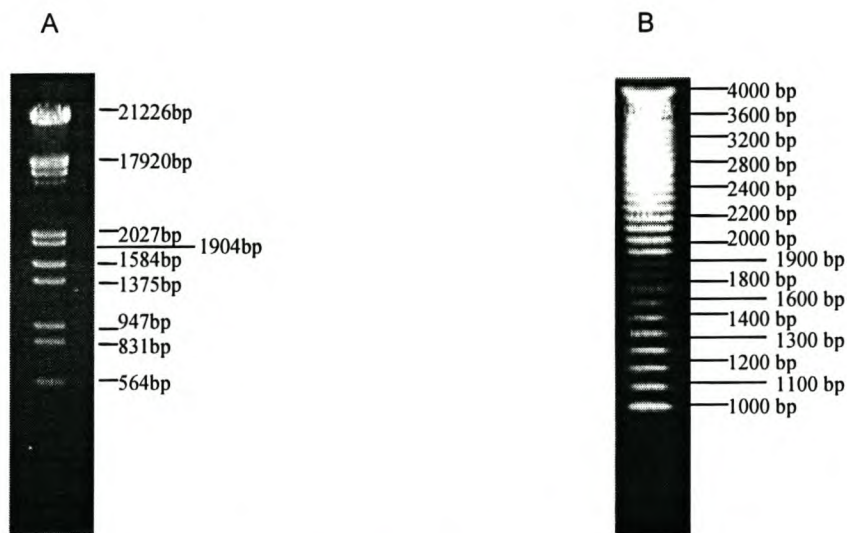
Figure 3.3 Electrophoretic separation of lambda molecular weight marker [III] and a 100bp DNA stepladder. Lambda molecular weight marker [III] (A) was prepared by digesting $20\mu g$ of bacteriophage lambda DNA (Boehringer Mannheim) with 20 units of each of the restriction enzymes *Eco* RI and *Hind* III for three hours at 37°C. A $200\mu l$ volume of TE buffer (pH 8.0) and $30\mu l$ loading dye (40% (v/v) glycerol; 1mM EDTA; 0.02% (w/v) bromophenol blue; 0.02% (w/v) xylene cyanol) was added to the digested sample. Ten microliters of marker [III] was used per lane for agarose gel electrophoresis. The 100bp DNA stepladder (Promega) (B) was used to differentiate between fragments less than 100 base pairs apart. A 1x final concentration blue/ orange loading dye (Promega) was added to $1.5\mu l$ of the step ladder and used per agarose gel lane.

3.2.12 Purification of PCR products

PCR fragments were excised from the gel with a sterile scalpel blade and the DNA purified using the Qiaquick Gel Extraction Kit (Qiagen). Gel slices were weighed in colourless 1.5ml tubes and $300\mu l$ solubilising buffer was added to every 100mg of gel slice. The samples

were incubated at 50∘C for 10min with vortexing every 2min to completely dissolve the agarose. The mixture was immediately added to a DNA binding column placed on top of a collection tube. The sample was centrifuged at 12 000rpm in a Sorvall® MC 12V microcentrifuge for 60sec to bind the DNA, and the flow-through was discarded. The column was washed with 750μl buffered ethanol and centrifuged at 12 000rpm for 60sec. Centrifugation at 12 000rpm for 60sec was repeated to remove any residual ethanol. The DNA was eluted with 30μl sterile water by adding the water directly to the center of the column, incubating at room temperature for 5min, then centrifuging at 12 000rpm for 60sec. The DNA was stored at -20∘C.

3.2.13 Concentration of purified DNA

Qiagen extracted DNA had to be precipitated or concentrated for cloning and sequencing applications.

DNA was precipitated using 0.1x volume of 3M sodium acetate (pH 5.2) and 2.5x volume ice-cold 100% ethanol (Moore, 1999). The sample was incubated at –20∘C for 30min before pelleting the DNA by centrifugation at 14 000rpm for 30min at 4∘C in a microcentrifuge. The pellet was washed with 70% ethanol, air-dried and re-suspended in a smaller volume of sterile water.

Alternatively, DNA was concentrated in a Savant SpeedVac SC 110 until the volume had decreased to approximately 5μl.

3.2.14 DNA quantification

DNA was quantified using either unrestricted lambda standards on an agarose gel, or fluorometry. Occasionally both fluorometry and lambda standards were used and the DNA concentrations obtained by both methods were compared and adjusted accordingly.

Lambda DNA (Boehringer Mannheim) standards of 62.5ng, 125ng, 250ng and 500ng were run alongside an aliquot of the DNA sample on a 1% (w/v) agarose gel. The DNA

concentrations were estimated by comparing the intensity of the lambda standards to that of the sample DNA band.

The Hoefer® DyNA Quant 200 fluorometer was used to determine the concentration of $2\mu l$ of DNA in 2ml of assay solution. A low range assay solution, $0.1\mu g/ml$ Hoechst 33258 dye in 1x TNE (0.2M NaCl, 10mM Tris-Cl, 1mM EDTA, pH 7.4), was prepared fresh from concentrated stocks for DNA estimated to be 10- 500ng/ml. The high range assay solution, $1\mu g/ml$ Hoechst 33258 dye in 1x TNE, was required for higher concentrations of DNA (100- 5 000ng/ml). The fluorometer was calibrated using an undiluted 1mg/ml DNA standard (Calf thymus DNA, ultrapure, Sigma D4764) for the high range assay or a 1:10 dilution of the DNA standard for the low range assay. The instrument was zeroed with 2ml of assay solution before and after calibration, and between DNA samples. An average of three fluorometer readings per sample was recorded as the DNA concentration.

3.2.15 DNA quality assessment

The quality of extracted DNA was assessed by measuring the absorbance of a 1:10 dilution of the sample at 260nm (A_{260}) and 280nm (A_{280}) on a Beckman DU® 7500 spectrophotometer. Since pure DNA has a A_{260}/A_{280} ratio of 1.8 (Sambrook *et al.*, 1989), DNA samples registering values as close as possible to the latter ratio were selected for PCR.

3.2.16 Cloning PCR fragments in *E.coli*

Insertion of PCR fragments into the vector

Ligation reactions containing 50ng pGEM®-T Easy Vector, 1x Rapid Ligation buffer, T4 DNA Ligase (Promega), an appropriate amount of PCR product for a 3:1 insert: vector ratio, and 3 Weiss units of T4 DNA ligase in a final volume of 10 μl, were incubated at 4∘C overnight. The amount of PCR product required for a 3:1 insert: vector ratio was calculated as follows:

$$\frac{50\text{ng vector X PCR product size in kb}}{\text{kb vector}} \quad \frac{X \quad 3}{1}$$

Ligation reaction mixtures were used directly for transformation or stored at 4°C.

Preparation of JM109 competent cells

Competent cells were prepared according to the protocol described by Sambrook *et al*. (1989). A two hundred millilitre volume of LB broth was inoculated with 5ml of an overnight culture of JM109 cells. The culture was incubated at 37∘C with 150rpm agitation until the absorbance, at a wavelength of 650nm, measured between 0.2- 0.4 units. The cells were then chilled on ice for 10min and centrifuged at 645xg in 4x 50ml tubes for 10min at 4∘C. Each pellet, maintained on ice, was re-suspended in 10ml transformation buffer (60mM $CaCl_2$, 10mM pipes-NaOH, pH 7.0, 15% (v/v) glycerol). The contents of two tubes were pooled and incubated on ice for 30min. The cells were pelleted once more, by centrifugation at 1 008xg for 10 min at 4∘C, and re-suspended in 6ml transformation buffer. Two hundred microlitres of the cells were aliquoted into 1.5ml sterile tubes, flash frozen in liquid nitrogen and stored at –80∘C.

Preparation of selection plates

LB agar plates containing 100$_\mu$g/ml of the antibiotic ampicillin were used within 30 days of preparation. Twenty microlitres 50mg/ml X-gal and 100$_\mu$l 0.1M IPTG, for blue/white selection, were spread on to the surface of the plate using a glass spreader and allowed to absorb at 37°C for 30min prior to use.

Transformation of E. coli with plasmid vector

Fifty microliters of JM109 competent cells were added to 2$_\mu$l of the ligation reaction in a Falcon® tube on ice. The tube was gently mixed by flicking, and kept chilled on ice for 20min, before the cells were heat shocked for 45-50sec at 42∘C. The cells were then placed on ice for a further 2min and 950$_\mu$l SOC medium (2% (w/v) tryptone, 0.5% (w/v) yeast extract, 10mM NaCl, 2.5mM KCl, 20mM Mg^{2+}, 20mM glucose, pH 7.0) at room temperature, was added. After incubation at 37∘C for 1.5hr with agitation (150rpm), 100$_\mu$l

of each transformation mixture was plated on to duplicate selection plates. The remaining cells were pelleted by centrifugation at 1 000xg for 10min, re-suspended in 200μl SOC medium, and 100μl plated onto separate selection plates. The plates were incubated at 37∘C overnight and then at 4∘C for 2hr to facilitate blue/ white selection.

Selecting positive transformants using direct PCR of bacterial cells

White colonies, selected as putative positive transformants based on blue/ white screening, were subcultured on to a master LB agar plate with 100μg/ml ampicillin using a sterile toothpick, and subsequently tested for the insert using PCR. One PCR reaction contained in a final volume of 15μl: 1x *Taq* DNA Polymerase magnesium free buffer (Promega), 1.5mM MgCl$_2$, 100μM dNTPs, 0.2μM SP6 primer, 0.2μM T7 primer, 0.1 unit *Taq* DNA polymerase (Promega) and a toothpick-end aliquot of cells from selected white colonies. Cycling conditions were as described in section 3.2.10 for the PFP-B1/ PFP-B8 primers. The expected fragment size for positive transformants, as shown by agarose gel electrophoresis, was calculated to be approximately 133bp larger than the insert size, due to the flanking vector sequence.

3.2.17 Plasmid isolation

Positive colonies were grown in 5ml LB broth, with 100μg/ml ampicillin, by overnight incubation at 37∘C with 150rpm agitation. Plasmids were isolated using either the Nucleobond Plasmid Isolation Kit (Machery-Nagel) or a rapid plasmid isolation protocol.

Nucleobond Plasmid Isolation Kit

The transformed cells were harvested by centrifugation at 3 000xg for 10min at 4∘C and the supernatant discarded. Plasmid DNA was isolated from the bacterial cells using the Nucleobond AX20 Plasmid Isolation Kit (Macherey-Nagel). The pelleted bacterial culture was re-suspended in suspension buffer (50mM Tris-HCl, 10mM EDTA, 100μg/ml RNase A, pH 8.0). Subsequently, lysis buffer (200mM NaOH, 1% SDS) was added, the solution mixed gently, and incubated at room temperature for 5min. Immediately upon adding the denaturing buffer (2.8M K acetate, pH 5.1), the suspension was incubated on ice for 5min to denature the proteins. The bacterial lysate was clarified of debris by centrifugation at

12 000xg for 15min at 4°C and subsequently loaded on to a Nucleobond cartridge pre-equilibrated with equilibration buffer (100mM Tris, 15% ethanol, 900mM KCl, pH 6.3). The cartridge was washed three times with wash buffer (100mM Tris, 15% ethanol, 1 150mM KCl, pH 6.3), before the plasmid DNA was eluted with the elution buffer (100mM Tris, 15% ethanol, 1 000mM KCl, pH 8.5). The DNA was precipitated with isopropanol at room temperature, washed with 70% ethanol and re-suspended in 50μl sterile water.

Rapid plasmid DNA isolation protocol

Cultures were centrifuged at 1.4xg for 10min in a swinging bucket rotor at 4°C. The supernatant was poured off and tubes drained on a paper towel before 200μl cold STET buffer [8% (w/v) sucrose, 5% (w/v) Triton X-100, 50mM Tris (pH 8.0), 50mM EDTA] was added to the pellet. The pellets were re-suspended by vortexing and subsequently transferred to 1.5ml tubes maintained on ice. Twenty microlitres of freshly prepared 10mg/ml lysozyme (Roche) was added to each tube and boiled for 1min. The samples were centrifuged at 14 000 rpm in a microcentrifuge at 4°C for 30min and the supernatant removed. An equal volume of phenol: chloroform: isoamylalcohol [25:24:1(v/v)] was added to the supernatant, vortexed briefly and centrifuged at 12 000rpm at room temperature for 10min. The supernatant was retained and another phenol: chloroform: isoamylalcohol [25:24:1 (v/v)] extraction performed. The supernatant was then subjected to chloroform:isoamylalcohol [24:1 (v/v)] extraction by adding an equal volume of the latter solution, vortexing and centrifuging for 10min at 12 000rpm at room temperature. The plasmid DNA was precipitated from the supernatant by adding 200μl of ice-cold isopropanol, maintaining the sample at –80°C for 10min, followed by 1-2hr at –20°C and centrifugation at 14 000rpm for 45min at 4°C. The pellet was dried under a partial vacuum in the SpeedVac and re-suspended in 30μl sterile water.

Plasmid DNA was quantified as described in section 3.2.14.

3.2.18 DNA sequence determination

Sequencing makes use of multicolour fluorescent dye terminator chemistry where terminal bases are covalently attached to different dyes, which can be differentiated by their

respective wavelengths. Sequencing reactions were performed using a Big Dye Sequencing Kit (Perkin Elmer). Each 20μl reaction contained 6μl terminator premix (A-dye terminator, C-dye terminator, G-dye terminator, T-dye terminator, dGTP, dATP, dCTP, dTTP, Tris-HCl (pH 9.0), MgCl$_2$, thermostable polymerase and Ampli*Taq* DNA polymerase FS), 500ng plasmid DNA from selected positive clones and 0.6μM forward or reverse primer. The polymerisation products were extended using 25 cycles of the following thermoprofile: 96∘C for 10sec, 50∘C for 5sec, 60∘C for 4min.

The extension products were precipitated by incubation with 60% (final concentration) non-denatured ethanol at room temperature for 10-15min, followed by centrifugation at maximum speed in a microcentrifuge. Once the supernatant was aspirated, the pellet was washed twice with 100μl 70% ethanol, recovered by centrifugation, aspiration, and dried under a partial vacuum (SpeedVac). Eighteen microlitres of template suppression reagent (Perkin Elmer) was added to the pellet and the mixture denatured in a boiling water bath for 2min. The sample was quenched on ice for 1min, before being centrifuged briefly and transferred to autosampler tubes which were subsequently loaded into the ABI Prism 310 Genetic Analyser (Perkin Elmer).

The extension products were automatically injected into a single sequence capillary filled with a pre-formulated liquid polymer, Performance Optimised Polymer-6 (POP-6) (Perkin Elmer). Extension products were resolved by electrophoretic separation at 6V/cm for approximately 2.5hr and nucleotide sequences were read by the Genetic Analyser.

3.2.19 Sequence analysis

Sequence Navigator

Sequences were manually edited on Sequence Navigator v1.0.1, 1991 (Perkin Elmer), a software programme installed on the automated sequencer. Ambiguities were replaced with the base, at the position of the ambiguity, corresponding to the coloured peak on the electropherogram considered the most likely candidate.

DNASIS

The Contig Manager or Shotgun fragment matching function of DNASIS v2.6 (Hitachi) was used to arrange the forward and reverse sequences of a particular clone with respect to each other. The function checks the ends of sequences in both the sense and antisense orientations and locates regions of overlap. Gaps may also be inserted in overlapping regions. However, the degree of mismatch is governed by the matching stringency parameter. A minimum matching homology of 60% in overlapping regions and a minimum overlap size of 10 bases were specified in order for matches to be scored as contigs. Contig Manager generates a full-length consensus sequence of clones. Thus, possible sequence artifacts were resolved by overlapping several forward and reverse sequences of each clone. Restriction enzyme site analyses were also performed using DNASIS.

International database searches

Sequence identity was queried using the National Center for Biotechnology Information (NCBI) Basic Local Alignment Tool (BLAST). Full-length consensus sequences or separate forward and reverse sequences were submitted to the non-redundant database using either the BLASTn (Altschul *et al.*, 1990) or BLASTx (Altschul *et al.*, 1998) program. BLASTn compares the nucleotide query sequence to the database target sequence and a statistical score is assigned to each alignment. Scores above 80 represent significant homology to a particular sequence in the database. In the same manner, the BLASTx program compares the six-frame translation products of the nucleotide query against sequences in the protein database. Nucleotide sequences were also queried against the BLAST EST (Expressed Sequence Tag) database to determine their homology to known expressed sequences. All BLAST queries were filtered to remove scores with low sequence homology and gapped alignments were allowed.

3.2.20 Southern hybridisation analysis

Restriction digestion of genomic DNA

Fifteen microgram samples of genomic DNA were incubated with 1x appropriate restriction enzyme buffer (Roche), 40 units of enzyme (Roche) and 8mM spermidine in a 400μl total

volume, at 37∘C for 16hr. The reaction was terminated by placing the samples at 4∘C for 2hr. Success of restriction was checked by separating 10μl aliquots of each digest on a 0.8% (w/v) agarose gel. If the DNA had not restricted completely, the samples were incubated for another 16hr with a further 10 units of enzyme, and restrictions confirmed electrophoretically.

Purification of restricted DNA

The restricted DNA was incubated at 37∘C for 30min with 1μl RNase A (10mg/ml) (Boehringer Mannheim) to remove any residual RNA present in the sample. The samples were then subjected to phenol: chloroform: isoamylalcohol [25:24:1 (v/v)] purification to ensure that the DNA was free of enzyme and other protein contamination. An equal volume of tris-saturated phenol: chloroform: isoamylalcohol [25:24:1 (v/v)] was added to the samples, vortexed and centrifuged at 14 000rpm in a microcentrifuge for 10min. The DNA was precipitated from the aqueous phase by the addition of 0.05x volume 5M NaCl and 2.5x volume absolute ethanol, incubation at -80∘C for 30min, and centrifugation at maximum speed for 30min in the cold. The pellet was air-dried and re-suspended in 25μl TE buffer (pH 8.0). To ensure complete solubilisation of the DNA, the samples were incubated at 4∘C overnight and at 37∘C for 2hr.

Gel electrophoresis and visualisation of DNA

Restricted DNA was separated on 0.8% (w/v) agarose gel by electrophoresis at 50V in 0.5x TBE until the dye front had migrated to the end of the gel. Gels were stained, de-stained and visualised according to section 3.2.10. A Hoefer ruler, fluorescent under UV light, was photographed alongside the gel to assist with size-determination of DNA fragments after Southern hybridisation.

Upward capillary blotting

The gel containing the separated, restricted DNA was depurinated in 0.125M HCl for 10min with agitation on the Belly Dancer® (Stovall), rinsed in polished water and subsequently denatured in transfer buffer (1.5M NaCl, 0.5M NaOH) for 30min. Stacks for upward capillary blotting were set-up according to Figure 3.4. After overnight blotting, the gels were

stained and viewed to ensure that all the DNA had been transferred to the membrane. The membrane was then UV cross-linked in a Hoefer UV Crosslinker at 120 000$_\mu$J/cm^2 with the DNA side up, air-dried and stored at 4$\circ$C.
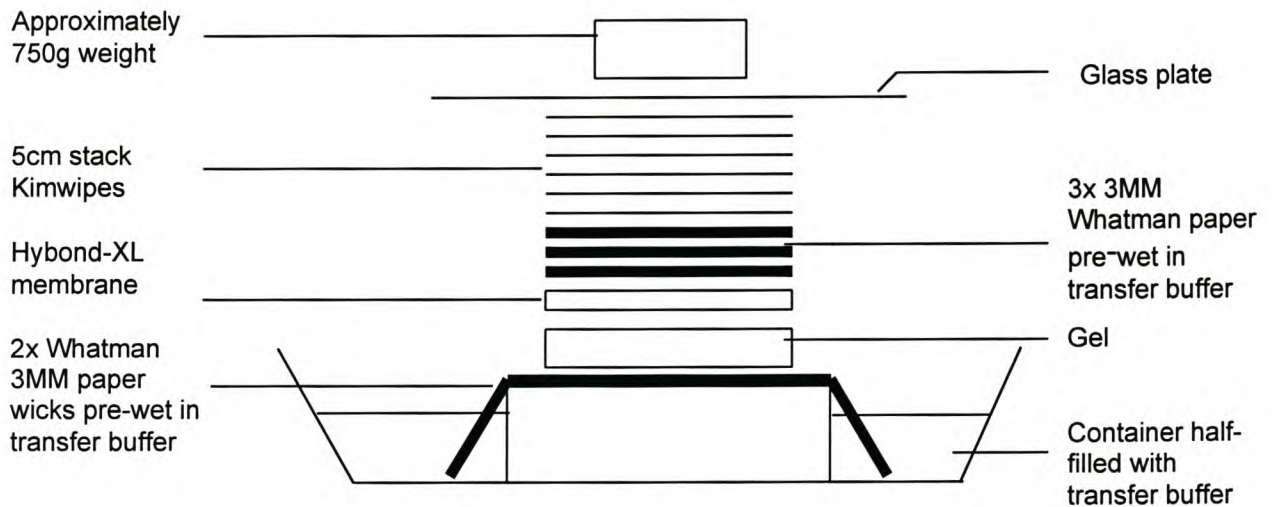
Figure 3.4 Diagram representing the upward capillary apparatus for Southern blotting. Depurinated and denatured 0.8% (w/v) agarose gels containing 15μg restricted DNA samples were blotted overnight by the above method. The Hybond-XL membrane (Amersham) was cut on one corner to orientate the membrane in relation to the gel, and UV cross-linked once transfer of DNA was complete.

Probe generation and purification

Cloned PFP-B1/B8 fragments were amplified from plasmid DNA using the specific primers PFP-B1 and PFP-B8. The potato PFP-$_\beta$ cDNA clone was amplified using the vector primers T7 and T3, and the castor bean PFP-$_\beta$ cDNA clone with SP6 and T7 primers. A 30$_\mu$l PCR reaction contained 1x *Taq* DNA polymerase magnesium free buffer (Promega), 1.5mM MgCl$_2$, 0.1mM dNTPs, 0.2$_\mu$M each primer, 0.1 unit *Taq* DNA polymerase and an equivalent of 20ng plasmid DNA. Cycling conditions were as described in section 3.2.10 (*Amplification using the PFP-B1/ PFP-B8 primers*). The PCR products were separated on a

0.8% (w/v) agarose gel and extracted using the Qiaquick Gel Extraction Kit (Qiagen) as described in section 3.2.12.

Probe labeling

The purified amplicon was labeled using the Amersham Megaprime Labeling Kit. Initially, 20ng of amplicon DNA was boiled together with 5μl of random nonamer primer solution in a total volume of 10μl for 5min. Ten microlitres of labeling buffer (dATP, dGTP and dTTP in Tris-HCl pH 7.5, 2-mercaptoethanol and $MgCl_2$), 2μl of Klenow enzyme, 5μl $\alpha^{32}P$ radioactively labeled dCTP (specific activity 6000 Ci/mmol, Amersham) and polished water to 50μl were added at room temperature. The reaction was mixed by pipetting and incubated at 37∘C for 1hr. The reaction was stopped by the addition of 20μl 50mM EDTA (pH 8.0).

The efficiency of the labeling reaction was checked by paper chromatography and autoradiography. One microlitre of the probe was spotted on to a piece of cellulose paper (25TLC plastic sheets, 20 x 20cm, PET cellulose, Merck) and solvent (0.75M Na_2HPO_4) was allowed to rise 10-12cm up the paper. The paper was placed in a cassette and exposed to X-ray film (Hyperfilm™-MP, Amersham) for 15min before the film was developed. A successfully labeled probe was represented by a large dark spot, where the probe had been spotted, and occasionally the appearance of slight smearing towards the top of the paper (indicative of unincorporated nucleotides).

Purification of radiolabeled DNA

If the labeled probe contained a large proportion of unincorporated nucleotides, the probe solution was purified using a NucTrap® Probe Purification Column (Stratagene). The column with blue caps removed, was prewet with 70μl 1x STE buffer (100mM NaCl, 20mM Tris-HCl pH 7.5, 10mM EDTA).The buffer was added to the top of the column and forced down the length of the column using a syringe with a Luer-Lok tip screwed onto the column. A 1.5ml collection tube was placed in the Beta shield base and the NucTrap® Probe Purification column placed three fourths of the way in the column-locking mechanism with the end of the column above the tube. The radiolabeled sample (70μl)

was added to the top of the column and the syringe (with the plunger extended) applied to the top of the column. The column was subsequently moved down until the syringe fitted snugly against the column beta shield and the syringe cover was placed over the syringe. Direct, constant pressure was applied so that the sample was forced through the column into the tube. The syringe was removed and the column rinsed by forcing 70μl 1x STE buffer through the column using the syringe. A bolus of air was pushed through the rinsed column to dislodge any remaining liquid. The sample was checked with a Geiger-Muller counter (series 900 mini-monitor). A high radioactive signal (above 20 counts per second) was an indication that the probe labeling had been successful.

Preparation of fragmented salmon sperm DNA

Salmon sperm DNA (250mg, Sigma) was dissolved overnight in 50ml TE (pH 8.0) at room temperature after which 8ml aliquots were dispensed in to 15ml Corning™ vials. The DNA was sonicated, using a Sonicator® Ultrasonic Processor (Misonix) in 1min pulses for 7 to 10min with a 1min cooling interval between each pulse on ice. In order to determine whether the fragments were of the correct size (500bp), the samples were combined and 1μl run on a 0.8% agarose gel at 50V for 1hr. The DNA was cleaned by performing two phenol: chloroform: isoamylalcohol [25:24:1 (v/v)] extractions. The supernatant was retained and the DNA precipitated by adding 0.1x volume 3M Na acetate, pH 7.0, and 2x volume 100% ethanol, placing on ice for 2hr and centrifuging at 23 200xg for 30min at 4°C. The pellet was washed with 70% ethanol, air-dried and re-suspended in 2ml of TE (pH 8.0). The concentration of the DNA was determined using fluorometry, diluted to a concentration of 20mg/ml and stored at –20°C. The DNA was thawed, denatured by boiling for 5min and snap cooled on ice prior to use.

Southern hybridisation

A 100x Denhardt's solution (Denhardt, 1966) was prepared by adding 2% (w/v) each of bovine serum albumin (BSA), polyvinylpyrrolidone (PVP) and Ficoll™. Membranes were pre-hybridised and hybridised in a solution containing 5x SSC (750mM NaCl, 75M sodium citrate, pH 7.0); 5x Denhardt's solution, 0.5% (w/v) SDS and 25μl of 20mg/ml denatured, fragmented salmon sperm DNA; pre-equilibrated to 65°C. The volume of pre-hybridisation

solution added was calculated as $70\mu l/cm^2$ of the membrane surface. The membrane was incubated in prehybridisation solution overnight, with rotation, in a Hybaid oven at 65∘C. The probe was boiled for 5min, quenched on ice and added to fresh hybridisation solution in which the membrane was incubated.

After overnight hybridisation at 65∘C, the membrane was rinsed briefly at room temperature and then again for 10min at 65∘C with 2x SSC, 0.1x SDS. The membrane was washed twice with 1x SSC, 0.1% (w/v) SDS for 15min each at 65∘C, followed by two washes with 0.1x SSC, 0.1% (w/v) SDS for 10min each at 65∘C. To ensure removal of the unbound probe, the volume of wash solution used was approximately half the volume of the hybridisation bottle. The membranes were rinsed in distilled water, briefly drained and placed between a sheet of folded plastic. The radioactive signal read between 10 and 20 counts per second on the Geiger-Muller counter.

Autoradiography

The membrane was exposed to X-ray film (Hyperfilm™-MP, Amersham) in a cassette for 5 days at –80∘C. Following exposure, the X-ray was developed for 5min in developer (Ilford Phenisol), briefly rinsed in stop solution [3% (v/v) acetic acid], and fixed for 2min in fixative solution (Ilford Hypam). The X-ray was washed under tap water and allowed to dry before being viewed on the visible light box (Hoefer). In instances where the radioactive signal was weak, the exposure time was increased (14 days).

Phosphorimaging

The Packard Cyclone™ Storage Phosphor system was used for phosphorimaging. A storage phosphor screen, which consists of a thin layer of $BaFBr:Eu^{2+}$ phosphor crystals, was placed, phosphor layer down, over the membrane. The membrane and screen were maintained in the dark, at room temperature, for 1-4 days between perspex covers clamped together. The phosphor crystals absorb the energy emitted by the radioactivity on the membrane and in turn, emit this stored energy in the form of blue light once exposed to a red laser. The intensity of energy emitted is proportional to the amount of radioactivity on the membrane. The phosphor screen was wrapped around the carousel drum and placed

into the Cyclone_{TM} scanner where the entire screen was scanned and an image of the signal obtained. The OptiQuant_{TM} software was used to edit and analyse the image. Prior to reuse, the screen was exposed to a light box, with UV light filtered out, for 2min to erase the signal on the screen.

Stripping membranes for reprobing

Membranes were kept damp before stripping. A boiling solution of 0.1% (w/v) SDS was poured over the membranes and subsequently incubated at 65∘C for 1hr. The membranes were checked with a Geiger-Muller counter for any signs of radioactivity before being rinsed in 2x SSC, blotted dry and stored at 4∘C. Membranes were stripped and probed three times before being discarded.

3.3 Results

3.3.1 Preliminary genomic DNA extraction

Sugarcane genomic DNA was extracted from selected sugarcane varieties using the modified Dellaporta method (section 3.2.8). Figure 3.5 shows a typical absorbance spectrum of a DNA sample extracted using this method.
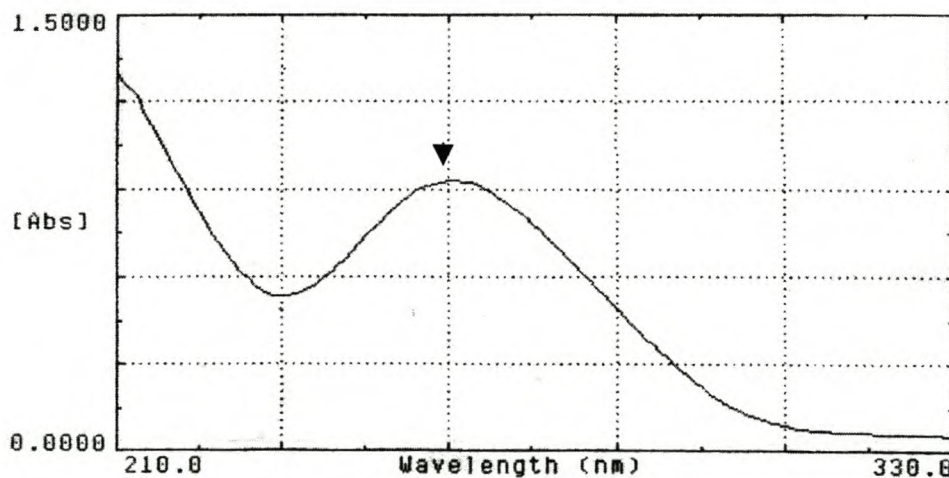
Figure 3.5 Absorbance spectrum of a DNA sample extracted using the modified Dellaporta

method. The arrow indicates the absorbance at 260nm (A_{260nm}= 0.915).

The purity values of the samples were determined by calculating the A_{260nm}/A_{280nm} ratio. The purity of the sample described above was:

$$A_{260}/A_{280}$$
$$= 0.915/0.545$$
$$= 1.7$$

This suggests that the sample was relatively pure since pure DNA has an A_{260}/A_{280} ratio of 1.8. The purity of genomic DNA extracted from ancestral and commercial sugarcane genotypes is shown in Table 3.3.

Table 3.3 The range of purity of genomic DNA from different sugarcane varieties. DNA was isolated using the modified Dellaporta method.

Variety	Number of extractions	Average yield in ng/μl	Range of purities (A_{260nm}/A_{280nm})
Black Cheribon	6	274.00	1.7 – 2.0
Coimbatore	6	175.33	1.8 – 1.9
N21	4	288.80	1.5 – 1.8
N19	4	191.00	1.8 – 2.2

Pure RNA has an A_{260}/A_{280} ratio of 2.0 (Sambrook *et al.*, 1989). Purity values greater than 1.8, as shown in Table 3.3, may be indicative of RNA contamination in the DNA preparations, while values less than 1.8 may be due to residual protein contamination which could theoretically decrease the A_{260}/A_{280} ratio (Sambrook *et al.*, 1989).

3.3.2 Optimisation of PFP-B1/ B8 PCR using sugarcane genomic DNA as template

Harvey _et al_. (1999) obtained two distinct PFP-$_\beta$ fragments when amplifying with the PFP-B1 and PFP-B8 primers on ancestral and commercial genotypes. For the purposes of the current investigation, it was important to reliably reproduce the latter results of Harvey _et al_., (1999) in order to confirm the presence of variants of the PFP-$_\beta$ gene. Initially, using the DNA extracted as in 3.3.1 and the PCR conditions described by Harvey _et al_., (1999), the two PFP-$_\beta$ fragments were observed. Figure 3.6 shows the results of initial PFP-B1/B8 PCR amplification using the ancestral and commercial sugarcane genotypes.
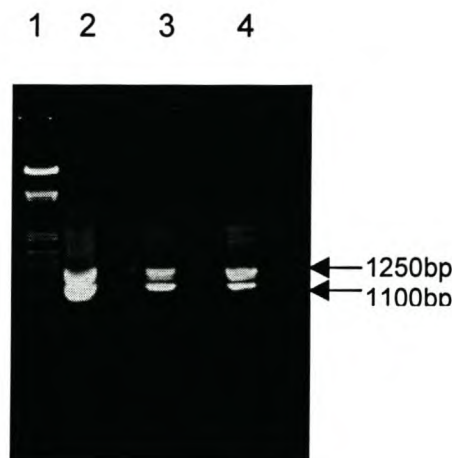
Figure 3.6 Agarose gel showing initial PFP-B1/B8 PCR amplification products using the protocol described by Harvey _et al_. (1999). The gel shows both the 1100bp and 1250bp fragments from the commercial genotype N19 (lane 2) and two ancestral species representatives: Black Cheribon (lane 3) and Coimbatore (lane 4). Lane 1, lambda molecular weight marker III.

The PCR conditions were sub-optimal and the results were not consistently reproducible. Figure 3.7 shows a set of variable results. Many PCR reactions using the same protocol produced entirely negative results (not shown).
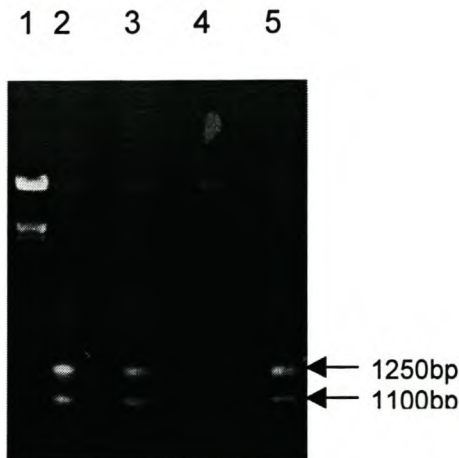
Figure 3.7 Inconsistent results of subsequent PCR amplification using the PFP-B1/B8 primer pair and original PCR protocol. Lane 1, lambda molecular weight marker III; lanes 2 and 3, N21; lane 4, Coimbatore and lane 5, Black Cheribon.

Poor amplification may be attributed to several factors involved in the PCR reaction. According to Erlich (1989), PCR conditions have to be optimised for each template and primer combination. In order to improve the PFP-B1/B8 PCR on genomic DNA, different variables were considered for manipulation.

Template considerations

Successful PCR amplification depends on the quality and amount of the template DNA, therefore these factors were investigated using N21 genomic DNA.

i) Template quality

The reagents commonly used to purify nucleic acids (e.g. salts, organic solvents and EDTA) are strong inhibitors of *Taq* DNA polymerase. Therefore the purity of DNA samples obtained by different extraction techniques were compared using N21 as the source material. These were, in addition to the modified Dellaporta method, the Honeycutt method,

the PhytoPure Plant Extraction Kit (Promega) method and a 'quick' DNA extraction method (section 3.2.7). Each DNA sample was checked using spectrophotometry and the purity of the sample assessed from the respective A_{260}/A_{280} ratios. Table 3.4 summarises the purity of the N21 genomic DNA samples extracted using different isolation techniques.

Protein contamination, which could physically impede the activity of *Taq* DNA polymerase, was removed by three phenol: chloroform: isoamylalcohol [25:24:1 (v/v)] washes, three chloroform: isoamylalcohol [24:1 (v/v)] washes and subsequent precipitation of the DNA sample (section 3.2.12). This process ensured that residual phenol, which could significantly decrease the purity of the DNA samples (Sambrook *et al.*, 1989), was removed.

The yield of the DNA samples was assessed using both fluorometry and quantification on a gel using lambda DNA as a standard. Figure 3.8 shows a lambda DNA quantification gel containing aliquots of the DNA samples extracted using the different techniques. The average yield obtained for each DNA sample is summarised in Table 3.4.

Table 3.4. The yield and purity of N21 genomic DNA samples extracted using different DNA isolation techniques.

Extraction technique	Yield (ng/μl)	A_{260nm}/A_{280nm}
Dellaporta Method	325	1.5 – 1.8
Honeycutt Method	250	1.5 – 1.6
PhytoPure Kit	160	1.1 – 1.5
'Quick' method	730	1.6 – 1.7

As shown by the A_{260}/A_{280} ratios in Table 3.4, the Dellaporta method yielded DNA of a higher purity in comparison to the other techniques. Although the DNA samples extracted using the 'quick' method of extraction had relatively high purity (Table 3.4), there was evidence of RNA contamination in this sample on the quantification gel (Figure 3.8, lane 5)

as well as in those samples extracted using the Dellaporta method (Figure 3.8, lanes 8 and 9).
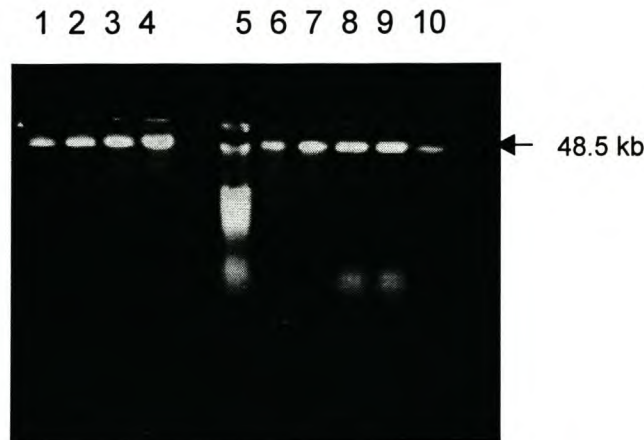
Figure 3.8 DNA quantification using lambda DNA standards as references. Lanes 1-4: Lambda DNA samples. An equivalent of 62.5ng (lane 1), 125ng (lane 2), 250ng (lane 3) and 500ng (lane 4) DNA was loaded. Lanes 5-10, N21 genomic DNA samples extracted using different methods. One microliter aliquot of each sample was loaded. Lane 5, 'quick' extraction method; lanes 6 and 7, Honeycutt method, lanes 8 and 9, modified Dellaporta method and lane 10, PhytoPure kit extraction.

The presence of contamination in the DNA samples resulted in the overestimation of DNA concentration using gel quantification. Fluorometry on the other hand, provided an accurate means of quantification, however this technique was sensitive to residual contaminants and pipetting errors. Therefore an average of three fluorometer readings were selected and averaged.

The DNA samples obtained by the different extraction techniques were used as template DNA in PFP-B1/B8 PCR. No amplification was observed (results not shown) which suggested that poor template quality was not exclusively responsible for the failed PCR.

Those DNA samples extracted using the Dellaporta method, which showed high purity ratios (close to 1.8), were selected for use in further PCR optimisation experiments.

ii) Template quantity

An incorrect amount of template may result in poor amplification. A range of DNA concentrations was used for PCR to determine the appropriate amount of template required for the PFP-B1/B8 PCR amplification. Figure 3.9 shows the results of the PFP-B1/B8 amplification using different template concentrations.
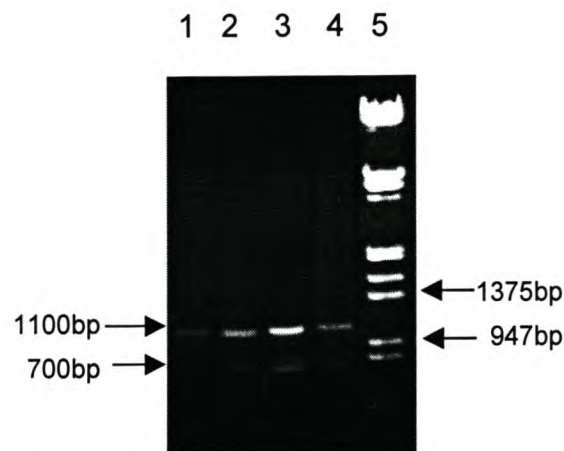
Figure 3.9 The effect of increasing template concentration on the PFP-B1/B8 PCR. Template concentrations were 30ng (lane 1), 50ng (lane 2), 75ng (lane 3) and 100ng (lane 4) of N21 genomic DNA. Lane 5, lambda molecular weight marker [III].

The 1100bp fragment amplified under the different template concentrations and an additional fragment was obtained (Figure 3.9). While the 1250bp PCR product was not amplified in the reaction, interpretations could be made about template concentrations. DNA concentrations lower than 50ng and greater or equal to 100ng appeared to decrease the yield of the PCR product. The use of 75ng template DNA resulted in the strongest amplification. This concentration of DNA, which was used in the original PFP-B1/B8 PCR, was considered an adequate template concentration for amplification of the 1250bp and 1100bp fragments.

iii) Genomic DNA complexity

The complexity of the sugarcane genome may confound specific PCR amplification. It was thought that by reducing the complexity of the sugarcane genomic DNA and targeting a specific size range of fragments, the PFP-B1/B8 PCR amplification could be improved. Two approaches were undertaken to decrease the complexity of sugarcane genomic DNA.

The first approach involved the restriction of $10\mu g$ genomic DNA with an enzyme that did not cut within the 1250bp fragment (_Eco_ R$_I$) (based on preliminary restriction site analysis of fragments obtained by Harvey _et al._, 1999). The resulting fragments were separated on a 1% (w/v) agarose gel and fragments between 0.8 and 1.5kb in size were excised. The excised fragments were purified using the QIAquick Gel Extraction Kit (Qiagen) and $1\mu l$ and a 1:10 dilution was used for PCR. Figure 3.10 shows the separation of the restricted fragments on a 1% (w/v) agarose gel (A), the excision of the 0.8 –1.5kb fragments (B) and the results of the PFP-B1/B8 amplification on aliquots of the restricted sample (C).
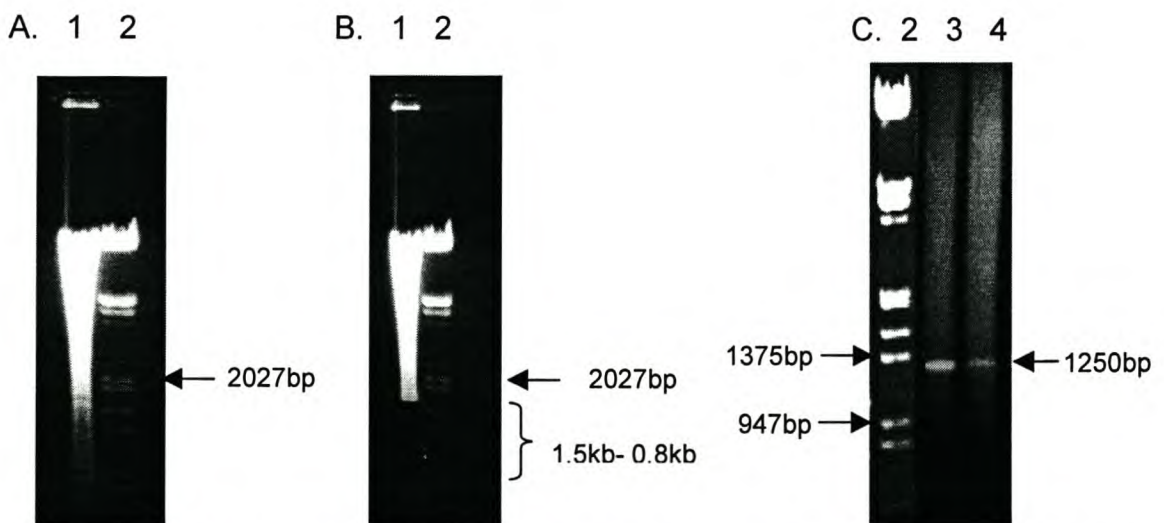
Figure 3.10 Effect of reducing template complexity, by restriction digestion, on PCR. (A) Agarose gel showing the separation of _Eco_ R^I digest of $10\mu g$ N21 genomic DNA. (B) Fragments between 0.8 and 1.5kb (as outlined by the box) were excised and gel extracted. (C) Results of PFP-B1/B8 amplification using aliquots of the 0.8 –1.5kb, _Eco_ R^I restricted fragments of N21 genomic DNA as template. Lane 1, _Eco_ RI digest of $10\mu g$ of N21 genomic DNA; lane 2, lambda molecular weight marker ^{III}; lane 3, $1\mu l$ aliquot of restricted fragments and lane 4, $1\mu l$ aliquot of a 1:10 dilution of the restricted fragments.

As expected, the 1250bp fragment amplified since the _Eco_ RI enzyme restricts within the 1100bp fragment. The higher template concentration resulted in stronger amplification (Figure 3.10, lane 3), however the presence of non-specific products is evident on the gel. The second approach to reducing the complexity of the sugarcane genomic DNA involved using a sugarcane genomic DNA library as template in PFP-B1/B8 PCR. The library was prepared using _Xho_ I, an enzyme that does not restrict within either fragment (based on preliminary restriction site analysis of the PFP-B1/B8 fragments obtained by Harvey _et al_., 1999). A 1:10 dilution, 1μl and 10μl aliquots of the N12/ N19 genomic library (section 3.2.2) were used as template according to the protocol described in section 3.2.10. Results of the PFP-B1/B8 amplification are shown in Figure 3.11.
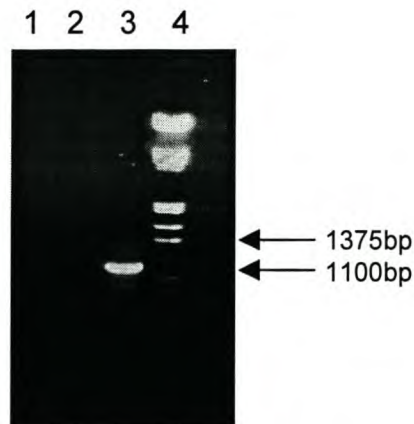
Figure 3.11 Results of PFP-B1/B8 amplification using a sugarcane genomic DNA library as template. Template volumes were 0.1μl (lane 1), 1μl (lane 2), and 10μl (lane 3) of the library. Lane 4, lambda molecular weight marker III.

A single PFP-B1/B8 amplification product was obtained (1100bp) (Figure 3.11). However, the 1250bp fragment failed to amplify from the genomic DNA library, which suggested that another PCR parameter had to be adjusted.

The 1100bp and 1250bp products were excised from the gel and purified using the protocol outlined in section 3.2.12. Aliquots of the products were used as template in separate PFP-B1/B8 PCR as described in section 3.2.10. These attempts at reamplification of the bands were unsuccessful and resulted in smearing on the agarose gel (results not shown).

Magnesium concentration

The performance of _Taq_ DNA polymerase depends on the concentration of magnesium in the reaction. High Mg^{2+} concentrations favour the production of primer dimers, decreases the enzyme fidelity and increases the level of non-specific products. Conversely, in the absence of adequate free magnesium, _Taq_ DNA polymerase is inactive and PCR product yield is low (Erlich, 1989). A Mg^{2+} titration experiment was performed to determine the optimal $MgCl_2$ concentration for the PFP-B1/B8 PCR of genomic DNA. Figure 3.12 shows the results of the Mg^{2+} experiment.
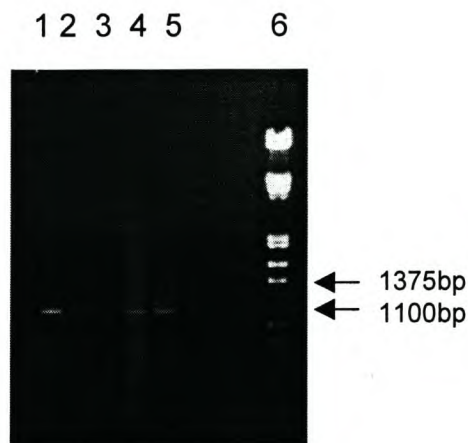
Figure 3.12 The effect of increasing the magnesium concentration on PFP-B1/B8 PCR. Lane 1, 1.5mM $MgCl_2$; lane 2, 2mM $MgCl_2$; lane 3, 4mM $MgCl_2$; lane 4, 6mM $MgCl_2$; lane 5, 8mM $MgCl_2$ and lane 6 lambda molecular weight marker [III].

Although amplification was observed for all the magnesium concentrations tested, the reaction containing 1.5mM $MgCl_2$ appeared to produce the strongest amplification. Higher magnesium concentrations (6mM and 8mM) increased the incidence of non-specific products, which were evident as excessive smearing flanking the PCR product. The optimal $MgCl_2$ concentration for the PFP-B1/B8 primer and genomic DNA template

combination appears to be 1.5mM final concentration in accordance with other PCR methods (Erlich, 1989).

PCR additives

A variety of PCR additives have been shown to increase the yield, specificity and consistency of PCR reactions.

Bovine serum albumin (BSA) concentrations of 0.1mg/ml or as much as 1mg/ml in the PCR may be beneficial in that it stabilises *Taq* DNA polymerase (Erlich, 1989). Glycerol at recommended concentrations of 5-15% (v/v) also stabilises the enzyme and improves the yield of the PCR product (Erlich, 1989).

DMSO in concentrations between 5-10% (v/v) has been suggested to increase amplification efficiency and specificity of PCR. These adjuvants have been shown to improve the amplification of some products and decrease the amplification of others (Erlich, 1989). DMSO is thought to reduce the secondary structure and has been particularly useful for GC-rich templates however, Erlich (1989) found that 10% DMSO had a slight inhibitory effect on *Taq* DNA polymerase and decreased the overall yield of the PCR product.

These additives were tested for the PFP-B1/B8 primer and genomic DNA template combination. Figure 3.13 shows the results of PFP-B1/B8 amplification using different combinations of PCR additives.

Although amplification was poor, the PCR containing 5% (v/v) DMSO appeared to have resulted in amplification of both the 1100bp and 1250bp fragments. In an attempt to confirm the effect of DMSO on the PFP-B1/B8 PCR, two successive experiments were performed using the concentration of DMSO that resulted in amplification. The results are shown in Figure 3.14.
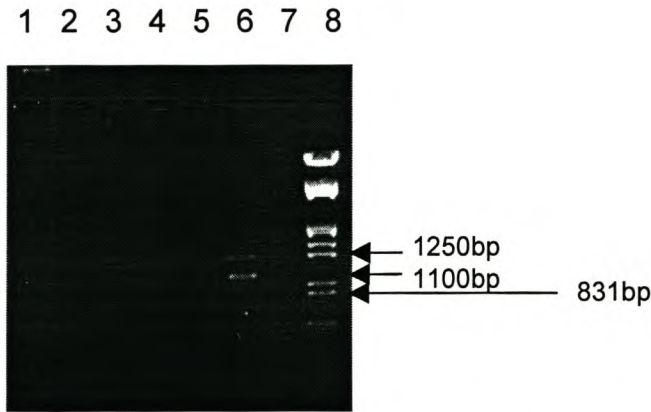
Figure 3.13 The effect of PCR additives DMSO, BSA and glycerol on the PFP-B1/B8 PCR. The reactions contained: no additives (Lane 1); 0.1mg/ml BSA (lane 2); 0.8mg/ml BSA (lane 3); 0.1mg/ml BSA, 5% (v/v) DMSO and 5% (v/v) glycerol (lane 4); 0.8mg/ml BSA, 5% (v/v) DMSO and 5% (v/v) glycerol (lane 5); 5% (v/v) DMSO (lane 6) and 5% (v/v) glycerol (lane 7). Lane 8, lambda molecular weight marker III.
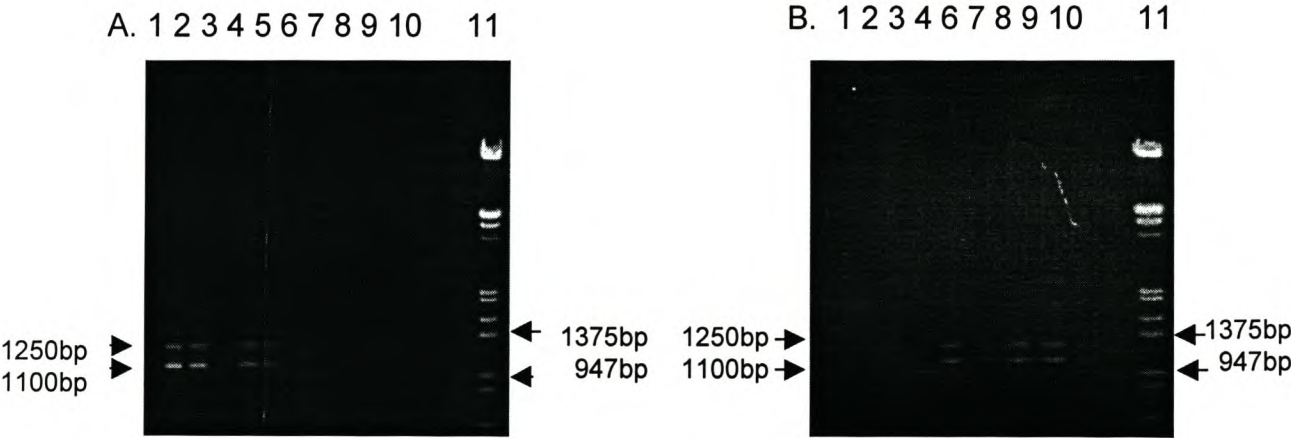
Figure 3.14 The effect of 5% (v/v) DMSO on two successive PFP-B1/B8 PCR performed under the same conditions (A and B). Lanes 1-5, 5% (v/v) DMSO; lanes 6-10, no DMSO and lane 11, lambda molecular weight marker III.

Initially both the 1100bp and 1250bp fragments were amplified under PCR conditions that included 5% (v/v) DMSO while reactions without DMSO failed. However, in a subsequent

reaction (Figure 4.14 B), the opposite was observed. It was therefore concluded that the addition of DMSO was not reliable in improving amplification of the PFP-β products.

Alternative PCR protocol

The Ready- to- Go® PCR beads (Pharmacia Biotech) were used as an alternative PCR protocol. The beads, which have been optimised for PCR reactions, contain buffer, nucleotides and *Taq* DNA polymerase. Each reaction contained 1.5 units *Taq*, 10mM Tris-HCl (pH 9.0), 50mM KCl, 1.5mM MgCl$_2$, 200μM each dNTP and stabilisers, including BSA, in a final reaction volume of 25μl. Three reactions were set up using PFP-B1 and PBP-B8 primers and 75ng template DNA. Reaction 1 contained 0.2μM primer, reaction 2: 0.3μM each primer and reaction 3: 0.4μM each primer. All three reactions failed to amplify the 1100bp and 1250bp fragments (results not shown).

Primer considerations

i) Primer stocks

Multiple freeze-thaw cycles of primer stocks could decrease the primer integrity, so the PFP-B1 and PFP-B8 primers were re-synthesized and fresh stocks were prepared for PCR. The primer stocks did not appear to be responsible, since the 'new' primer stocks also resulted in poor amplification.

ii) Primer design

Poor PCR amplification was thought to be due to the design of the PFP-B1/B8 primer pair. Although the PFP-B8 primer contained inosine for loose binding, the primers were specific and would therefore not be able to anneal to templates that contained base mutations at the targeted primer sites. The degenerate CODEHOP primers (PFP-B SRfor and PFP-B SRrev) were designed to anneal to variable bases at the targeted primer sites.

Thermocycling Profile

Due to the degenerate nature of the PFP-B SRfor and SRrev primers, it was necessary to perform a touchdown PCR. This procedure involves decreasing the annealing temperature by one degree every cycle, followed by cycling at a low annealing temperature for 10 or more cycles. This enriches for the correct product over any incorrect product since any differences in the Tm between the incorrect and correct annealing resulted in a two-fold difference in product amount per cycle (Don *et al.*, 1991).

In order to favour the amplification of the correct product, the thermocycling profile also included a hot start with the Ampli*Taq* Gold™ enzyme (Perkin Elmer). This method is designed to prevent the non-specific annealing of primer to template. Results of the PFP-B SR for/ rev amplification using the hot start, touchdown protocol is shown in Figure 3.15.
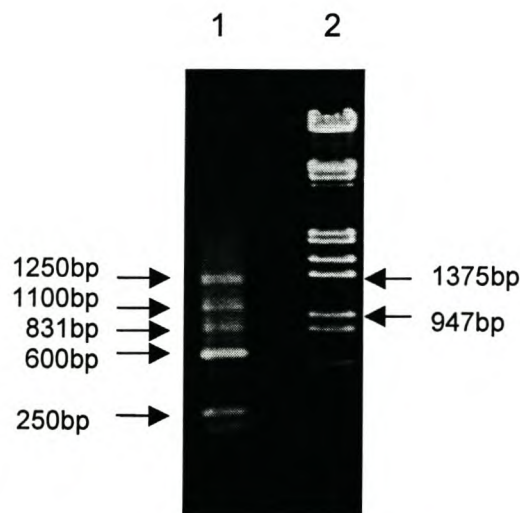
Figure 3.15 Results of the PFP-B SRfor and PFP-B SRrev PCR amplification of genomic DNA using a combined touchdown, hot start protocol (lane 1). Lane 2, lambda molecular weight marker III.

As shown in Figure 3.15, the PFP-BSRfor and PFP-BSRrev primers amplified the desired products (1100bp and 1250bp fragments). However, extra amplification was obtained (831bp, a dominant 600bp and faint 250bp fragment). Even after increasing the annealing temperature, the 831bp and 600bp fragments were observed (results not shown). It was thought that some of the extra fragments produced by the PFP-BSRfor and PFP-BSRrev degenerate primers (Figure 3.15) may be non-specific. However these fragments warranted further investigation which is described in section 3.3.3.

Since the PFP-B SRfor and PFP-B SRrev primers produced extra fragments, the PFP-B1 and PFP-B8 primers were used in combination with the hot start and touchdown protocol. Figure 3.16 shows the results of the PFP-B1/B8 amplification.
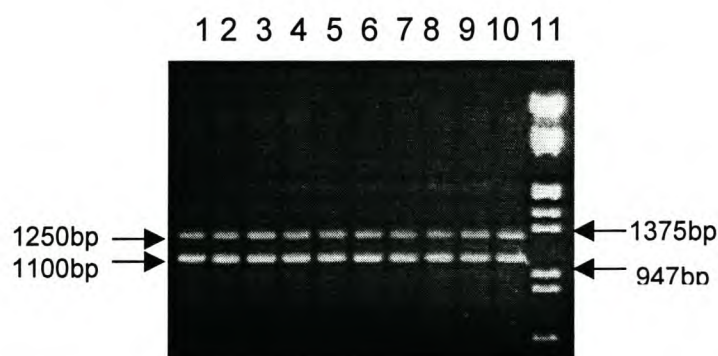
Figure 3.16 Results of the PFP-B1/B8 amplification of N21 genomic DNA using the combined hot start, touchdown PCR protocol. Lanes 1-10, replicates of the PFP-B1/B8 PCR using a hot start and touchdown protocol and lane 11, lambda molecular weight marker III.

Both the 1100bp and 1250bp fragments amplified with high reproducibility (Figure 3.16). The faint 700bp PFP-B1/B8 amplification product obtained previously (Figure 3.9) was not observed in the PFP-B1/B8 amplification using the combined hot start/ touchdown protocol (Figure 3.16). This suggested that the 700bp fragment observed in Figure 3.9 was a non-specific PCR product produced due to sub-optimal conditions of the PCR parameters. The

hot start, touchdown PFP-B1/B8 PCR was also performed on N19 genomic DNA and similar results (as shown in Figure 3.16) were obtained. Thus, using this optimised PCR protocol, two PCR fragments, were consistently obtained when genomic DNA from two commercial varieties, N19 and N21, was amplified with the PFP-B1 and PFP-B8 primers. These fragments were designated PFP-B1/B8$_{1100bp}$ and PFP-B1/B8$_{1250bp}$ in reference to their respective sizes and the primers used to amplify them.

3.3.3 Sequence identity of the PFP-B SRfor and PFP-B SRrev genomic amplicons

The sequence identity of the extra amplification products obtained from the PFP-B SRfor and PFP-B SRrev amplification (approximately 831bp and 600bp fragments shown in Figure 3.15) was investigated in order to determine whether the fragments originated from PFP-β. The 831bp and 600bp fragments were excised from the gel, extracted (as described in section 3.2.12) and sequenced directly without the intervention of cloning. These sequences were compared to the international database using the BLASTn search tool. The results indicated that the two fragments had no homology to PFP-β. Instead, the 831bp fragment showed low homology to part of the *Chlamydia muridarum* genome (score (bits)= 42 and E= 0.23) while the 600bp fragment had low homology to part of the *Beta vulgaris* mitochondrion (score (bits)= 40 and E= 0.82). The PFP-B SRfor and PFP-B SRrev primers were not reliable in amplifying the PFP-β genic region. Thus, only the PCR products produced with the PFP-B1/B8 primers (Figure 3.16) were considered for further analysis.

3.3.4 Cloning and preliminary sequence analysis of the PFP-B1/B8 fragments

The PFP-B1/B8$_{1100bp}$ and PFP-B1/B8$_{1250bp}$ fragments from N21 and N19 were cloned using the pGEM®-T Easy Vector System (Promega) and introduced into JM109 *E. coli* competent cells by transformation. White colonies were PCR amplified using the SP6 and T7 primers, as described in section 3.2.16, in order to verify that the PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ fragments were inserted. Figure 3.17 shows the result of this PCR on whole bacterial cells.
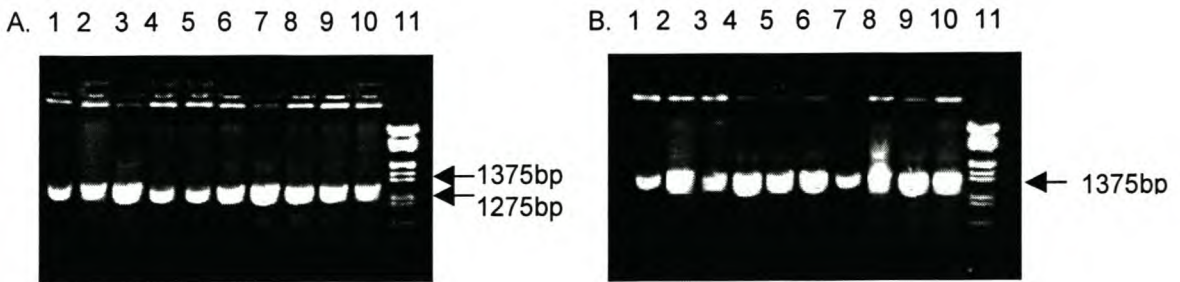
Figure 3.17 Results of SP6 and T7 PCR on whole bacterial cells to select clones containing the PFP-B1/B8$_{1100bp}$ insert (A, lanes 1-10) and the PFP-B1/B8$_{1250bp}$ insert (B, lanes 1-10). Lane 11, lambda molecular weight marker [III]. The size of the SP6/T7 PCR products, which included the vector arms, were approximately 175bp larger than the insert sizes.

Using PCR, 67% of the white colonies tested for the N19 PFP-B1/B8$_{1250bp}$ insert were positive, while 71% of those tested for the N19 PFP-B1/B8$_{1100bp}$ insert were positive. Seventy five percent and 67% of the white colonies subjected to colony PCR were positive for the PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ fragments (derived from N21) respectively. Twenty positive PFP-B1/B8$_{1100bp}$ clones, 10 from N19 and 10 from N21, and 20 positive PFP-B1/B8$_{1250bp}$ clones, 10 each from N19 and N21, were selected for plasmid extractions.

Once the clones were grown overnight and subjected to plasmid extractions, the 1100bp and 1250bp fragments were sequenced using the forward (T7) and reverse (SP6) vector primers. An internal sequencing primer (1250int) was designed in order to generate a significant overlap within the 1250bp fragment. This primer sequence (synthesised by Roche Diagnostics, Germany) was derived from conserved regions of the 1250bp fragment forward sequence, and had the following sequence:
5' - CTG CAC ATT TCA GAG GGC AAG C - 3'.

Full-length sequences were generated in the Contig Manager function of DNASIS, which arranges pieces of sequences from a particular clone with respect to one another. Figure 3.18 represents an example of the full-length sequence for a 1250bp clone.
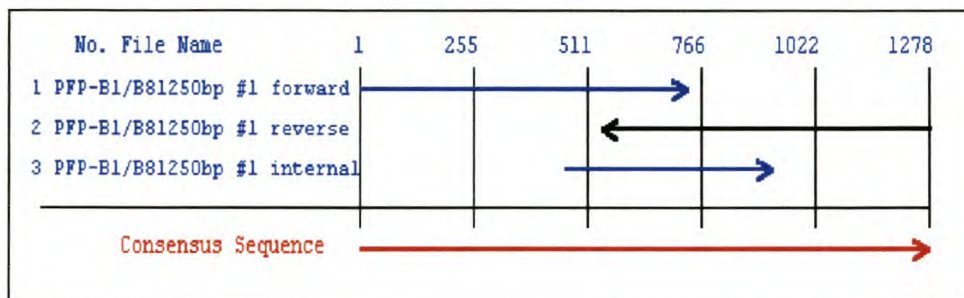
Figure 3.18 An example of the overlap of the forward, reverse and internal PFP-B1/B8$_{1250bp}$ sequences in Contig Manager (DNASIS) to generate a full-length sequence for a single PFP-B1/B8$_{1250bp}$ clone. Similar overlaps were conducted to obtain full-length sequences for each PFP-B1/B8 clone.

Initial sequence analyses were performed in DNASIS and sequence identity was determined using the BLASTn database (Altschul *et al.*, 1990). Results of the cloning and preliminary sequence analyses of the PFP-B1/B8$_{1100bp}$ clones are summarised in Table 3.5 and those of the PFP-B1/B8$_{1250bp}$ clones in Table 3.6.

All the PFP-B1/B8$_{1250bp}$ clones showed homology to the PFP-β gene of rice, castor bean, potato and citrus while the PFP-B1/B8$_{1100bp}$ clones were homologous to the PFP-β gene of rice and potato. The E-value, which measures the expected number of sequences in the database that would produce a given score by chance, was very low. This indicates that the fragments were significantly homologous to PFP-β (Brenner, 1998). Not all the individual PFP-B1/B8$_{1250bp}$ clones had the same BLASTn score. This was also true for the PFP-B1/B8$_{1100bp}$ clones which suggests that there may be sequence variation within the 1100bp and 1250bp amplicon populations. From the size range of the full length PFP-B1/B8$_{1100bp}$ and PFP-B1/B8$_{1250bp}$ fragments, presented in Table 3.5 and Table 3.6 respectively, it is evident that the PFP-B1/B8$_{1250bp}$ fragment is approximately 1277-1365bp and much larger compared to the size estimated by agarose gel electrophoresis (1250bp), while the size of the PFP-B1/B8$_{1100bp}$ fragment ranges from 1070-1084bp, smaller than that estimated by gel electrophoresis.

Table 3.5 Summary of cloned N19 and N21 PFP-B1/B8$_{1100bp}$ fragments obtained and preliminary sequencing results. The BLASTn search results are also indicated. Only those scores producing the most significant results are presented.

Variety	Clone	Forward sequence length size (bp)	Reverse Sequence length size (bp)	Full length sequence size (bp)	Organism	BLASTn results; PFP-β identity	
						Score (bits)	E-value
N19	1	450	530	1081	Potato	78	4e-12
	2	585	583	1073	Potato	80	2e-12
	3	669	688	1078	Rice	88	1e-14
	4	623	340	1079	Rice	88	8e-15
	5	659	704	1076	Rice	80	3e-12
	6	679	536	1070	Potato	88	1e-14
	7	675	690	1075	Rice	88	1e-14
	8	614	232	1070	Rice	80	3e-12
	9	661	672	1150	Potato	86	4e-14
	10	647	518	1081	Potato	80	3e-12
N21	1	552	516	1084	Rice	88	8e-15
	2	555	576	1075	Rice	88	1e-14
	3	669	688	1056	Rice	88	1e-14
	4	595	603	1075	Rice	88	1e-14
	5	723	395	1076	Potato	72	6e-10
	6	653	536	1076	Rice	88	1e-14
	7	646	626	1082	Rice	82	6e-13
	8	648	614	1071	Rice	88	1e-14
	9	638	627	1075	Rice	88	1e-14
	10	588	667	1083	Rice	88	1e-14

Table 3.6 Summary of cloned N19 and N21 PFP-B1/B8$_{1250bp}$ fragments obtained and preliminary sequencing results. The BLASTn search results are indicated. Only those scores producing the most significant results are presented.

Variety	Clone	Forward sequence length size (bp)	Reverse sequence length size (bp)	Internal sequence length size (bp)	Full length sequence size (bp)	Organism	BLASTn results; PFP-β identity	
							Score (bits)	E-value
N19	1	713	436	698	1278	Rice	105	5e-20
	2	635	734	721	1281	Rice	113	2e-22
	3	601	755	575	1277	Rice	92	8e-16
	4	393	543	487	1348	Rice	92	8e-16
	5	217	663	436	1376	Rice	92	8e-16
	6	199	643	200	1339	Rice	92	8e-16
	7	274	707	470	1346	Rice	92	8e-16
	8	378	255	514	1344	Rice	92	8e-16
	9	393	653	594	1355	Rice	92	8e-16
	10	536	254	402	1346	Rice	60	3e-06
N21	1	602	580	558	1343	Rice	92	8e-16
	2	623	310	438	1357	Rice	92	8e-16
	3	614	538	595	1346	Rice	92	8e-16
	4	674	581	606	1355	Rice	78	1e-11
	5	217	712	593	1365	Rice	72	8e-10
	6	199	707	656	1362	Rice	92	8e-16
	7	560	538	570	1339	Rice	78	1e-11
	8	567	712	607	1347	Rice	92	8e-16
	9	672	707	517	1355	Rice	92	8e-16
	10	660	668	611	1349	Rice	92	8e-16

The wide range of sequence sizes for the PFP-B1/B8$_{1100bp}$ and PFP-B1/B8$_{1250bp}$ fragments, presented in Tables 3.5 and 3.6, may be attributed to the use of single pass sequences,

which may contain some sequence errors. On the other hand, the discrepancy in the size of the PFP-B1/B8 fragments may be due to genuine sequence variation within the PFP-B1/B8 fragment populations.

3.3.5 Sequence variation within the PFP-B1/B8 fragment populations

It was thought that further variants of the PFP-β gene would exist within the PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ fragment populations. Variation within the fragment populations was investigated by performing sequence analysis of individual clones.

Comparative alignment of sections of PFP-B1/B8 clones

The PFP-B1/B8$_{1100bp}$ sequences were aligned using the multiple alignment function of DNASIS. Several individual sequences were identical in different sections along the length of the fragment while other sequences differed by a few bases. Similar results were obtained upon alignment of individual PFP-B1/B8$_{1250bp}$ sequences. The regions of the PFP-B1/B8 sequences, used as to demonstrate the comparative alignments (shown in Figure 3.20), are indicated in Figure 3.19. Figure 3.20 shows the similar sequences (represented by a single consensus sequence) and the variant sequences within both fragment populations.
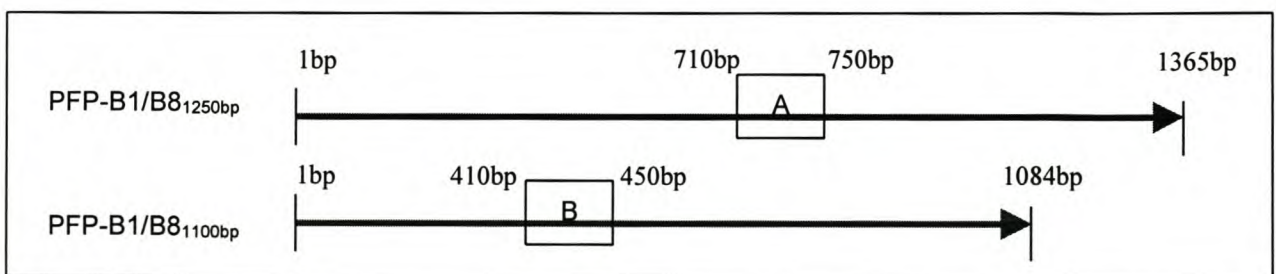
Figure 3.19 Diagram representing the PFP-B1/B8 fragments and the regions used to demonstrate the comparison of individual clones of each fragment. Region $\boxed{A}$ represents that used for the PFP-B1/B8$_{1250bp}$ fragment, while $\boxed{B}$ represents the region used for the PFP-B1/B8$_{1100bp}$ fragment. Since the individual sequences from the respective PFP-B1/B8 clones were of different sizes, approximate base numbers are indicated on the diagram.

```
A
                            710        720        730        740        750
    consensus    > G?CAATAACT TG?GTTGATA ?GTCTCCCAC ?TGAGGACG? CCA?CCC??C
    N19#5.DNA    > G?CAATAACT TG?GTTGATA ?GTCTCCCAC ?TGAGGA?GG CCA?CCC??C
    N19#6.DNA    > G?CAATAACT TG?GTTGATA AGTCTCCCAC ?TGAGGACG? CCA?CCC??C
    N19#8.DNA    > G?CAATAACT TG?GTTGATA ?GTCGCCCAC ?TGACGACGG CCA?CCC??C
    N21#1.DNA    > G?CAATAACT TG?GTTGATA ?GTCTCCCAC CTGAGGACG? CCA?CCC??C
    N21#4.DNA    > G?CAATAACT TG?GTTGATA ?GTCTCCCAC ?TGAGGACTG CCA?CCC??C
    N21#8.DNA    > GTCTCCCACT ?GAG??GAC? ?GCCTGACAC CTGAGGAGTG CCAGCCCATC
```

```
B
                            410        420        430        440        450
    consensus    > ATTT?A?GAG CTAGT?G?A? ??GA??GTTC CTTGAT?CCT ??CAGATT??
    N19#1.DNA    > CCT?CAAGAT TTGGTTGCAC CT?ACAG?A? ??TGTTT?CT G?C?GATTAC
    N19#4.DNA    > ACT?CA?GAG CTAGT?G?A? ??GA??GTAC CTTGATGCCC ??CAGATG??
    N21#1.DNA    > CCTTCA???T TTAGAAGCAA CTGACAGTAC AATGATAACT GGCAGATTA?
```

Figure 3.20 Section of an alignment produced using the Multiple Sequence function of DNASIS within the PFP-B1/B8$_{1250bp}$ fragment population (A) and the PFP-B1/B8$_{1100bp}$ population (B). The first sequence presented on each diagram represents the consensus sequence of the identical sequences within the section of the fragment. Bases within the variant sequences, which differ from the consensus sequence, are highlighted.

Minor sequence variation was evident between individual PFP-B1/B8$_{1250bp}$ sequences and between individual PFP-B1/B8$_{1100bp}$ sequences (Figure 3.20) with discrepancies of a few base pairs evident along the length of the fragments. Single base differences may be attributed to sequencing error which may be difficult to discern since single pass sequences of each clone were compared. *Taq* DNA polymerase which was used in PCR amplification, does not have 3'-5' exonuclease (proofreading) activity (Innis *et al.*, 1988; Holland *et al.*, 1991), another factor which may contribute to sequence errors (Tindall and Kunkel, 1988).

Similarity as shown by homology analysis

The multiple alignment (Higgins-Sharp) function of DNASIS is based on an assumed evolutionary relationship between the sequences. This is governed by the assumption that

more closely related sequences contain fewer mutations such as transitions, transversions and deletions. Thus, the large difference in size between the individual PFP-B1/B8 fragments (Tables 3.5 and 3.6) may be interpreted as deletions or substitutions and thus bias the alignment of the individual sequences. As an exercise, the aligned PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ individual clones were manually edited to remove obvious sequencing errors. These were identified as the presence of one or two base substitutions in one sequence compared to the absence of these bases in the majority of the sequences. For example, those bases that are indicated with a question mark in Figure 3.20 represent bases not present in the majority of the sequences and were thus considered to be sequencing errors. The bases were then shifted in relation to the sequences in order to obtain the best possible alignment. By performing this manual sequence editing of individual PFP-B1/B8$_{1100bp}$ clones, the lengths of the clones were reduced to the size range of approximately 1080bp. Similarly, sequence editing of the PFP-B1/B8$_{1250bp}$ clones resulted in sequence lengths of approximately 1300bp. The 20 edited, individual sequences of PFP-B1/B8$_{1100bp}$ clones derived from N19 and N21 were aligned with respect to each other using the multiple sequence function of DNASIS, and a homology tree was derived in order to determine the relationship between them. This was repeated for the 20 edited, individual sequences of PFP-B1/B8$_{1250bp}$ clones from N19 and N21. Figure 3.21 and Figure 3.22 represent the degree of homology between individual clones of the PFP-B1/B8$_{1100bp}$ and PFP-B1/B8$_{1250bp}$ populations respectively.

Sequence variation within the 1250bp and 1100bp fragment populations (Figure 3.21 and Figure 3.22) is not very significant with most sequences showing high homology to each other. The sequences from both commercial varieties (N19 and N21) appear to be quite similar to each other in that some of the PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ fragments from both sources are 100% homologous (Figures 3.21 and 3.22). Sequence variation within each of the PFP-B1/B8$_{1100bp}$ and PFP-B1/B8$_{1250bp}$ fragment populations was low and could have been due to sequencing errors and therefore was not considered sufficient to classify the different clones as separate variants.
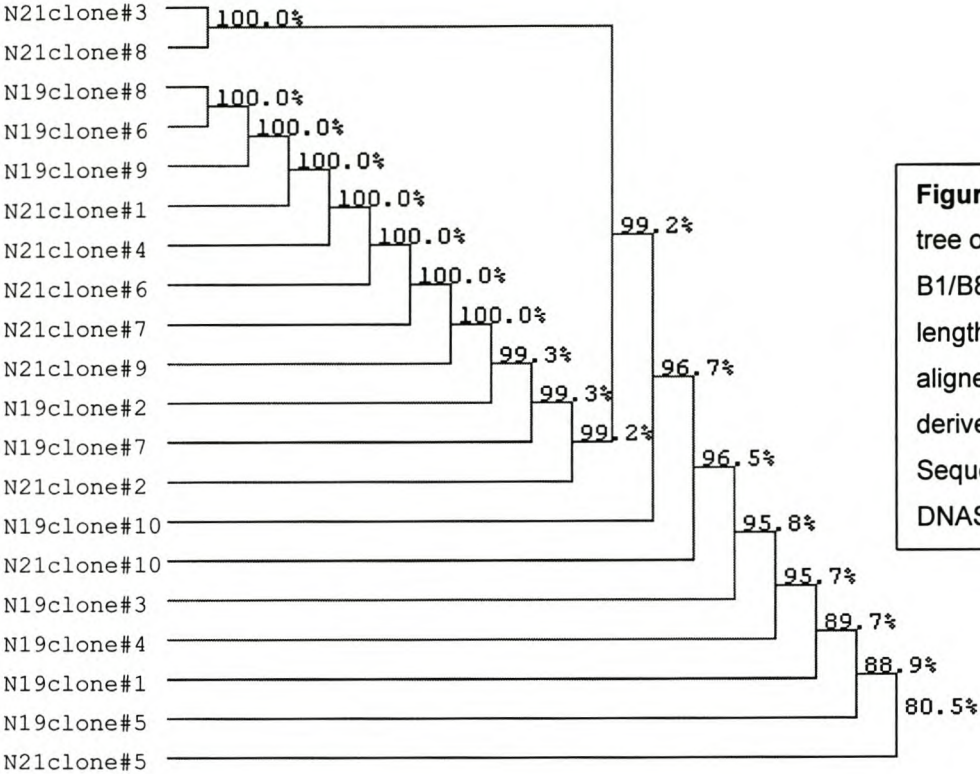
Figure 3.21 Homology tree of individual PFP-B1/B8$_{1100bp}$ clones. Full-length sequences were aligned and the tree derived in the Multiple Sequence function of DNASIS.
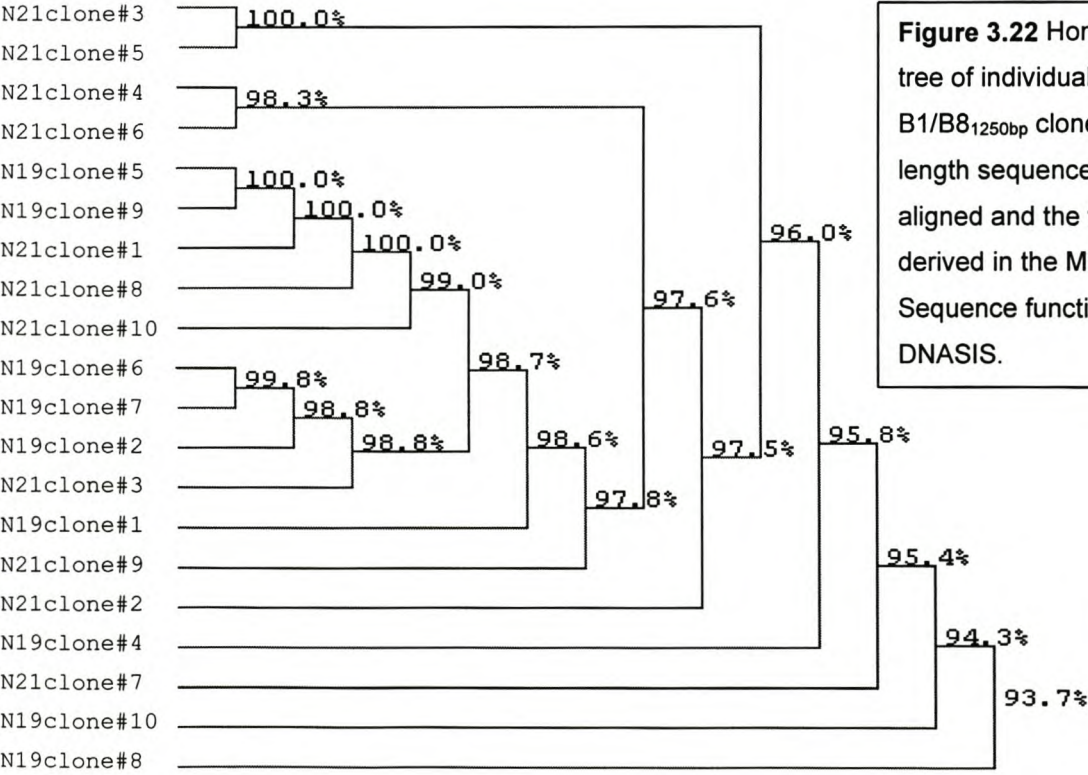
Figure 3.22 Homology tree of individual PFP-B1/B8$_{1250bp}$ clones. Full-length sequences were aligned and the tree derived in the Multiple Sequence function of DNASIS.

In order to minimise sequencing artifacts, a single PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ clone was sequenced three times and a full-length sequence generated using the Contig Manager function of DNASIS (as shown in Figure 3.18). The full-length sequences were aligned to obtain single consensus sequences for the selected PFP-B1/B8$_{1100bp}$ and PFP-B1/B8$_{1250bp}$ clones referred to as PFP-B1/B8$_{1100bp}$con and PFP-B1/B8$_{1250bp}$con respectively (primary sequence data not shown). The PFP-B1/B8$_{1250bp}$con sequence was then used as the first sequence in an alignment of all the individual, single-pass sequenced PFP-B1/B8$_{1250bp}$ clones using the Contig Manager function of DNASIS. This was repeated for the PFP-B1/B8$_{1100bp}$ sequences using the PFP-B1/B8$_{1100bp}$con sequence as the first sequence input into the Contig Manager alignment (alignment not represented). The resulting PFP-B1/B8$_{1100bp}$ and PFP-B1/B8$_{1250bp}$ consensus sequences, derived in this way, were used for all subsequent sequence analyses.

3.3.6 Sequence variation between the PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ fragments

The PFP gene fragments PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ were compared to each other to investigate the sequence differences between the variants. The sequences were compared using Maximum Matching (DNASIS). This comparative alignment is shown in Figure 3.23. Since the PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ sequences were different in size, it was possible that the variation between the sequences could be due to one large region of sequence variation between the sequences. Instead, the results shown in Figure 3.23 indicate that there are regions of low homology (areas of no matching bases) interspersed with regions of high homology (areas of matching bases) between the sequences. However, due to the different sizes of the fragments, the Maximum Matching function does not provide an accurate estimate of the homology between the sequences. In order to investigate the variation between the sequences further, the intron and exon regions of the respective sequences were compared. Initially, each sequence (the PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$) was aligned to the known castor bean sequence available on the database using Contig Manager (DNASIS). Figure 3.24 represents this alignment for the PFP-B1/B8$_{1250bp}$ sequence, while that of the PFP-B1/B8$_{1100bp}$ sequence is shown in Figure 3.25.

Figure 3.23 Alignment of the PFP-β variants PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ consensus sequences. The alignment was derived in Maximum Matching (DNASIS).
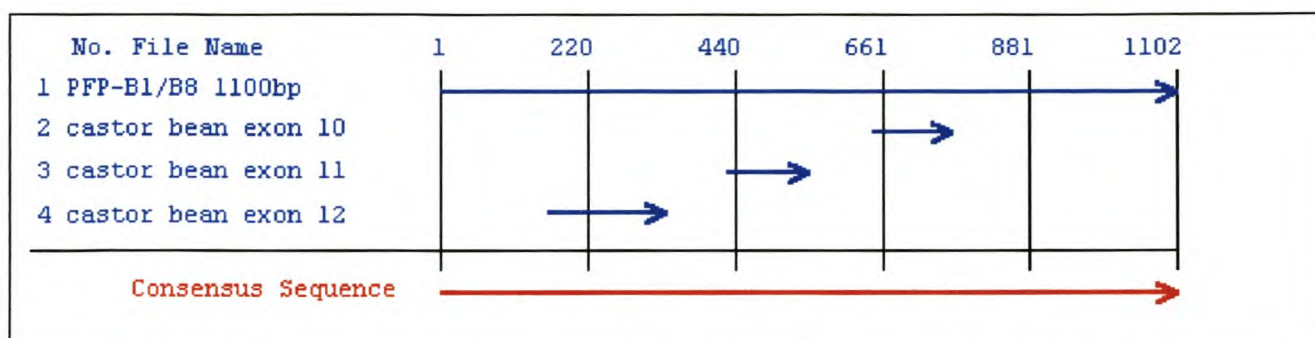
Figure 3.24 Diagram representing the alignment of the PFP-B1/B8$_{1100bp}$ consensus sequence with the known exon regions of castor bean using Contig Manager (DNASIS).
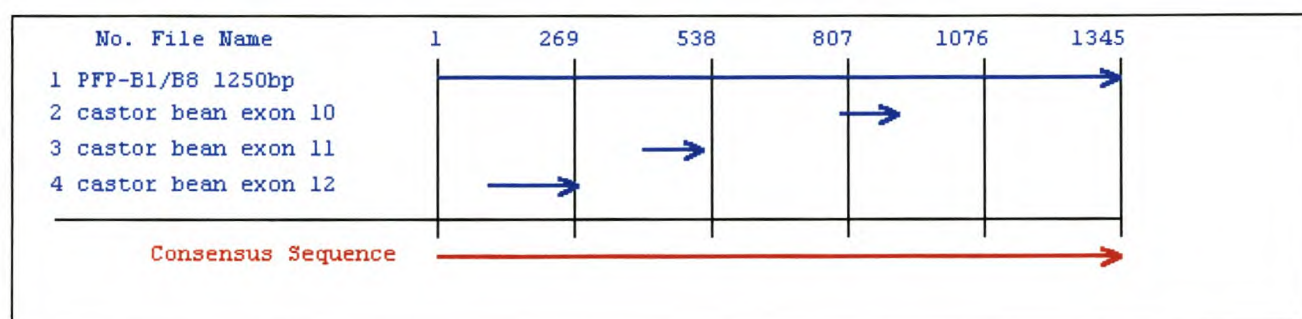
Figure 3.25 Diagram representing the alignment of the PFP-B1/B8$_{1250bp}$ consensus sequence with the known exon regions of castor bean using Contig Manager (DNASIS).

By aligning the castor bean exon regions with the PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ sequences, the conserved exon and intron regions of the fragments were identified. The corresponding exon and intron regions of the two fragments were then separately compared in the Maximum Matching function of DNASIS (preliminary data not shown). This function analyses the similarity of two sequences by using an algorithm that shifts sequences relative to each other and inserts gaps which are associated with a penalty. A score is then assigned to the most significant alignment and percent homology values are indicated for the alignment window. From these values, a diagram representing the homology between the sequences in their exon and intron regions was constructed. This diagram is indicated in Figure 3.26.
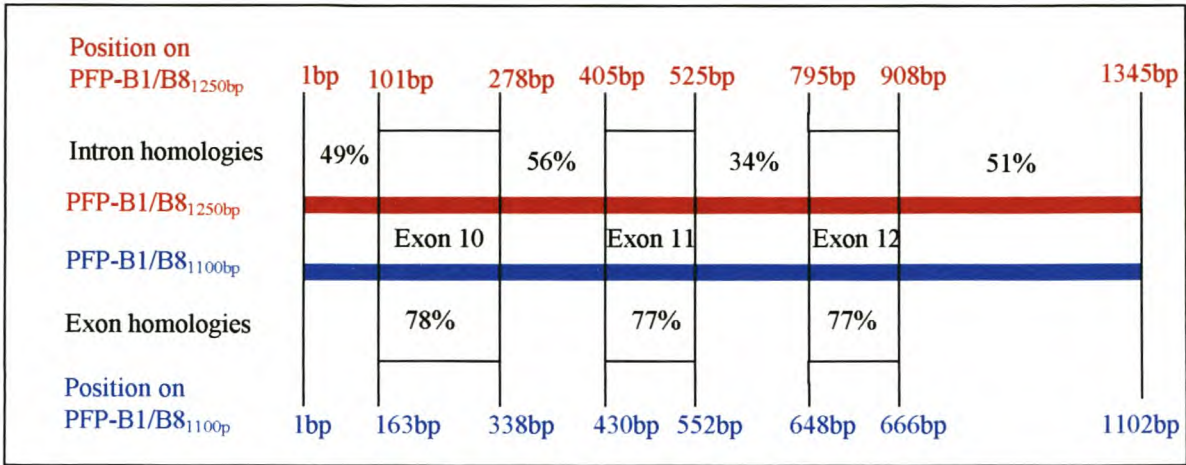
Figure 3.26 Diagram representing the homology between the PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ sequences in their respective putative intron and exon regions. The exon and intron regions were derived by comparison to the known castor bean exons. The percentage homology, computed using Maximum Matching (DNASIS) and intron and exon positions on the PFP-B1/B8 sequences are also indicated.

There appears to be high homology in the putative exon regions of the PFP-B1/B8$_{1100bp}$ and PFP-B1/B8$_{1250bp}$ fragments compared to the low homology in their respective putative intron regions (Figure 3.26). It was expected that higher sequence variation would be present in the intron region of these sequences since they are subject to less functional constraint than the coding or exon regions and therefore accumulate more mutational changes. These results confirm, at the nucleotide level, that there are two variants of the PFP-β gene in the commercial sugarcane cultivars N19 and N21, which are characterised by the PFP-B1/B8$_{1100bp}$ and PFP-B1/B8$_{1250bp}$ fragments.

3.3.7 Copy number of the sugarcane PFP-β gene

At the outset it had been hypothesised that PFP-β is a single copy gene in sugarcane, which suggests that the two PFP-β variants may represent alleles at a single locus. In order to investigate this, information regarding the organisation of the PFP-β gene in the sugarcane genome needed to be obtained. To this end, Southern hybridisation analysis was used. Initially restriction site analysis was performed on the PFP-B1/B8$_{1100bp}$ and PFP-

B1/B8$_{1250bp}$ variants using the DNASIS restriction site search function. Two enzymes i.e. *Hind* III, which restricts once within the PFP-B1/B8$_{1250bp}$ fragment and does not restrict within the PFP-B1/B8$_{1100bp}$ fragment, and *Eco* R$_I$, which restricts once within the PFP-B1/B8$_{1100bp}$ fragment and has no restriction site within the PFP-B1/B8$_{1250bp}$ fragment, were selected. Table 3.7 shows the selected restriction enzymes and their restriction sites within the PFP-B1/B8 variants.

Table 3.7 Selected restriction enzymes and their restriction sites within the PFP-B1/B8 sequences

Enzyme	Restriction site	Cutting position on the PFP-B1/B8$_{1250bp}$ sequence	Cutting position on the PFP-B1/B8$_{1100bp}$ sequence
Eco R$_I$	G!AATTC	-	841bp
Hind III	A!AGCTT	177bp	-

Fifteen micrograms of genomic DNA from Black Cheribon, Coimbatore, N21 and N19 was restricted using *Hind* III. Similarly, fifteen micrograms of DNA from the different varieties were restricted with *Eco* R$_I$. The restricted DNA was separated on a 0.8% (w/v) agarose gel and the DNA stained with ethidium bromide to ensure that the DNA was digested and that separation of the fragments had occurred. Figure 3.27 shows an agarose gel containing the restricted genomic DNA samples.

The *Eco* R$_I$ restriction produced fragments of higher molecular weights compared to the *Hind* III restriction which results in a much wider spread of fragments (Figure 3.27). The restricted DNA was transferred to a nylon membrane using upward capillary blotting and fixed as described in section 3.2.20. The membrane was then probed using the PFP-B1/B8$_{1100bp}$ fragment according to the protocol described in section 3.2.20. Figure 3.28 shows the results of this hybridisation. Once the membrane was stripped, it was re-probed using the PFP-B1/B8$_{1250bp}$ fragment as a probe. Figure 3.29 shows the results of the hybridisation with the PFP-B1/B8$_{1250bp}$ probe
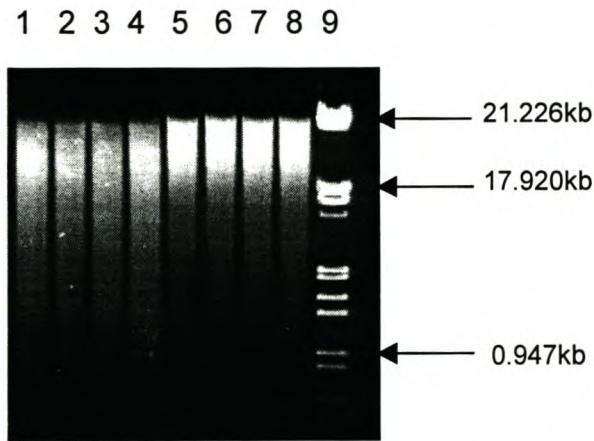
Figure 3.27 Results of electrophoretic separation of genomic DNA restricted with *Hind* III (lanes 1-4) and *Eco* RI (lanes 5-8) restriction enzymes. Lanes 1 and 5, Coimbatore; lanes 2 and 6, Black Cheribon; lanes 3 and 7, N21; lanes 4 and 8, N19 and lane 9, lambda molecular weight marker III.
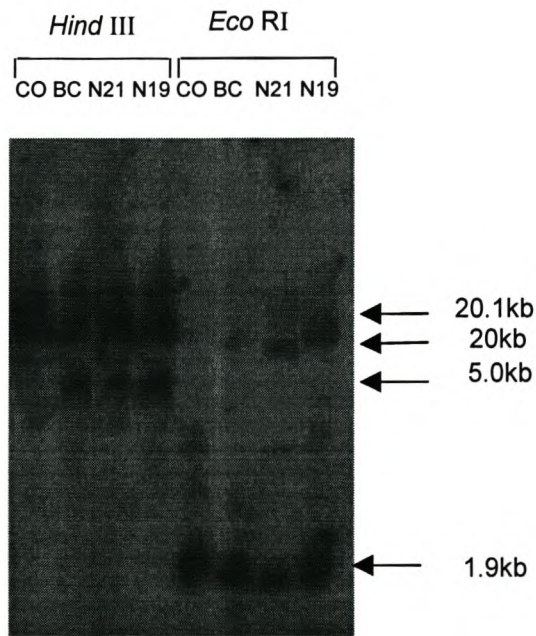
Figure 3.28 Results of Southern hybridisation analysis of the PFP-β gene using the PFP-B1/B8$_{1100bp}$ variant as a probe. Fifteen micrograms of genomic DNA was restricted with *Hind* III. Similarly, 15 μg of genomic DNA was restricted with *Eco* RI. Genomic DNA from Coimbatore (*Saccharum spontaneum*) (CO), Black Cheribon (*Saccharum officinarum*) (BC), N21 and N19 was used.

Results of the hybridisation indicated a more simple profile for the PFP-B1/B8$_{1100bp}$ fragment (Figure 3.28) compared to that produced by the PFP-B1/B8$_{1250bp}$ fragment (Figure 3.29). Hybridisation with the PFP-B1/B8$_{1100bp}$ fragment resulted in a similar banding profile for both the commercial varieties restricted with *Eco* R$_I$ and for those restricted with *Hind* III (Figure 3.28). Similar profiles were seen across the four varieties restricted with *Hind* III except for Coimbatore where a single band of approximately 20.1kb was obtained (Figure 3.28).
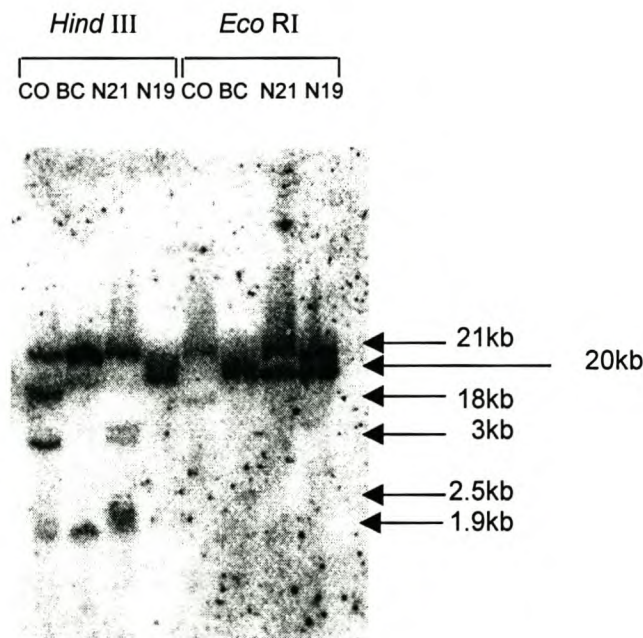
Figure 3.29 Results of Southern hybridisation analysis of the PFP-β gene using the PFP-B1/B8$_{1250bp}$ variant as a probe. Fifteen micrograms of genomic DNA was restricted with *Hind* III. Similarly, 15μg of genomic DNA was restricted with *Eco* RI. Genomic DNA from Coimbatore (*Saccharum spontaneum*) (CO), Black Cheribon (*Saccharum officinarum*) (BC), N21 and N19 was used.

Hybridisation with the PFP-B1/B8$_{1250bp}$ fragment (Figure 3.29) showed identical banding patterns for the two commercial varieties restricted with *Eco* R$_I$. In contrast, the banding

pattern observed for the commercial varieties, restricted with *Hind* III was distinctly different (Figure 3.29). The *Hind* III restriction profile obtained for variety N19 appeared to be similar to that obtained for Black Cheribon, while similar banding patterns were observed for N21 and Coimbatore (Figure 3.29). In addition, similar banding patterns were observed in the *Eco* RI restriction profile for variety N19 and Black Cheribon (Figure 3.29).

In order to make some interpretations on the Southern hybridisation pattern and speculate whether the pattern is suggestive of a single locus gene for PFP-β, the simplest genetic scenario that assumed all copies of the gene are the same, i.e. monomorphic, was taken into account. The actual results (derived from Figures 3.28 and 3.29) were compared to the predicted number of fragments that would be obtained if the gene were at a single locus per haploid genome. These results are summarised in Table 3.8.

Table 3.8 Summary of the predicted number of fragments that would be obtained if the PFP-β gene was a single locus gene and the actual number of fragments obtained by Southern hybridisation analysis.

	Predicted number of fragments for a single locus gene				Actual number of fragments obtained			
	PFP-B1/B8 1100bp fragment probe		PFP-B1/B8 1250bp fragment probe		PFP-B1/B8 1100bp fragment probe		PFP-B1/B8 1250bp fragment probe	
	Eco RI	*Hind* III	*Eco* RI	*Hind* III	*Eco* RI	*Hind* III	*Eco* RI	*Hind* III
N19	2	1	1	2	2	2	2	2
N21	2	1	1	2	2	2	2	3
Black Cheribon	2	1	1	2	2 (one faint)	2	2	2
Coimbatore	2	1	1	2	2 (one very faint)	1	2	4

The results presented in table 3.8 suggest that the number of bands obtained was not consistent with the assumed simplest genetic scenario.

3.4 Discussion

In this study it was hypothesised that the PFP-β gene was at a single locus in sugarcane and therefore an appropriate model for the study of allelic variation in this polyploid plant. Using previously designed specific primers which amplify across putative exons 10, 11 and 12 of the PFP-β gene, two amplicons of approximately 1250bp and 1100bp (designated PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ respectively) were obtained from genomic DNA of the commercial hybrid varieties N19 and N21. Initially, the conditions for PFP-B1/B8 PCR had to be optimised. After testing various PCR parameters, a combined hot start, touchdown protocol produced consistent amplification. Sloop and Rhodes (2000) recommend the use of hot start and touchdown PCR to amplify difficult targets. In their experiments, a hot start increased the specificity of the reactions. By combining the hot start technique and the touchdown strategy; specific amplification of the desired product was achieved. Harvey (1998) occasionally obtained a third PCR fragment after amplification with the PFP-B1/B8 primers on Coimbatore genomic DNA. The author speculated that this fragment was an artifact which arose during the PCR reaction, or genetic variation in the Coimbatore background, or a third variant which was present at a low dosage or with variation in the primer binding site for which amplification is difficult. Other spurious bands were also observed during amplification of the PFP-β region in this study (Figures 3.9 and 3.15). These bands were not consistently obtained and were subsequently shown to not be PFP-β.

Homology searches against international nucleotide databases indicated that both amplicon populations were highly homologous to PFP-β from other sources. Minor sequence variation was detected within the PFP-B1/B8$_{1100bp}$ and PFP-B1/B8$_{1250bp}$ fragment populations with individual clones showing single base differences. These differences may be attributed to single pass sequences being compared. The full-length sequences, which were generated from partial sequences in Contig Manager, may have poor homology in the overlap regions due to ambiguities at the beginning and end of electropherograms, which may be attributed to abnormal peak heights and widths. Sequence comparison within the PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ fragment populations suggested that only two

distinct variants of the PFP-β gene, represented by the PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ amplicons, were present.

The size difference between the two amplicons was due to diverse sequence variations dispersed throughout the length of the PFP-B1/B8$_{1100bp}$ and PFP-B1/B8$_{1250bp}$ sequences rather than a few large sequence disparities. More variation occurred in the putative intron regions compared to the exon regions. Some variation was observed in the form of mutations such as single and multiple base insertions or deletions. According to Flavell and Moore (1996), such mutations are likely to occur due to enzyme error during DNA replication. Mutations are more likely to be tolerated in introns which do not contain coding sequences, rather than the exon regions since sequence variation within the exon regions may contribute to changes in gene structure (Paterson, 1996).

The PFP-β and PFP-α genes are 40% identical at the amino acid level, 20% of which are conserved substitutions in the overlapping region (Carlisle *et al.*, 1990). It is surmised that the α and β subunit genes originated from a common ancestral gene. A duplication of the β subunit gene may have given rise to the α subunit and due to changes occurring in the catalytic site, regulatory interaction occurs once a heterotetramer has formed. A similar duplication event may have given rise to the two PFP-β variants described here. Since the alpha and beta subunit genes share regions of high homology, it was of interest to note that neither the PFP-B1/B8$_{1100bp}$ fragment nor the PFP-B1/B8$_{1250bp}$ genomic fragment showed BLASTn homology to the PFP-α subunit gene from any source. Instead, both genomic fragments showed high homology to PFP-β. This suggests that both the PFP-B1/B8 fragments represent variants of the same gene.

Allelic variants have been identified in several other plant enzymes. For example, Nowakowska (1998) describes the oxalate oxidase enzyme, which is encoded by two allelic variants in hexaploid wheat. Despite the 87% homology between the coding regions of the sequences from the two variants, the 5' flanking regions differ greatly from one another. The enzyme UDP-glucose pyrophosphorylase (UGPase) in potato is encoded by two allelic variants that differ by five amino acids (Sowokinos *et al.*, 1997). In barley, Erkillä

et al. (1998) describe three allelic variants of the β-amylase enzyme. The sequences of the three variants differed by the insertion/deletion of a palindromic 126bp sequence in intron III, the insertion/deletion of a 39bp sequence in intron III and two amino acid substitutions in the open reading frame. Similarly, as described by Erkillä *et al.* (1998), much of the variation observed between the PFP-B1/B8$_{1100bp}$ and PFP-B1/B8$_{1250bp}$ genomic sequences occurred in the intron regions.

Southern hybridisation was used to provide some evidence as to whether the PFP-β gene was a single locus gene in sugarcane. The results of the Southern hybridisation analysis on genomic DNA restricted with *Hind* III and probed with the PFP-B1/B8$_{1250bp}$ fragment indicates that the banding pattern observed for N19 was similar to that observed for Black Cheribon (*S. officinarum*) while the banding patterns for N21 and Coimbatore (*S. spontaneum*) were similar. It is interesting to note that N21 shows more *S. spontaneum*-like traits (tall stalks with thin leaves and erect canopy, low sucrose content, high fibre and hard stalks), while N19 has more *S. officinarum*-like characteristics (open growth habit with broad slightly drooping leaves, thicker stalks and higher sucrose content). This suggests that N21 has a higher proportion of the *S. spontaneum* genome and N19 more of the *S. officinarum* genome (M. Butterfield, personal communication[1]).

The Southern hybridisation experiments performed in the current study indicate that there are DNA polymorphisms for the genomic regions corresponding to each fragment. This is more consistent with a complex polyploid like sugarcane than the assumed simple genetic scenario that PFP-β is monomorphic. However, since the Southern hybridisation pattern obtained was relatively simple, this suggests that PFP-β is less polymorphic than might be expected for sugarcane. Blakeley *et al.* (1992) observed two fragments for PFP-β in a Southern hybridisation experiment using castor bean genomic DNA digested with *Hind* III and probed with a castor bean PFP-β cDNA fragment. Blakeley *et al.* (1992) also obtained a single fragment when castor bean genomic DNA was restricted with *Eco* R$_I$ and probed with a castor bean PFP-β cDNA fragment. These results suggested that there was a single gene for the PFP-β subunit.

---

[1] Mike Butterfield, SASEX, Mount Edgecombe, Durban

However, in the polyploid potato, two fragments respectively were observed in an Eco R$_I$ and $Hind$ III digest of potato genomic DNA probed with PFP-β cDNA. Blakeley et $al.$ (1992) surmised that this banding pattern resulted due to the polyploid nature of the plant and suggested that the hybridisation was indicative of a single gene for the PFP-β gene in potato. In sugarcane too, the presence of extra bands as opposed to single bands as expected for a single locus gene (Table 3.9) may be attributed to the polyploid nature of the plant.

Groenewald (1998) obtained two bands when genomic DNA from Black Cheribon, Coimbatore, N21, Nco310 (another commercial sugarcane variety), potato, castor bean and Zea $mays$ was restricted with Eco R$_I$ and probed using a PFP-β cDNA fragment. These results are consistent with the two fragments observed for the two commercial varieties N19 and N21, and the ancestral species representatives Black Cheribon and Coimbatore probed with the PFP-B1/B8$_{1250bp}$ fragment. However, the size of the fragments obtained by Groenewald (1998) were approximately 4.16kb and 2.3kb for the Black Cheribon, Coimbatore, N19 and N21 commercial varieties while the size of the fragments obtained in a similar experiment in this study were different. Hybridisation of the PFP-B1/B8$_{1250bp}$ fragment probe to the Eco R$_I$ restricted Coimbatore, Black Cheribon, N21 and N19 genomic DNA produced fragments of approximately 21kb and 20kb while hybridisation of the PFP-B1/B8$_{1100bp}$ fragment probe to the same restricted samples showed a single fragment of approximately 1.9kb and an additional band of approximately 20kb for the two commercial varieties.

There is a possibility that the PFP-B1/B8$_{1100bp}$ fragment probe hybridised to the 1250bp genomic fragment in the genomic DNA and in the same vein, the PFP-B1/B8$_{1250bp}$ fragment probe may have hybridised to the 1100bp fragment on the membrane since the two fragments are approximately 70% homologous to each other (as shown in the putative exon regions in Figure 3.26 and in the comparative alignment of the two fragments in Figure 3.23). This may account for the presence of extra bands in the Southern hybridisation analysis of the PFP-β gene compared to the predicted number of bands for a single locus gene for sugarcane.

The number of bands obtained in the Southern experiments is dependent on the degree of polymorphism within the PFP-β gene in sugarcane. Thus, in a complex, interspecific polyploid like sugarcane, it could be speculated that the banding pattern obtained by Southern hybridisation might indicate different alleles at one locus, which would be consistent with a single copy gene per haploid genome. Further evidence is required to indicate whether PFP-β is a single locus gene in sugarcane. An alternative technique to determine whether the variants are alleles is *in situ* hybridisation. The technique, which has been applied to sugarcane to characterise the double genome structure of modern sugarcane cultivars (D' Hont *et al.*, 1996), involves the hybridisation of labelled genomic DNA to chromosome spreads. The technique may also be applied to PCR amplified genomic fragments. By labelling the PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ genomic fragments and hybridising them to chromosome spreads of a commercial sugarcane variety (either N19 or N21), some insight as to whether the two fragments represent separate loci could then be obtained.

In this study it has been shown that two variants of the PFP-β gene exist in the commercial varieties N19 and N21. Previously, Harvey *et al.* (1999) also showed the presence of these two variants in the ancestral species representatives Coimbatore (*Saccharum spontaneum*) and Black Cheribon (*Saccharum spontaneum*). It is unlikely that the two variants maintained the same level of heterozygosity over thousands of years. Some variation in the ancestral germplasm would be expected. This suggests that the variants characterised by the PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ fragments are perhaps separate loci in sugarcane, which may have arisen as a result of duplication. It is thought that the two gene variants may not segregate since they are both present in the commercial and ancestral sugarcane germplasm (Harvey *et al.*, 1999). It may then be speculated that the gene forms are closely linked loci at one chromosome or on different chromosomes. The simple genomic pattern for PFP-β obtained by Southern hybridisation analysis is consistent with the suggestion that the two variants may be closely linked on the same chromosome.

CHAPTER 4

PFP-β EXPRESSION IN SUGARCANE

4.1 Introduction

In chapter 3, it was shown that two PCR fragments of approximately 1100bp and 1250bp were obtained from genomic DNA, isolated from the commercial cultivars N19 and N21, amplified with specific primers PFP-B1 and PFP-B8. Sequence analysis indicated that both fragments (designated PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ respectively) were identifiable as PFP-β based on nucleotide database homology searches. Despite minor sequence variation within the 1250bp and 1100bp fragment populations, the PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ fragments were poorly homologous to each other at the nucleotide level with higher levels of homology residing in the exon compared to the intron regions of the sequences. Thus, there are two distinct variants of the PFP-β gene in sugarcane. However, it was unknown whether the two genomic variants encoded an active PFP-β gene. Therefore, this study endeavored to determine expression of the variants in sugarcane.

Moore (1995) suggested that in sugarcane, sucrose accumulation is primarily regulated within the translocation system and at the level of the sink. Botha *et al*. (1996) noted that in sugarcane, sucrose increased from the young to the older internodes. PFP activity has also been shown to increase from internode 3 reaching optimal activity at internode 6 and 7 after which the enzyme declined to internode 10 (Whittaker and Botha, 1997). An inverse correlation between peak PFP activity/ expression levels and the prevailing sucrose content of internodal tissues was reported (Whittaker and Botha, 1997). In a study by Whittaker and Botha (1999), PFP activity appeared to be controlled by the expression of the β subunit of the protein.

It is possible that expression of the PFP-β variants is different at different stages of sucrose accumulation. Several techniques to analyse mRNA expression are available, three of

which, the RT-PCR technique, cDNA macroarray screening and Northern hybridisation, are described below.

RT-PCR (Reverse transcription-polymerase chain reaction) is the technique of choice for analysing extremely low abundance mRNA derived from cells or tissues (Soauze *et al.*, 1996). In RT-PCR, the target mRNA molecules are reverse transcribed to cDNA which is in turn amplified by PCR (Sambrook *et al.*, 1989). Northern hybridisation permits the analysis of mRNA and is commonly used as a quantitative assay (Omann *et al.*, 1993; Kim *et al.*, 1999; Thoiron and Briat, 1999; Swiderski *et al.*, 2000). This technique determines the size and abundance of specific mRNA molecules in preparations of total or poly $(A)^+$ RNA. Northern analysis is not as sensitive as the RT-PCR technique (Ra and Saris, 1995). cDNA macroarrays are a type of reverse Northern where cDNA clones are deposited onto a nylon filter and clones are probed with radiolabeled total cDNA. This technique allows for many clones to be screened at once. In recent years, a parallel technology, which involves miniaturisation and automation of the macroarray technology, has emerged i.e. the microarray. In microarray technology, DNA is deposited onto a glass slide and clones are isolated using fluorescent probes (vanHal *et al.*, 2000; Xiang and Chen, 2000).

Since it has been shown that the 1250bp and 1100bp fragments were obtained from the commercial cultivars N19 and N21 and that the fragments from both sources were highly homologous to each other, N19 was selected as the hybrid commercial variety for further expression analysis. The *Saccharum* species representatives, Black Cheribon and Coimbatore, were also selected for this study since it has been shown previously that the PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ fragments were also present in these varieties (Harvey *et al.*, 1999).

The specific objective of this study was to determine the expression of the PFP-β variants in tissues at different stages of sucrose accumulation of a commercial hybrid sugarcane variety (N19), and to investigate which variant is being expressed in the ancestral species representatives Black Cheribon and Coimbatore. Due to its sensitivity, the RT-PCR technique was selected to determine which variant is being expressed. Since the cDNA

macroarray allowed for a large number of clones to be screened at once, it potentially provided a quick method to isolate further expressed PFP-β gene sequences. In order to determine quantitative expression of the variants at different stages of sucrose accumulation, Northern hybridisation analysis was used.

4.2 Materials and Methods

4.2.1 Sugarcane varieties

The commercial sugarcane variety N19 and the two ancestral varieties, Coimbatore and Black Cheribon, collected as described in section 3.2.1, were used in PFP expression studies. Leaf, leafroll (meristematic region), bud scale, internode 4 (immature culm) and internode 7 (mature culm) material was acquired from 7- 12 month old plants of variety N19. Internode 1 was defined as the internode attached to the leaf with the uppermost visible dewlap. Leafroll tissue was also obtained from Black Cheribon and Coimbatore. Germinating sett tissue of variety N19 was also used in this study. Setts of N19 stalk (10-15cm) were cut and planted in potting soil (Grovida, SA) and maintained in the glasshouse for a week. Once germinated, the nodal parts of the setts were used for RNA extractions. The tissue was either used for RNA extractions immediately after acquisition, or prepared for RNA extractions by slicing the tissue, pulverising with a mallet and grinding to a fine powder under liquid nitrogen with a mortar and pestle. The tissue was then stored at -80°C until required.

4.2.2 Sugarcane cDNA libraries

Sugarcane cDNA libraries used for this study were previously prepared in the Biotechnology department at SASEX. Total cDNA libraries from meristematic tissue and mature culm were prepared from a commercial sugarcane variety (NCo376). Immature and mature culm subtractive libraries, prepared by reciprocal subtraction between internode 2 and internode 7, were derived from clone 126 from the South African sugarcane population AA40.

4.2.3 Sugarcane cDNA macroarray

A cDNA macroarray was prepared at the Clemson University Genomics Institute (USA) using a Qbot (Genetix). The randomly selected, bacterial clones (*E. coli* SOLR strain) were deposited onto a Hybond N+ nylon membrane (Amersham). The array contained 1400 clones in total; 500 of which were from the leafroll (total cDNA library), 500 from the mature culm (total cDNA library), 200 from the immature culm (subtractive cDNA library) and 200 from the mature culm (subtractive cDNA library).

4.2.4 Oligonucleotide primers

The PFP-B1 and PFP-B8 primers, as well as the vector primers SP6 and T7, were also used in this study (section 3.2.3). Primers PFP-B2 and PFP-B4 were previously designed from consensus sequences of potato and castor bean PFP-β gene sequences (Groenewald, 1998) and synthesised by Roche diagnostics (Germany). The primer sequences and target sites are summarised in Table 4.1.

Table 4.1 Summary of previously designed oligonucleotide primers and their sequence

Primer	Sequence 5'-3'	Size (bp)	Target site	Expected amplicon size
PFP-B4 Forward	AAT-TAA-CCC-TCA-CTA-AAG-GG	20	Sugarcane PFP-β gene (exon 8)	Approx. 250bp
PFP-B2 Reverse	TCA-TCI-ACA-ACA-TCA-TGI-GCC	21	Sugarcane PFP-β gene (exon 9)	

4.2.5 Precautions to prevent RNase contamination

All solutions used for procedures involving RNA were prepared with diethyl pyrocarbonate (DEPC) – treated water which was prepared as follows: 1ml of DEPC was added to 1 litre of ultra-pure water and allowed to dissolve overnight at 37°C. The solution was autoclaved for 20min to remove all traces of DEPC. Glassware, plasticware, spatulas, mortar and pestles were autoclaved and wiped with RNase Away (Molecular BioProducts, Whitehead Scientific) prior to use. Filter tips were used to prevent aerosol contamination from pipettes. Surfaces and electrophoresis apparatus were cleaned with RNase Away and double gloves were used to prevent RNase contamination from hands.

4.2.6 RNA isolation protocols

Small scale RNA extractions were performed with the SV Total RNA Isolation System (Promega) or the RNeasy® Plant Mini Kit (Qiagen).

SV Total RNA Isolation System (Promega)

RNA from N19 leaf and leafroll tissue was extracted using this method.

Using liquid nitrogen and a mortar and pestle, 100mg of tissue was ground to a fine powder which was transferred to a 1.5 ml Eppendorf containing 580μl lysis buffer (containing 2% (v/v) β-mercaptoethanol). Dilution buffer (350μl) was added to the tube and the sample was incubated for 3min at 70°C. This was followed by centrifugation for 10min at 12 000rpm in a microcentrifuge. The cleared lysate was transferred to a fresh tube containing 200μl 95% ethanol and the sample mixed by pipetting 3-4 times before being transferred to the spin basket assembly. After centrifugation at 12 000rpm for one minute, the eluate was discarded and 50μl of a freshly prepared DNase incubation mix (40μl yellow core buffer, 5μl 0.09M $MnCl_2$ and 5μl DNase I) added directly to the spin basket membrane. The tube was incubated at 37°C for 30min and 200μl DNase stop solution added to the spin basket which was then centrifuged at 12 000rpm for 1min in a microcentrifuge. The membrane was subjected to two washes using 600μl and 250μl wash solution respectively, each of which was followed by centrifugation at 12 000rpm for 1min. RNA was eluted by the

addition of 100μl nuclease-free water to the membrane and centrifugation at 12 000rpm for 1min.

RNeasy® Plant Mini kit (Qiagen)

This method was used to extract RNA from leaf, leafroll, internode 4, internode 7, bud scale and germinating sett tissue of variety N19. In addition this technique was also used to isolate RNA from leafroll tissue of Black Cheribon and Coimbatore.

The sample (200mg) was ground to a fine powder under liquid nitrogen using a mortar and pestle. The ground tissue was transferred to 450μl buffer RLT, vortexed and subsequently incubated at 56°C for 3min to help disrupt the tissue. The lysate was transferred to a QIAshredder spin column and centrifuged for 2min at 12 000rpm in a microcentrifuge. The cleared lysate was transferred to a new tube and 0.5x volume 96% ethanol added and mixed by pipetting. The sample was added to an RNeasy mini spin column and centrifuged at 10 000rpm for 15sec in a microcentrifuge. Buffer RW1 (700μl) was added to the column and centrifuged for 15sec at 10 000rpm. The RNeasy column was transferred to a new 2ml collection tube, 500μl buffer RPE added and the sample centrifuged for 15sec at 10 000rpm. Once the flow-through was discarded, a second wash using 500μl RPE followed by centrifugation at 12 000rpm for 2min was performed in order to dry the RNeasy membrane. The RNeasy column was transferred to a 1.5ml collection tube and 30μl of RNase-free water was pipetted directly onto the membrane. The sample was centrifuged for 1min at 10 000rpm to elute the RNA.

Large-scale RNA extraction protocol

The following protocol was used to obtain RNA from leaf, leafroll, internode 4 and internode 7 tissue of variety N19.

A RNA extraction buffer, containing 1% (w/v) SDS, 1mM aurintricarboxylic acid (ATA) (Sigma), 4% (w/v) p-aminosalicyclic acid (PAS) (Sigma), 10mM Tris, 1mM EDTA (pH 7.5), and 2% (v/v) β-mercaptoethanol was prepared and stored at 4°C. Four to five grams of tissue was ground under liquid nitrogen using a mortar and pestle. The tissue was

transferred to a pre-chilled 50ml Corning™ tube on ice containing 4ml of extraction buffer and 4ml of phenol: chloroform: isoamylalcohol [25:24:1 (v/v)]. The tissue was homogenised with the Ultra-Turrax (T25) homogeniser for 3-4min. The sample was centrifuged for 15-20min at 4 149xg and the top aqueous layer was transferred to a fresh 50ml Corning™ tube containing 20μl of 100mM ATA and 660μl of 12M LiCl. RNA precipitation was allowed to proceed overnight at 4°C. Samples were centrifuged for 10-15min at 4 149xg, the supernatant discarded and the pellet re-suspended in 1ml of 50mM ATA. The sample was transferred to a 1.5ml Eppendorf and centrifuged at 8 000rpm in the microcentrifuge in order to remove particulate matter. The supernatant was incubated with 170μl 12M LiCl in a fresh tube at 4°C overnight. After centrifugation at 11 000rpm for 10min, the supernatant was discarded and the pellet rinsed with 70% ethanol (chilled to –20°C). The pellet was re-suspended in 250μl 50μM ATA. Particulate matter was removed by centrifugation at 8 000rpm for 2min, the supernatant retained and 125μl of 7.5M NH_4 acetate and 750μl 95% ethanol were added to the sample, which was then incubated for 1hr at –20°C. The RNA pellet, recovered by centrifugation at 11 000rpm for 30min in an Eppendorf 5415C microcentrifuge at 4°C, was re-suspended in 20-200μl 50μM ATA (depending on the size of the pellet). All RNA samples were stored at –80°C.

4.2.7 Quantification and quality assessment of RNA

An aliquot of the RNA sample was diluted 1:100 or 1:200 in DEPC-treated water and the absorbance at 260nm determined spectrophotometerically. Since an absorbance of 1 unit at 260nm corresponds to 40μg of RNA per ml (A_{260} = 1 = 40ng/μl), the concentration of the RNA sample was calculated from the A_{260} value.

The purity of the RNA sample was assessed from the A_{260}/A_{280} ratio determined spectrophotometrically.

4.2.8 Assessing genomic DNA contamination by PCR

RNA samples were used as template in a random amplified polymorphic DNA (RAPD) PCR using a random primer. One reaction contained 1x *Taq* DNA polymerase magnesium free buffer (Promega), 1.5mM $MgCl_2$, 1mM dNTPs, 0.2μM random primer OPD-08 (Operon

technologies) and 1 unit of *Taq* DNA polymerase (Promega). Cycling conditions were one cycle of 94°C for 3min, 35°C for 1min, 72°C for 2min, 40 cycles of 94°C for 1min, 35°C for 1min, 72°C for 2min followed by one cycle of 72°C for 7min.

4.2.9 Reverse transcription (RT) and PCR

One microgram of RNA and 0.2μl 50μM oligo (dT)$_{16}$ primer (Perkin Elmer), in a total volume of 10μl, was boiled for 5min and snap cooled on ice. The following reagents were added to the tube to a total volume of 20μl: 1x Expand™ Reverse Transcriptase Buffer (Roche Molecular Bioproducts), 10mM dithiothrietol (DTT), 1mM dNTPs, 40 units RNase inhibitor (Roche Molecular Bioproducts), and 50 units Expand™ Reverse Transcriptase (Roche Molecular Bioproducts). The reaction was incubated at 50°C for 1hr and the enzyme subsequently heat inactivated by boiling for 5min and chilling on ice.

For PCR, the reaction was diluted 1:1 with sterile water. One microlitre, 2μl, 4μl and 8μl aliquots of the diluted product was used as template in PFP-B1/B8 PCR. One 30μl PCR reaction contained, in addition to template, 1x GeneAmp® PCR buffer II, 1.5mM MgCl$_2$, 0.1mM dNTPs, 0.2μM each primer, 2.5 unit Ampli*Taq* Gold™ (Perkin Elmer). The reaction was overlaid with 50μl mineral oil (Sigma) and cycled in a Hybaid OmniGene thermocycler. Cycling included a single heat activation cycle of 95°C for 12min, followed by a touchdown protocol of 1 cycle of 95°C for 1min, 50°C for 1min, 72°C for 2min; 1 cycle of 95°C for 1min, 49°C for 1min, 72°C for 2min; 1 cycle of 95°C for 1min, 48°C for 1min, 72°C for 2min; 1 cycle of 95°C for 1min, 47°C for 1min, 72°C for 2min; 1 cycle of 95°C for 1min, 46°C for 1min, 72°C for 2min; 35 cycles of 95°C for 1min, 45°C for 1min, 72°C for 2min, and two holds of 72°C for 5min and 35°C for 5min.

4.2.10 Gel electrophoresis and visualisation of cDNA

PCR products were electrophoresed and visualised as described in section 3.2.11.

4.2.11 Purification of PCR products

The electrophoresed PCR products were purified using the Qiaquick Gel Extraction Kit (Qiagen). The protocol, described in section 3.2.12, was followed.

4.2.12 Concentration of purified RT-PCR products

Purified RT-PCR fragments were concentrated in the SpeedVac for 30min or until the volume had reduced to approximately 5µl.

4.2.13 Quantification of purified cDNA fragments

The RT-PCR fragments were quantified using fluorometry as described in section 3.2.14.

4.2.14 Cloning cDNA fragments in *E. coli*

The pGEM®-T Easy Vector System (Promega) was used for cloning the cDNA fragments. The cloning protocol was followed according to section 3.2.16.

4.2.15 Plasmid isolation

Plasmid was extracted from overnight cultures of selected positive colonies using the Nucleobond AX20 Plasmid Isolation Kit (Machery-Nagel) or the rapid extraction protocol (as described in section 3.2.17).

4.2.16 cDNA sequence determination

The sequences of the cDNA fragments were determined as described in section 3.2.18.

4.2.17 Sequence analysis

Sequences were analysed using Sequence Navigator (v 1.0.1) and DNASIS (v 2.6). Sequence identity searches were performed using the BLASTN program (Altschul *et al.*, 1990).

4.2.18 cDNA macroarray screening

Probe generation and purification

Plasmids containing the RT-PCR fragment were amplified using the PFP-B1/B8 primers as described in section 3.2.10 with the exception that an equivalent of 20ng of plasmid DNA was used as template.

Probe labeling

PFP-B1/B8 cDNA fragments were labeled with the Megaprime Labeling kit (Amersham). Five microlitres of random nonamer primer solution (Amersham) and 25ng of DNA, in a total volume of 10μl was boiled for 5min and snap cooled on ice. Five microlitres of reaction buffer, 4μl each of supplied solutions of dTTP, dCTP, dGTP, 5μl of $\alpha^{33}P$ dATP (specific activity 3000Ci/mmol, Amersham) and 2 units Klenow enzyme (DNA polymerase I Klenow fragment (cloned) in 100mM potassium phosphate pH 6.5, 10mM 2-mercaptoethanol and 50% (v/v) glycerol) was added at room temperature. The reaction was made up to a total volume of 50μl and incubated at 37°C for 1hr. Twenty microlitres of 50mM EDTA was added to stop the reaction.

Probe purification

Probes were purified using the NucTrap® Probe Purification Column as described in section 3.2.20.

Hybridisation

A 200μl volume of 10mg/ml denatured, fragmented salmon sperm DNA (prepared as described in section 3.2.20) was added to 20ml of hybridisation solution (0.5M sodium phosphate pH 7.2, 7% (w/v) SDS, 1mM EDTA) and the membrane incubated in this solution at 65°C overnight. The labeled probe was denatured by boiling for 5min, snap cooled on ice and added to fresh hybridistion solution in which the membrane was incubated for a further 24hr at 65°C.

The membrane was washed twice with 1x SSC, 0.1% (w/v) SDS for 20min each, followed by two washes with 0.5x SSC, 0.1% (w/v) SDS for 20min each. All washes were performed at 65°C.

Phosphorimaging

Membranes were checked for signs of radioactivity using the Geiger-Muller counter, sealed in plastic and exposed to a phosphor screen for up to 3 days.

Stripping cDNA macroarrays for reprobing

The macroarray, which was kept damp before stripping, was stripped by two incubations in a solution of 0.1M NaOH, 10mM EDTA, 0.1% (w/v) SDS, for 10min with agitation on the Belly Dancer® (Sorvall), followed by a rinse in polished H_2O and two incubations in 5x SSPE (0.9M NaCl; 0.05M NaHPO4, 0.05M EDTA) of 10min each. The membrane was allowed to dry before being sealed in plastic and stored at 4°C. Membranes were stripped and reprobed three times before being discarded.

4.2.19 Northern hybridisation analysis

Precipitation of RNA for Northern blotting

Four to six RNA samples from the same tissue, extracted as described in section 4.2.6 (*large scale RNA extraction method*), were pooled and precipitated in order to concentrate the RNA. A 0.5x volume of cold 7.5M NH_4 acetate and 3x volume of 96% ethanol (chilled at -20°C) were added to the sample which was then chilled at -80°C for 10min. The sample was placed at -20°C for 2hr before being centrifuged at 11 000rpm at 4°C for 30min in an Eppendorf 5415C microcentrifuge. The RNA pellet was dried briefly under a partial vacuum in the SpeedVac, re-suspended in 10-30μl of DEPC-H_2O, re-quantified as described in section 4.2.7 and stored at -80°C.

Fractionation of RNA by agarose-formaldehyde gel electrophoresis

A 1.2% (w/v) agarose gel was prepared in the following manner: agarose was dissolved in 1x MOPS buffer [0.2M 3-(N-morpholino)-propanesulfonic acid pH 7.0, 0.5M sodium acetate, 0.01M EDTA]. Formaldehyde [0.17% (w/v)] was added to the gel and mixed

gently. Ten micrograms of RNA, 0.5x (v/v) formamide, 1x MOPS buffer and 0.1x (v/v) formaldehyde in a total volume of 25μl was incubated at 65°C for 5min and chilled on ice. Two and a half microlitres of loading dye [0.4% (w/v) bromophenol blue, 0.04% (w/v) xylene cyanol and 25% (w/v) Ficoll$^{™}$] was added to the RNA. RNA Molecular Weight Marker 0.3 – 6.9kb (Roche), treated in the same manner as the RNA, was also loaded alongside the RNA samples on the gel. Samples were electrophoresed for 3 hours at 60 volts or until the bromophenol dye front had migrated 7–8cm from the wells. The part of the gel containing the marker and a representative RNA sample was stained in 5μg/ml ethidium bromide for 30min, destained in DEPC water for 1-2hr and photographed under UV light alongside a fluorescent ruler. The remaining gel was used for Northern blotting.

Downward capillary blotting

Stacks for downward capillary blotting were set up as described in Figure 4.1 and allowed to proceed overnight. The stack was dismantled and the gel and membrane excised on the top left hand corner in order to orientate the membrane, which was subsequently dried for 1hr, UV crosslinked at 120 000 μJ/cm^2, sealed in plastic and stored at 4°C. The gel was stained (as described in section 4.2.10), to ensure that all the RNA had been transferred.
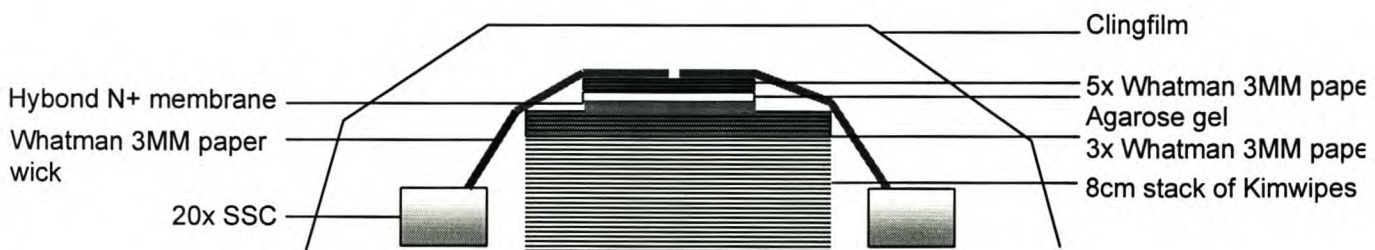
Figure 4.1 Diagram representing the downward capillary apparatus for Northern blotting. Sheets of Whatman 3MM filter paper were prewet in 20x SSC. Air bubbles were removed by gently rolling a sterile plastic 10ml pipette over the surface of the gel. The exposed surface of the filter paper was covered with cling wrap. Two Whatman 3MM paper wicks were placed on top of the stack and placed into containers half-filled with 20x SSC. The entire stack was covered with clingfilm and blotting was allowed to proceed overnight.

Hybridisation

Prior to prehybridisation, the membrane was briefly rinsed in a solution of 5x SSPE at room temperature. The membrane was prehybridised in 20ml of buffer containing 0.5M sodium phosphate, pH 7.2, 7% (w/v) SDS, 1mM EDTA and 20μl 10mg/ml denatured, fragmented salmon sperm DNA, at 65°C for 4-8hr. The probe was denatured in a boiling water bath, snap cooled on ice and added to the prehybridisation buffer. The membrane was allowed to hybridise overnight at 65°C.

The membrane was washed twice in a solution of 2x SSC, 0.1% (w/v) SDS for 10min each at room temperature followed by a wash at 65°C for 15min in 1x SSC, 0.1% (w/v) SDS, and a wash of 10 min in 0.1% (w/v) SSC, 0.1% (w/v) SDS at 65°C. The corner of the membrane was checked for signs of radioactivity using the Geiger-Muller counter. If the signal was high (greater than 15 counts per second), the final wash was repeated.

Phosphorimaging

The membrane was sealed in plastic and exposed to a phosphor screen as described in section 3.2.20.

Stripping membranes for reprobing

A boiling solution of 0.5% (w/v) SDS was poured over the membrane and allowed to cool to room temperature for 1hr. The membrane was air-dried for up to an hour before being sealed in plastic and stored at 4°C. The membrane was probed three times before being discarded.

4.3 Results

4.3.1 Selection of a small-scale RNA extraction technique for RT-PCR

It was important to isolate RNA of high quality for use in RT-PCR. In order to determine which RNA extraction technique was most suitable, RNA was extracted using two available small-scale extraction kits and the quality and quantity of the RNA compared.

General yield and purity of RNA samples

RNA was extracted from leafroll tissue of variety N19 using the RNeasy Plant Mini Kit (Qiagen) and the SV Total RNA Isolation System (Promega). The RNA was quantified spectrophotometrically. Figure 4.2 represents a typical RNA absorption spectrum.
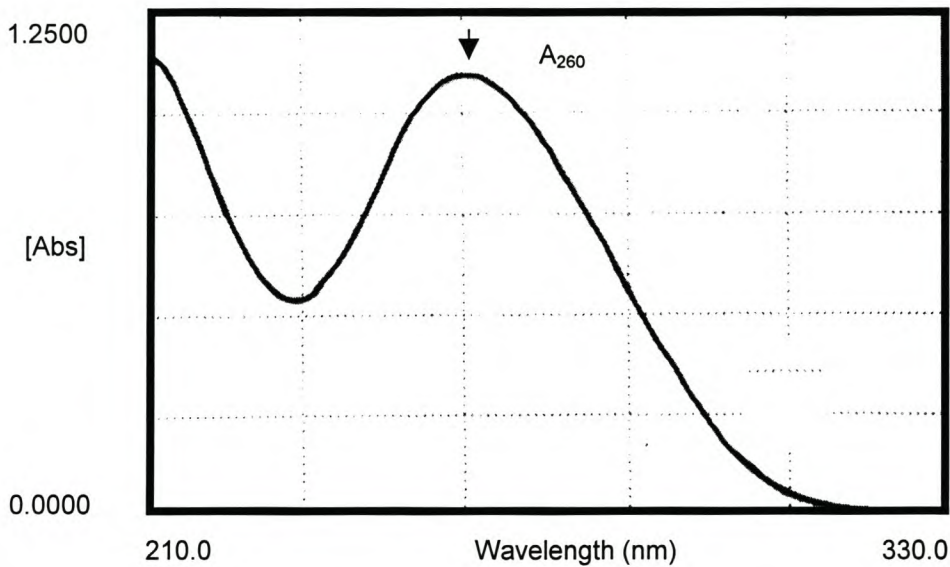
Figure 4.2 Typical absorbance profile of an RNA sample extracted from variety N19 using the RNeasy Plant Mini Kit (Qiagen).

The concentration of the RNA was calculated from the A_{260} value since an absorbance of 1 unit at 260nm corresponds to 40µg RNA per ml. The concentration of the above sample was calculated as follows:

$$A_{260} \text{ X } 40 \text{ X dilution factor}$$
$$= 1.0945 \text{ X } 40 \text{ X } 20$$
$$= 875.6 \text{ng/µl}$$

Pure RNA has an A_{260}/A_{280} of 2.0 (Sambrook *et al.*, 1989). The purity of the above RNA sample was:

$$\text{Purity} = A_{260}/A_{280}$$
$$= 1.0945 / 0.6330$$
$$= 1.7$$

Several replicate extractions were performed using the different kits and their yields compared. Table 4.2 represents the typical yields and purity of N19 leafroll RNA obtained with the two kits.

Table 4.2 Summary of the yields and purity of RNA samples extracted using different methods.

RNeasy Plant Mini Kit (Qiagen)		SV Total RNA Isolation Kit (Promega)	
Yield (ng/ mass original tissue)	Purity	Yield (ng/ mass tissue)	Purity
7.86	1.7	0.82	1.8
7.29	1.6	0.60	1.6
4.34	1.7	0.45	1.8
7.02	1.5	1.30	1.9
6.05	1.7	2.00	1.8

Although the SV Total RNA Isolation Kit yielded RNA of higher purity (average purity = 1.8) compared to the RNeasy Plant Mini Kit (average purity = 1.6), the latter kit yielded more RNA (Table 4.2).

Integrity of RNA samples

The integrity of the RNA was checked by running 10μg of an RNA sample on a denaturing formaldehyde/ agarose gel. Figure 4.3 shows the result of denaturing gel electrophoresis of RNA samples extracted from leafroll tissue of variety N19 using the SV Total RNA Isolation Kit and the RNeasy Plant Mini Kit.

1   2
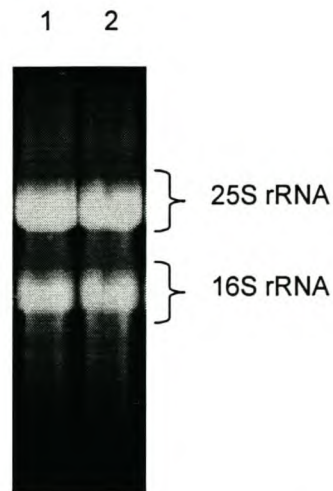
Figure 4.3 Denaturing gel electrophoresis of 10μg RNA isolated using the RNeasy Plant Mini Kit (lane 1) and the SV Total RNA Isolation Kit (lane 2).

In RNA samples that are not degraded, the 25S ribosomal RNA band is present in approximately twice the amount of the 16S RNA. Degradation of the sample is apparent if the inverse ratio is observed, and the ribosomal bands are not sharp, but appear as a smear towards the smaller sized RNAs. The results shown in Figure 4.3 confirmed that the RNA samples were not degraded.

DNA contamination

For RT-PCR, which is sensitive to residual DNA contamination in RNA extracts, it is important that the RNA is DNA-free. The SV Total RNA Isolation Kit (Promega) includes a DNase treatment step that removes any contaminating DNA in the RNA preparation. The RNeasy Plant Mini Kit does not include such a step. In order to confirm that the RNA extracted using the SV Total RNA Isolation Kit was free from DNA contamination, 50ng, 150ng and 300ng aliquots of RNA, extracted using this kit, was used as template in a PCR with random primer OPD-08 (Operon Technologies) and a RAPD PCR protocol. In addition, 75ng of genomic DNA was also amplified using this protocol. Figure 4.4 shows the results of the RAPD PCR on the different samples.
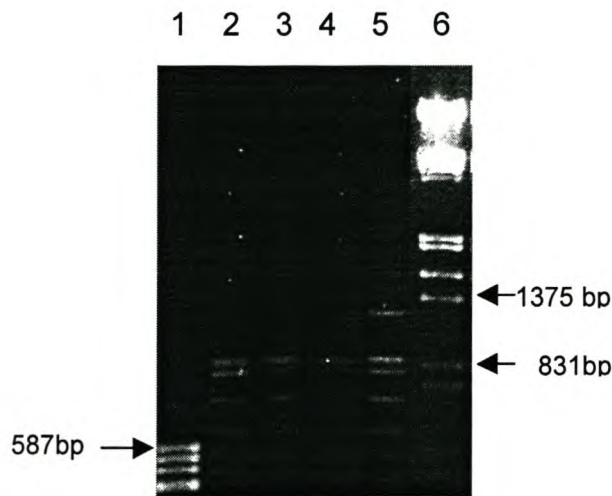
Figure 4.4 Results of RAPD PCR on RNA from N19 leafroll tissue extracted using the SV Total RNA Isolation Kit . Fifty nanograms (lane 2), 150ng (lane 3) and 300ng (lane 4) of RNA and 75ng aliquots of genomic DNA (lane 5) was amplified using the RAPD PCR protocol. Lane 6, lambda molecular weight marker III.

Amplification products were obtained even with low concentrations of RNA (Figure 4.4, lane 2). The banding profile obtained for all the concentrations of RNA were comparable to that obtained for genomic DNA (figure 4.4, lane 5). This suggests that genomic DNA was present in the RNA samples. Thus, the results indicate that the SV Total RNA Isolation System is not reliable in producing DNA-free RNA. One approach to prevent DNA contamination from confounding RT-PCR results is the use of intron spanning primers. PFP-B1 and PFP-B8 are such primers therefore it was not crucial to obtain DNase-free RNA for RT-PCR. Since the yields obtained with the RNeasy Plant Mini Kit were high (Table 4.2), this kit was selected for further RNA extractions.

4.3.2 Transcriptional expression of PFP-β gene variants

From the results of chapter 3, it was evident that two PFP-β variants were present in the genomes of commercial sugarcane varieties. However, it was unknown which of these

variants were being expressed. The RT-PCR technique was used to determine the transcriptional expression of the variants. It was hypothesised that at different stages of sucrose accumulation, different variants of the gene might be expressed. Therefore, transcriptional expression of the PFP-β gene variants was investigated in tissues of variety N19 at different stages of sucrose accumulation.

i) Expression of PFP-β gene variants in the maturing culm

One microgram of RNA from leaf, leafroll, internode 4 and internode 7 tissue of variety N19 was reverse transcribed using reverse transcriptase and the cDNA amplified using the specific primers PFP-B1 and PFP-B8. In addition, PFP-B1/B8 PCR was conducted on 75ng of genomic DNA as a control. The results of this amplification are shown in Figure 4.5.
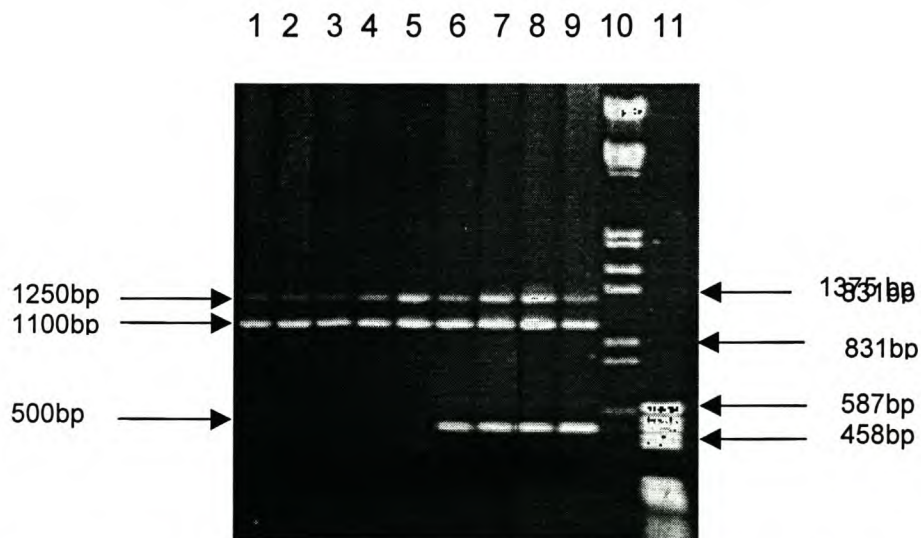
Figure 4.5 Agarose gel electrophoresis of PFP-B1/B8 RT-PCR products from tissues at different stages of sucrose accumulation. Lane 1-5, replicate PFP-B1/B8 amplification products from genomic DNA; lanes 6-9 are the PFP-B1/B8 RT-PCR products from leaf (lane 6), leafroll (lane 7), internode 4 (lane 8), and internode 7 (lane 9). Lanes 10 and 11 are molecular weight markers III and V respectively.

A 500bp fragment was consistently obtained from PFP-B1/B8 RT-PCR on leafroll, leaf, internode 4 and internode 7 tissue (Figure 4.5). In addition, two fragments (1250bp and

1100bp) were obtained. PFP-B1/B8 PCR on genomic DNA samples (Figure 4.5, lanes 1-5) confirmed that these were the same size fragments obtained from genomic DNA (section 3.3.2). Thus it was deduced that the extra fragments present in lanes 6-9 (Figure 4.5) were due to residual genomic DNA contamination in the RNA preparations.

The 500bp RT-PCR products from leaf, leafroll, internode 4 and internode 7 of variety N19 were cloned using the pGEM®-T Easy Vector System (Promega). Approximately 60-70% white colonies were obtained after overnight growth on selection plates. These colonies were tested using the protocol outlined in section 3.2.16.
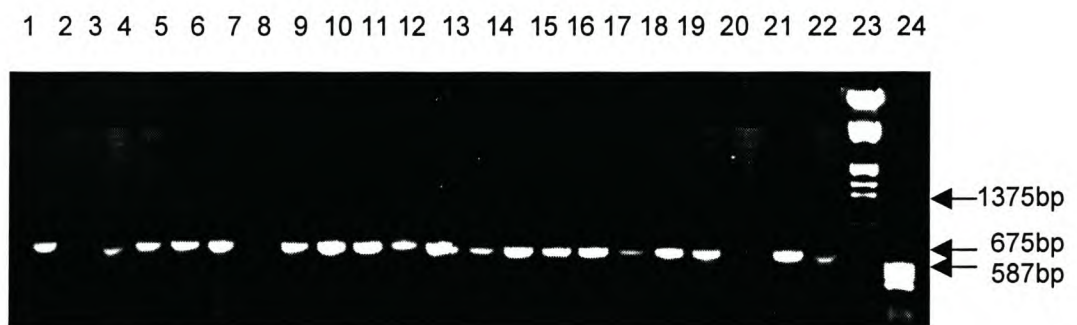
Figure 4.6 Results of SP6/T7 PCR on bacterial cells for the selection of positive PFP-B1/B8 RT-PCR transformants (lanes 1-22). Lane 23, lambda molecular weight marker III and lane 24, molecular weight marker V.

Figure 4.6 shows the results of insert verification using SP6 and T7 PCR. Ten positive colonies from each tissue were selected for sequencing using the vector primer T7 and the Big Dye Sequencing Kit (Applied Biosystems). Table 4.3 summarises the number of positive clones obtained per PFP-B1/B8 RT-PCR fragment from the different tissues and the preliminary results of sequencing.

Table 4.3 Results of cloning and preliminary sequencing of PFP-B1/B8 RT-PCR clones from different tissues of variety N19.

Tissue type from N19	Number of positive clones	Nucleotide sequence length, in bp, of 10 clones									
		#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
Leaf	16	304	473	473	473	478	473	474	473	473	475
Leafroll	15	482	484	485	373	486	484	486	493	485	455
Internode 4	19	485	483	484	472	484	472	484	474	483	441
Internode 7	16	498	475	409	450	484	484	482	481	473	495

Full-length sequences were obtained for almost all of the clones using the T7 forward vector primer (Table 4.3). In order to eliminate sequencing artifacts that may have arisen due to sequencing, one clone from each tissue was sequenced three times and any nucleotide inconsistencies were manually edited. These sequences were compared to the international database using BLASTn to determine the sequence homology to PFP-β. Table 4.4 shows the results of the BLASTn search for the PFP-B1/B8 RT-PCR sequence.

Table 4.4 BLASTn search results of the PFP-B1/B8 RT-PCR sequence.

Organism/ gene product	Score (bits)	E value
Oryza sativa PFP-β	157	6e-36
Citrus X paradisi PFP-β	72	3e-10
Solanum tubersosum PFP-β	66	2e-08
Ricinus communis PFP-β	54	6e-05

As shown in Table 4.4, all the sequences had high homology to PFP-β. An alignment of the 10 PFP-B1/B8 RT-PCR clones from leafroll was performed in DNASIS to generate a consensus sequence for the expressed sequence from this tissue. Consensus sequences were derived similarly for PFP-B1/B8 RT-PCR clones from internode 4, internode 7 and leaf.

The PFP-B1/B8 RT-PCR consensus sequences from each tissue were compared using the Maximum Matching (DNASIS). The results of this alignment are shown in Figure 4.7.

```
                        10         20         30         40         50         60
Internode 7    1 ATGGATTTCA TGCCGGAGGT TCAAAAACTC ATCGCAGAAT TGAATGAAAT TTTGGCACAT
leaf           1 ATGGATTTCA TGCCGGAGGT TCAAAAACTC ATCGCAGAAT TGAATGAAAT TTTGGCACAT
Leafroll       1 ATGGATTTCA TGCCGGAGGT TCAAAAACTC ATCGCAGAAT TGAATGAAAT TTTGGCACAT
Internode 4    1 ATGGATTTCA TGCCGGAGGT TCAAAAACTC ATCGCAGAAT TGAATGAAAT TTTGGCACAT
                        70         80         90        100        110        120
Internode 7   61 GATGTGGTTG ATGAGGCAGG GGCCTGGAAA AGCAAGCTTC AGCCTGAATC AAAGGAGCTG
leaf          61 GATGTGGTTG ATGAGGCAGG GGCCTGGAAA AGCAAGCTTC AGCCTGAATC AAAGGAGCTG
Leafroll      61 GATGTGGTTG ATGAGTCAGG GGCCTGGAAA AGCAAGCTTC AGCCTGAATC AAAGGAGCTG
Internode 4   61 GATGTGGTTG ATGAGGCAGG GGCCTGGAAA AGCAAGCTTC AGCCTGAATC AAAGGAGCTG
                       130        140        150        160        170        180
Internode 7  121 TTTGAGTTTT TGCCCAAAAC TATTCAGGAG CAACTTATGC TTGAAAGGGA CCCCCATGGC
leaf         121 TTTGAGTTTT TGCCCAAAAC TATTCAGGAG CAACTTATGC TTGAAAGGGA CCCCCATGGC
Leafroll     121 TTTGAGTTTT TGCCCAAAAC TATTCAGGAG CAACTTATGC TTGAAAGGGA CCCCCATGGC
Internode 4  121 TTTGAGTTTT TGCCCAAAAC TATTCAGGAG CAACTTATGC TTGAAAGGGA CCCCCATGGC
                       190        200        210        220        230        240
Internode 7  181 AATGTTCAGG TTGCAAAAAT TGAAACCGAG AAAATGCTTA TTAGCATGGT GGAAACTGAA
leaf         181 AATGTTCAGG TTGCAAAAAT TGAAACCGAG AAAATGCTTA TTAGCATGGT GGAAACTGAA
Leafroll     181 AATGTTCAGG TTGCAAAAAT TGAAACCGAG AAAATGCTTA TTAGCATGGT GGAAACTGAA
Internode 4  181 AATGTTCAGG TTGCAAAAAT TGAAACCGAG AAAATGCTTA TTAGCATGGT GGAAACTGAA
                       250        260        270        280        290        300
Internode 7  241 CTGGAGAAGA GAAAAGCAGA GGGGAGATAC TCTGCACATT TCAGAGGGCA AGCTCATTTC
leaf         241 CTGGAGAAGA GAAAAGCAGA GGGGAGATAC TCTGCACATT TCAGAGGGCA AGCTCATTTC
Leafroll     241 CTGGAGAAGA GAAAAGCAGA GGGGAGATAC TCTGCACATT TCAGAGGGCA AGCTCATTTC
Internode 4  241 CTGGAGAAGA GAAAAGCAGA GGGGAGATAC TCTGCACATT TCAGAGGGCA AGCTCATTTC
                       310        320        330        340        350        360
Internode 7  301 TTTGGGTACG AAGGAAGATG TGGCCTTCCT ACCAATTTTG ATTCTAACTA TTGCTATGCA
leaf         301 TTTGGGTACG AAGGAAGATG TGGCCTTCCT ACCAATTTTG ATTCTAACTA TTGCTATGCA
Leafroll     301 TTTGGGTACG AAGGAAGATG TGGCCTTCCT ACCAATTTTG ATTCTAACTA TTGCTATGCA
Internode 4  301 TTTGGGTACG AAGGAAGATG TGGCCTTCCT ACCAATTTTG ATTCTAACTA TTGCTATGCA
                       370        380        390        400        410        420
Internode 7  361 TTAGGCTATG GTGCTGGTGC CCTTCTCCAA AGTGGGAAGA CAGGACTTAT TTCATCGGTT
leaf         361 TTAGGCTATG GTGCTGGTGC CCTTCTCCAA AGTGGGAAGA CAGGACTTAT TTCATCGGTT
Leafroll     361 TTAGGCTATG GTGCTGGTGC CCTTCTCCAA AGTGGGAAGA CAGGACTTAT TTCATCGGTT
Internode 4  361 TTAGGCTATG GTGCTGGTGC CCTTCTCCAA AGTGGGAAGA CAGGACTTAT TTCATCGGTT
                       430        440        450        460        470
Internode 7  421 GGCAACCTTG TGGCTCCAGT AGAAGAATGG ACTGTTGGTG GAACAGCATT GAC.......
leaf         421 GGCAACCTTG TGGCTCCAGT AGAAGAATGG ACTGTTGGTG GAACAGCATT GAC.......
Leafroll     421 GGCAACCTTG TGGCTCCAGT AGAAGAATGG ACTGTTGGTG GAACAGCATT GAC.......
Internode 4  421 GGCAACCTTG TGGCTCCAGT AGAAGAATGG ACTGTTGGTG GAACAGCATT GAC.......
```

Figure 4.7 Alignment of the PFP-B1/B8 RT-PCR consensus sequences from tissues at different stages of sucrose accumulation. The alignment was derived using the multiple sequence comparison function of DNASIS (v 2.6).

The alignment in Figure 4.7 indicates that the PFP-B1/B8 RT-PCR products from tissues at different stages of sucrose accumulation are identical. Since the same variant was being expressed in tissues at different stages of sucrose accumulation, expression of the PFP-β gene variants was investigated in other tissues with different physiology to those originally tested.

ii) Expression of PFP-β gene variants in germinating sett nodal tissue, budscales and roots
PFP-B1/B8 RT-PCR was conducted on budscale, germinating sett nodal tissue and root tissue from variety N19. The same RT-PCR fragment (approximately 500bp) was obtained in all these tissues (gel electrophoresis results not shown).

iii) Expression of PFP-β gene variants in ancestral Saccharum species
In order to determine whether the same variant is expressed in the ancestral species, PFP-B1/B8 RT-PCR was conducted on RNA isolated from the leafroll tissue of Black Cheribon and Coimbatore. The results of this amplification also indicated a single RT-PCR product of approximately 500bp (gel electrophoresis results not shown).

Comparison of the transcriptional product from different sources
The RT-PCR products from Black Cheribon and Coimbatore leafroll tissue and from germinating sett nodal tissue, budscale and root tissue of N19, were sequenced directly without the intervention of cloning. The sequences from these tissue sources were aligned with the consensus sequences from leaf, leafroll, internode 4 and internode 7 using Contig Manager (DNASIS). Figure 4.8 represents this alignment using the exon 10 region as an example.

The transcriptionally expressed product in all the different tissue sources were identical (Figure 4.8).

```
                              10        20        30        40        50
Black Cheribon          1  CCTGAATCAA AGGAGCTGTT TGAGTTTTTG CCCAAAACTA TTCAGGAGCA
Bud Scale (N19)         1  CCTGAATCAA AGGAGCTGTT TGAGTTTTTG CCCAAAACTA TTCAGGAGCA
Germinating sett (N19)  1  CCTGAATCAA AGGAGCTGTT TGAGTTTTTG CCCAAAACTA TTCAGGAGCA
Coimbatore              1  CCTGAATCAA AGGAGCTGTT TGAGTTTTTG CCCAAAACTA TTCAGGAGCA
Root (N19)              1  CCTGAATCAA AGGAGCTGTT TGAGTTTTTG CCCAAAACTA TTCAGGAGCA
Leaf (N19)              1  CCTGAATCAA AGGAGCTGTT TGAGTTTTTG CCCAAAACTA TTCAGGAGCA
```

Figure 4.8 An alignment of part of the exon 10 region of the PFP-B1/B8 RT-PCR sequences from the different tissues of N19 (leafroll, leaf, internode 4, internode 7, root, bud scale and germinating sett nodal tissue) and the leafroll tissue of Black Cheribon and Coimbatore. The alignment was derived in Contig Manager (DNASIS).

Matching the transcribed product to the gene variant

The RT-PCR fragment, which was designated PFP-B1/B8$_{500bp}$, was compared to the derived exon 10, 11 and 12 regions of the PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ genomic DNA sequences using the Maximum Matching function of DNASIS (v 2.6). An example of this alignment is shown in Figure 4.9. Table 4.5 summarises the percentage nucleotide homology between the expressed sequence and the exon regions of the genomic DNA sequences.


```
A Matching Percentage (Total Window:
100%, Alignment Window: 100%)

TTGCAAAAATTGAAACCGAGAAAATGCTTATTAGCATGG
|||||||||||||||||||||||||||||||||||||||
TTGCAAAAATTGAAACCGAGAAAATGCTTATTAGCATGG

TGGAAACTGAACTGGAGAAGAGAAAAGCAGAGGGGAGAT
|||||||||||||||||||||||||||||||||||||||
TGGAAACTGAACTGGAGAAGAGAAAAGCAGAGGGGAGAT

ACTCTGCACATTTCAGAGGGCAAGCTCATTTCTTTGG
|||||||||||||||||||||||||||||||||||||
ACTCTGCACATTTCAGAGGGCAAGCTCATTTCTTTGG
```

```
B Matching Percentage (Total Window: 77%,
Alignment Window: 77%)

TTGCAAAAATTGAAACCGAGAAAATGCTTATTAGCA-TGG
|||||||||||||||||| |||||||||||||||| ||| |||
TTGCAAAAATTGAAACTGAGAAAATGCTTATT-GCAATGG

TGGAAACTGAACTGGAGAAGA-GAAAAG-CAGAGGGGAGA
| ||||| ||| | |||| || ||||||| | || || | |
TTGAAACGGAATTAGAGA-GAAGAAAAGGCTGAAGGAA-A

-TACTCTGCACA-TTTCAGAGGGCAAGCTCATTTCTTTGG
   |||   |   || ||||||   || |||  | ||  |||||||||
GTACAAGGG-CAGTTTCATTGGTCAATCCCACTTCTTTGG
```

Figure 4.9 Alignment of the exon 11 region of the genomic fragments PFP-B1/B8$_{1250bp}$ (A) and PFP-B1/B8$_{1100bp}$ (B) with the RT-PCR fragment. Alignments were made in the Maximum Matching function of DNASIS.

Table 4.5 Nucleotide sequence homology between the PFP-B1/B8$_{500bp}$ cDNA sequence and the PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ fragments respectively, derived from alignments made in the Maximum Matching function of DNASIS.

Exon Region	% Homology to PFP-B1/B8$_{1250bp}$	% Homology to PFP-B1/B8$_{1100bp}$
10	96	77
11	100	77
12	98	79

From the results presented in Figure 4.9 and Table 4.5, it was evident that the expressed sequence was closely homologous to the 1250bp genomic sequence but imperfectly homologous to the 1100bp sequence. This suggested that the variant characterised by the PFP-B1/B8$_{1250bp}$ fragment is being expressed in all the tissues sampled.

4.3.3 Comparison of PFP-β RT-PCR clones with expressed PFP-β sequences from other sugarcane sources

By comparing further expressed sugarcane PFP-β sequences to the PFP-B1/B8$_{500bp}$ RT-PCR sequence, the results shown in Figure 4.9 and Table 4.5, which indicate that the PFP-B1/B8$_{1250bp}$ variant is being expressed, could be confirmed. Thus, searches were made for further sugarcane PFP-β sequences. One resource available to obtain this was a cDNA macroarray prepared in-house (D. Carson[1]).

cDNA macroarrays

The cDNA macroarray consisted of a limited random selection of 1400 clones from 1), total cDNA libraries prepared from leafroll and the mature stalk and 2), subtractive cDNA libraries prepared from the immature and mature stalk (section 4.2.3). Thus, there was a possibility that a PFP-β clone would be present on the array. Also, the macroarray potentially provided a quick method for the isolation of further expressed PFP-β clones. To this end, the PFP-B1/B8 amplicons were used as probes to screen the macroarray.

---

[1] Deborah Carson, SASEX, Mount Edgecombe, Durban

i) Macroarray screening using genomic fragments as probes

The macroarray was screened using the PFP-B1/B8$_{1250bp}$ fragment as probe. No positive PFP-β clone was obtained (results not shown). Probing with the PFP-B1/B8$_{1100bp}$ fragment was also unsuccessful (results not shown). Since the genomic DNA fragments contained introns that could complicate hybridisation between the probe and the cDNA clone on the membrane, a PFP-β cDNA fragment was used as probe to screen the macroarray.

ii) Macroarray screening using cDNA fragments as probes

A 10μl aliquot of the sugarcane cDNA leafroll library was amplified using primers PFP-B2 and PFP-B4 (section 4.2.9). The results of this amplification are shown in Figure 4.10.

The 250bp cDNA fragment (Figure 4.10), designated PFP-B2/B4$_{250bp}$, was cloned and sequenced. International database scores indicate that the sequence was homologous to the exon 9 region of the PFP-β gene of potato (score (bits) = 78; E value = 2e-12) and castor bean (score (bits) = 70; E value = 5e-10). Figure 4.11 indicates the probes used in relation to the potato PFP-β gene sequence available on the GenBank database.
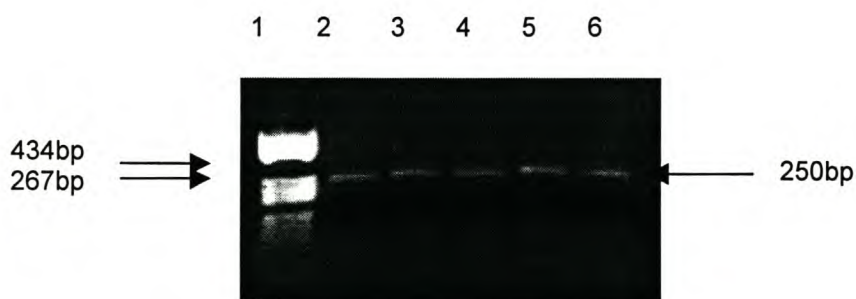
Figure 4.10 Agarose gel electrophoresis of PFP-B2/B4 PCR products from sugarcane cDNA leafroll library. Lane 1, molecular weight marker V; lanes 2-6, 250bp cDNA product from PFP-B2/B4 amplification.

Potato PFP-β cDNA sequence

PFP-B1/B8 1250bp genomic DNA sequence

PFP-B1/B8 1100bp genomic DNA sequence

PFP-B2/B4 250bp cDNA sequence

Exon 9 | Exon 10 | Exon 11 | Exon 12 | Exon 13
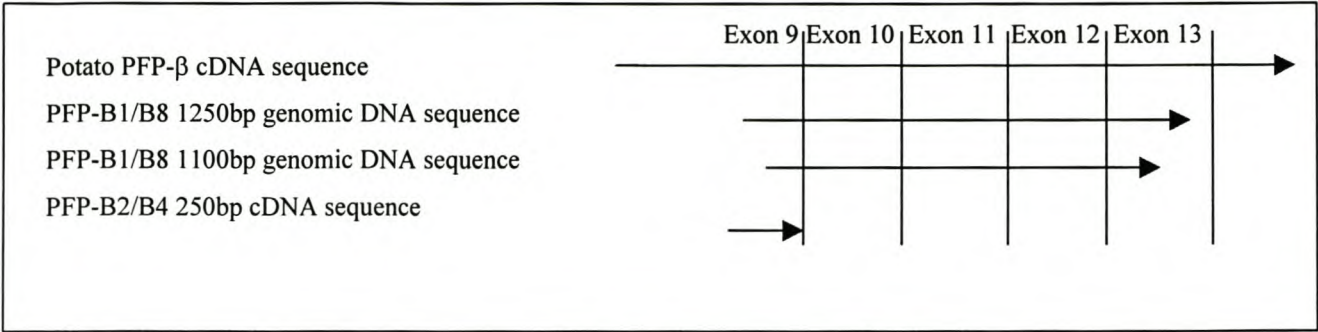
Figure 4.11 Diagram representing the PFP-B2/B4$_{250bp}$ and PFP-B1/B8 genomic DNA amplicons in relation to the potato PFP-β gene.

The PFP-B2/B4$_{250bp}$ amplicon was used as a probe to screen the macroarray. The results of this hybridisation are shown in Figure 4.12.

C

1 2 3 6
4 5 4 7
3 2 1 7
5 6 8 8
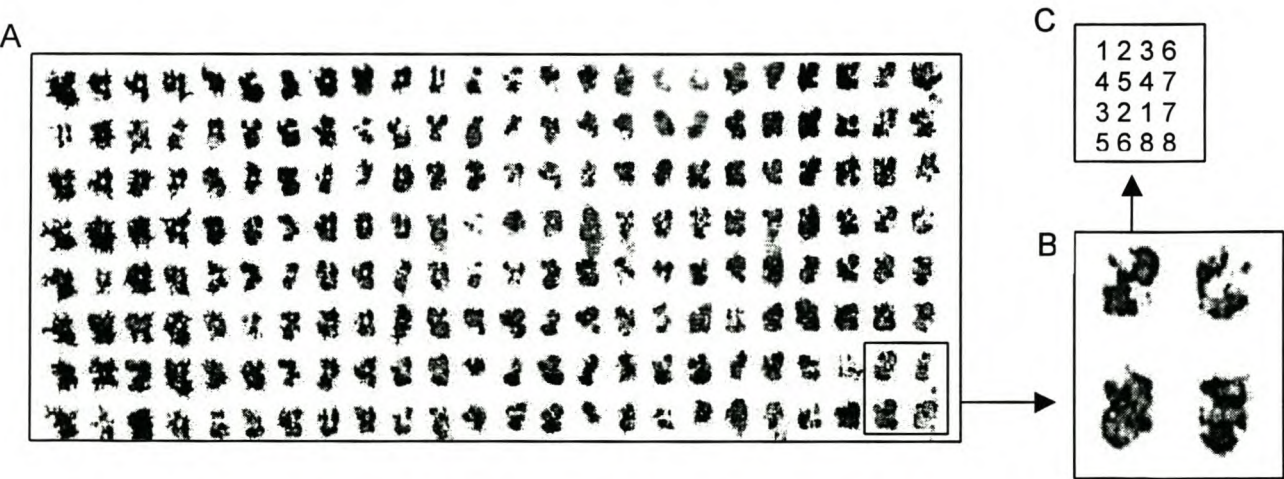
Figure 4.12 Results of cDNA macroarray screening using the PFP-B2/B4$_{250bp}$ cDNA fragment as a probe (A). (B) A magnified portion of the cDNA macroarray showing individual 4x4 matrices. (C) Diagram representing the 4x4 duplication pattern of clone deposition on the macroarray. Each number corresponds to an individual clone.

The PFP-B2/B4$_{250bp}$ probe hybridised to numerous clones on the membrane (Figure 4.12 (A)). Due to the manner in which the clones were deposited on the macroarray (Figure 4.12 (C)), a successful probing event would result in hybridisation to duplicate copies of the homologous clones on the membrane. However, this duplicate hybridisation was not observed. Instead non-specific hybridisation, as shown in Figure 4.12 (B), was obtained. Thus, no PFP-β clone could be isolated from the array. Unsuccessful probing was initially thought to be due to the nature of the probes used. The PFP-B1/B8$_{1250bp,}$ PFP-B1/B8$_{1100bp}$ and PFP-B2/B4$_{250bp}$ probes, which targeted specific regions of the PFP-β gene, may not have been homologous to any PFP-β clones on the membrane. Therefore, the array was probed with a full-length PFP-β cDNA clone from potato. This also resulted in non-specific hybridisation (results not shown). The presence of hybridisation artifacts indicated that conditions for the specific cDNA probe needed be optimised. The macroarray has been used successfully at SASEX to identify clones exhibiting variable levels of expression using total cDNA probes from a variety of tissue types (D. Carson, personal communication). The hybridisation conditions, which were used in screening the macroarray for PFP-β clones were similar to those used for total cDNA probes and may have been sub optimal for the specific target probes (PFP-B2/B4$_{250bp}$ and the full-length potato PFP-β cDNA fragments). Further optimisation of hybridisation conditions was not considered a practical approach to obtaining PFP-β clones since this exercise would have been time-consuming. Also, due to the limited number of clones on the membrane, PFP-β could be absent from the array altogether.

Confirmation of PFP-β variant expression by sequence comparison to other PFP-β sequences

Groenewald (1998) previously isolated a sugarcane PFP-β cDNA clone from the mature culm library which was sequenced and shown to include putative exons 10, 11 and 12. This sequence, which was available at SASEX, provided a means of confirming the results of section 4.3.2, which suggested that the PFP-B1/B8$_{1250bp}$ variant is expressed in sugarcane. The putative exon regions of the PFP-B1/B8$_{1250bp}$ sequence (derived as described in section 3.3.6) were compared to the cDNA sequence obtained by Groenewald

(1998). This comparative alignment was performed using the Maximum Matching function of DNASIS and is shown in Figure 4.13.

```
Matching Percentage (Total Window: 95%, Alignment Window: 95%)

    1   GTGCAGAATA-TCATTGCTGAATTGAATGAAATTTTGGCACATGATGTTG       50
        || || || |  |||| ||  |||||||||||||||||||||||||||| |
    1   GTTCAAAA-ACTCATCGCAGAATTGAATGAAATTTTGGCACATGATGTGG       50

   51   -TTG-ATG-AGGCAGGGGCCTGGAAAAGCAAGCTTCAGC-C-TGAATCAA      100
         |||  |||  |||||||||||||||||||||||||||||  | ||||||||
   51   GTTGGATGGAGGCAGGGGCCTGGAAAAGCAAGCTTCAGCGCCTGAATCAA      100

  101   AGGAGCTGTTTGAGTTTTTGCCCAAAA-CTATTCAGGAGCAAC-TTATGC      150
        |||||||||||||||||||||||||||| |||||||||||||| ||||||
  101   AGGAGCTGTTTGAGTTTTTGCCCAAAAACTATTCAGGAGCAACCTTATGC      150

  151   TTGAAAGGGGCCCCCATGGCAATGTTCAGGTTGCAAAAATTGAAACCGAG      200
        ||||||||| ||||||||||||||||||  ||||||||||||||||||||
  151   TTGAAAGGGACCCCCATGGCAATGTTCAG-TTGCAAAAATTGAAACCGAG      200

  201   AAAATGCTTATTAGCATGGTGGAAACTGAACTGGAGAAGAGAAAAGCAGA      250
        ||||||||||||||||||||||||||||||||||||||||||||||||||
  201   AAAATGCTTATTAGCATGGTGGAAACTGAACTGGAGAAGAGAAAAGCAGA      250

  251   GGGGAGATACTCTGCACATTTCAGAGGGCAAGCTCATTTCTTTGGGTACG      300
        |||||||||||||||||||||||||||||||||||||||||||||| ||||
  251   GGGGAGATACTCTGCACATTTCAGAGGGCAAGCTCATTTCTTTGG-TACG      300

  301   AAGGAAGATGTGGCCTTCCTACCAATTTTGATTCTAACTATTGCTATGCA      350
        ||||||||||||||||||||||||||||||||||||||||||||||||||
  301   AAGGAAGATGTGGCCTTCCTACCAATTTTGATTCTAACTATTGCTATGCA      350

  351   TTAGGCTATGGGGCTGGTG-CCCTTCTCC-AAAGTGGGAAGACAGGACTT      400
        ||||||||||| |||||||  |||||||| |||||||||||||||||||||
  351   TTAGGCTATGGTGCTGGTGGCCCTTCTCCCAAAGTGGGAAGACAGGACTT      400

  401   ATTTCATC........................................      450
        ||||||||
  401   ATTTCATC........................................      450
```

Figure 4.13 An alignment of the putative exon 10, 11 and 12 regions of part of the PFP-β gene sequence isolated by Groenewald (1988) from a mature culm cDNA library (first sequence) with the exon regions of PFP-B1/B8$_{1250bp}$ sequence (second sequence). The alignment was derived in Maximum Matching (DNASIS). Putative exon regions were derived for the PFP-B1/B8$_{1250bp}$ by comparison with the known castor bean exons using Contig Manager (DNASIS). Putative exon 10 is indicated in green, while putative exons 11 and 12 are indicated in red and blue respectively.

The high sequence homology between the PFP-B1/B8$_{1250bp}$ cDNA sequence with the expressed PFP-β sequence obtained by Groenewald (1988) (figure 4.13), indicated that the variant characterised by the PFP-B1/B8$_{1250bp}$ sequence is expressed in sugarcane tissues.

4.3.4 Re-analysis of the PFP-B1/B8 genomic fragments

Comparison of the PFP-B1/B8 RT-PCR sequence to the derived exon regions of the PFP-B1/B8$_{1250bp}$ consensus sequence reveals high homology which suggests that the variant characterised by the PFP-B1/B8$_{1250bp}$ sequence is being expressed in sugarcane (Table 4.7). However, the results of the alignment of individual PFP-B1/B8$_{1250bp}$ sequences (Figure 3.20) and the homology tree in Figure 3.22 suggests that there may be sequence variation within the PFP-B1/B8$_{1250bp}$ fragment population. Little to no sequence variation was detected in the PFP-B1/B8 RT-PCR sequences from the respective tissue sources. The discrepancy between the lack of sequence variation within the PFP-B1/B8 RT-PCR sequences and the presence of some sequence variation within the PFP-B1/B8$_{1250bp}$ fragment population suggests that there may be different expressed sequences within the PFP-B1/B8$_{1250bp}$ fragment population. This was investigated using sequence analysis. Individual PFP-B1/B8$_{1250bp}$ sequences were aligned to the known castor bean PFP-β exons using the Contig Manager function of DNASIS (as described in Figure 3.26) and the putative exons 10, 11 and 12 derived. These exon regions were then compared to the exon regions of the PFP-B1/B8 RT-PCR sequence using the Maximum Matching function of DNASIS and homology values computed (as described in Figure 4.9). The percentage homology between the putative exon regions of the individual PFP-B1/B8$_{1250bp}$ sequences with the exon regions of the PFP-B1/B8 RT-PCR sequence is presented in Table 4.6.

In the same vein, the alignment shown in Figure 3.20 and the homology tree (Figure 3.21) suggests that there may be sequence variation within the PFP-B1/B8$_{1100bp}$ fragment. This implies that there may be some sequences within the PFP-B1/B8$_{1100bp}$ fragment population which may show higher homology to the PFP-B1/B8 RT-PCR sequence than those values indicated in Table 4.5. To investigate this, the derived exon regions of the individual PFP-B1/B8$_{1100bp}$ fragments were compared to the exon regions of the PFP-B1/B8 RT-PCR

fragment. Exon regions were derived using Contig Manager and comparisons were performed using Maximum Matching as described for the individual PFP-B1/B8$_{1250bp}$ sequences. The percentage homology values (computed in Maximum Matching) for the comparison of the putative exon regions from individual PFP-B1/B8$_{1100bp}$ sequences to the PFP-B1/B8 RT-PCR sequence are presented in Table 4.7.

Table 4.6 Summary of the homology between the putative exon regions of individual PFP-B1/B8$_{1250bp}$ clones and the PFP-B1/B8 RT-PCR sequence. Putative exon regions of the PFP-B1/B8$_{1250bp}$ sequences were identified by comparing each clone to the exon 10, 11 and 12 regions of castor bean PFP-β. The percentage similarities were derived in the Maximum Matching function of DNASIS.

Variety	Clone number	% homology to exon 10 of PFP-B1/B8 cDNA	% homology to exon 11 of PFP-B1/B8 cDNA	% homology to exon 12 of PFP-B1/B8 cDNA
N19	1	96	100	98
	2	96	100	98
	3	96	100	96
	4	96	100	98
	5	96	96	98
	6	95	96	98
	7	96	96	96
	8	96	96	98
	9	95	97	98
	10	96	96	98
N21	1	96	100	97
	2	95	100	96
	3	95	100	98
	4	96	99	96
	5	96	100	98
	6	95	100	96
	7	95	100	96
	8	96	99	98
	9	96	100	96
	10	96	99	98

The percentage homologies between the putative exon regions of the PFP-B1/B8$_{1250bp}$ fragment and the exon regions of the PFP-B1/B8 RT-PCR sequence are not significantly

different from each other (Table 4.6). This suggests that there is a single expressed sequence within the PFP-B1/B8$_{1250bp}$ sequence.

Table 4.7 Summary of the homology between the putative exon regions of individual PFP-B1/B8$_{1100bp}$ clones and the PFP-B1/B8 RT-PCR sequence. Putative exon regions of the PFP-B1/B8$_{1100bp}$ sequences were identified by comparing each clone to the exon 10, 11 and 12 regions of castor bean PFP-β in Contig Manager. The percentage similarities were derived in the Maximum Matching function of DNASIS.

Variety	Clone number	% homology to exon 10 of PFP-B1/B8 cDNA	% homology to exon 11 of PFP-B1/B8 cDNA	% homology to exon 12 of PFP-B1/B8 cDNA
N19	1	75	77	79
	2	77	77	79
	3	76	77	79
	4	77	76	79
	5	77	77	79
	6	77	76	79
	7	77	77	78
	8	77	77	77
	9	77	75	79
	10	77	76	79
N21	1	77	77	79
	2	77	77	78
	3	75	77	79
	4	77	77	79
	5	72	76	79
	6	77	77	78
	7	77	77	78
	8	77	77	79
	9	76	77	77
	10	77	76	78

All the individual PFP-B1/B8$_{1100bp}$ sequences show similar percentage homologies in their derived exon regions compared to the exon regions of the PFP-B1/B8 RT-PCR sequence (Table 4.7). The results indicate that all the PFP-B1/B8$_{1100bp}$ sequences are poorly homologous to the PFP-B1/B8 RT-PCR sequence compared to the high percentage homology obtained for the individual PFP-B1/B8$_{1250bp}$ sequences (shown in Table 4.6).

4.3.5 Quantitative expression of the PFP-β gene in relation to sucrose accumulation

Previous investigations suggest that PFP activity increases from the younger to the older internodes of sugarcane (Botha *et al.*, 1996; Whittaker and Botha, 1997; Whittaker and Botha, 1999). This suggests that there could be differential expression of the PFP-β gene in different tissues of sugarcane. The results of the RT-PCR indicate expression of a single transcript for the PFP-β gene, however this technique is sensitive to trace amounts of mRNA and does not reveal the level of expression of the transcripts in relation to each other. In order to determine the differential expression of the PFP-β gene at different stages of sucrose accumulation in sugarcane, Northern hybridisation analysis was performed on leaf, leafroll, internode 4 and internode 7 tissues of variety N19. Initially, a membrane containing 10μg of RNA was probed with the PFP-B1/B8500bp cDNA fragment from internode 4. Figure 4.14 (B) shows the result of this hybridisation. Even after 10 days of exposure to the phosphorscreen, only a faint signal was observed in leafroll tissue. According to S. Groenewald (personal communication[1]) no signal could be observed using 10μg of RNA on a Northern blot probed with a PFP-β cDNA fragment. The concentration of RNA was increased to 20μg before a signal could be obtained. Therefore, a second membrane was prepared using 20μg of RNA from the respective tissues and reprobed using the PFP-B1/B8500bp cDNA fragment from internode 4. The results of this hybridisation are shown in Figure 4.14 (C).

A single mRNA species of approximately 2.0kb was obtained in leafroll and internode 4 tissues with higher level of expression in the leafroll. No PFP-β expression was detected in leaf and internode 7 tissues (Figure 4.14 (C)).

Figure 4.14 (A) shows that the RNA samples were loaded equally in all the lanes and that the RNA was not degraded in that the ratio of ribosomal bands were not reversed (as discussed in section 3.4.1). The presence of an extra band in lane 1 of figure 4.14 (A) is perhaps due to the 5S rRNA.

---

[1] Sarita Groenewald, Institute for Plant Biotechnology, University of Stellenbosch
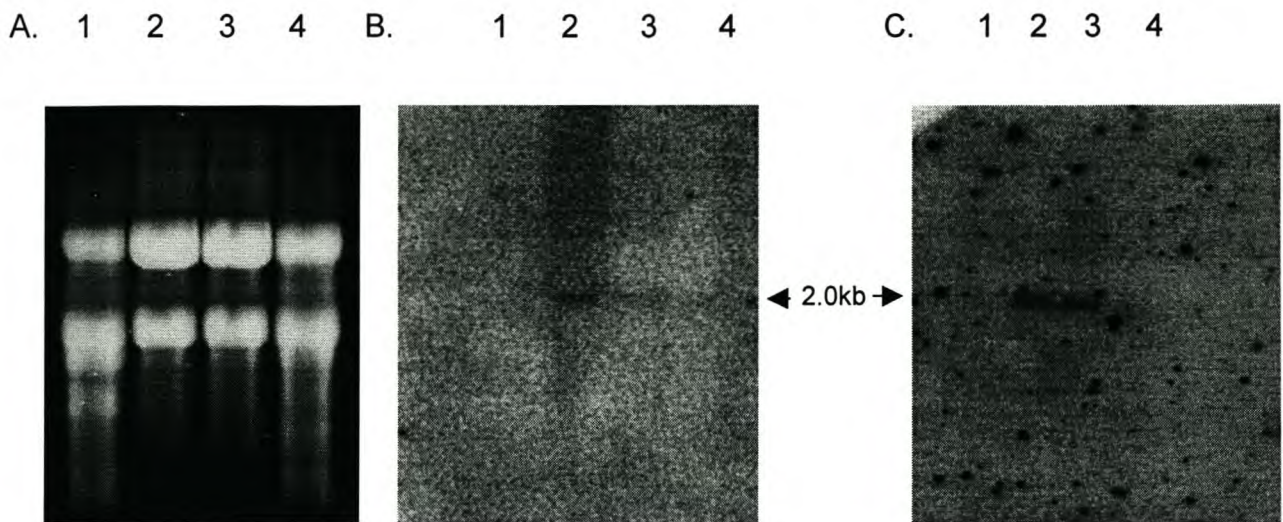
Figure 4.14 Northern hybridisation analysis of the PFP-β gene in tissues at different stages of sucrose accumulation. Formaldehyde/agarose gel electrophoresis of the RNA samples is shown in (A). Ten micrograms (B) and 20μg (C) of RNA was separated and the PFP-B1/B8 RT-PCR fragment used as probe. RNA from leaf, leafroll, internode 4 and internode 7 tissues were loaded in lanes 1, 2, 3 and 4 respectively.

4.4 Discussion

Since significant distinctions had been resolved in the two variants of the PFP-β gene at the genomic sequence level (Chapter 3), it was seen as important to sequence the transcripts.

A single PFP-B1/B8 RT-PCR fragment was isolated from tissues at different stages of sucrose accumulation in the stalk and in budscale, germinating sett nodal tissue and root tissues of a commercial sugarcane variety (N19). This transcript was also obtained from the ancestral varieties, Black Cheribon and Coimbatore. International database searches revealed that the fragment is homologous to PFP-β. Alignments of the RT-PCR sequences obtained from the different tissue sources indicate that the cDNA sequences are identical. Comparison of the putative exon regions of the genomic variants (PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$) with the transcriptionally expressed 500bp sequence suggests that the variant characterised by the PFP-B1/B8$_{1250bp}$ fragment is the one being expressed. These

results are consistent with the sequence comparison by Harvey (1998) which suggested that a cDNA fragment obtained from a sugarcane leafroll library was more homologous to the exon regions of the PFP-B1/B8$_{1250bp}$ fragment than to that of the PFP-B1/B8$_{1100bp}$ sequence.

Intapruk *et al.* (1993) also used the RT-PCR, cloning and sequencing approach to determine the exon regions of the corresponding genomic DNAs encoding *Arabidopsis thaliana* peroxidase isozymes. The authors showed that both genes were being expressed however, the one gene was subjected to organ-specific expression. Although there are no known isozymes for PFP, in this study it was thought that the PFP-β gene variants (PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$) might be subjected to tissue specific expression. PFP has been described as an adaptive enzyme as its activity is responsive to environmental stresses such as phosphate nutrition and anaerobiosis, and to developmental or tissue-specific cues (Duff *et al.*, 1989; Stitt, 1990; Mertens, 1991; Theodorou *et al.*, 1992; Theodorou and Plaxton, 1993). In the present study, transcriptional expression of the PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ gene variants was not investigated under these varied conditions. The PFP-B1/B8$_{1100bp}$ variant may be expressed during conditions of environmental stresses and other developmental stages of sugarcane and should thus be investigated in this context in the future.

As shown by Harvey *et al.* (1999), the two PFP-β variants PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ are also present in the ancestral species representatives Black Cheribon and Coimbatore. In the current study it was shown that the same variant (PFP-B1/B8$_{500bp}$) is expressed as mRNA in these ancestral varieties and in the commercial hybrid variety (N19). This suggests that the gene characterised by the PFP-B1/B8$_{1250bp}$ variant may be an ancestral gene that has persisted over evolutionary history. According to Paterson (1996), essential genes may remain unchanged over generations since mutations in the coding region may prove fatal to the organism.

Northern hybridisation analysis of tissues at different stages of sucrose accumulation revealed a mRNA species of approximately 2kb for the PFP-β gene. This compared

favourably with the results obtained by Carlisle *et al.* (1990) who obtained a RNA transcript of 2.1kb in potato tissues and with Blakeley *et al.* (1992) who found a single mRNA species of 2kb for PFP-β in castor bean.

Differential expression of PFP-β was observed in tissues at different stages of sucrose accumulation. Other authors also report differential expression of PFP-β. In potato tissues, differential expression was observed in potato tuber, leaf and sprout tissue with lowest level of expression in the leaf and highest in the tuber (Carlisle *et al.*, 1990). According to Blakeley *et al.* (1992) there is transcriptional and tissue-specific regulation of expression of the PFP subunits in developing castor seeds. Their results suggested that the regulation of expression is posttranscriptional. Similarly, Botha and Botha (1991) showed that the concentration of the PFP subunits varied independently in the different tissues of *Citrillus lanatus*.

Low PFP-β expression and activity in leaf tissue has been observed in other plant species. For example in potato (Carlisle *et al.*, 1990) and in the mature leaf of *Citrillus lanatus* (Botha and Botha, 1991). This is in accordance with the concept that PFP is generally associated with sink tissue (Black *et al.*, 1987). Thus it was expected that very low expression of PFP-β, the catalytic subunit of the enzyme, would be found in leaf (a source tissue).

Northern analysis indicates that PFP-β expression has fallen off by internode 7. This is in sharp contrast to the results of Whittaker and Botha (1999) who found high PFP enzyme activity corresponding to high protein expression for the beta subunit of the enzyme in this tissue. Whittaker and Botha (1999) did not conduct Northern analysis on internode 7 tissue and in the scope of this study, PFP-β protein analysis was not performed thus the results are not directly comparable. However, it is possible that during the transition stage in the maturing stalk the mRNA levels may have fallen well in advance of the decline of the protein. This may explain the discrepancy between the results of Whittaker and Botha (1999) and the results presented here.

Whittaker and Botha (1999) looked at the activity of PFP across sugarcane varieties and showed an inverse correlation with sucrose. This may indicate that expression might depend on the sucrose status of the internode, which could vary from individual plants even within internode 7. Whittaker and Botha (1999) sampled sugarcane plants which were over a year old, while in this study plants that were approximately seven months to a year old were used. Although both studies sampled the sugarcane cultivar N19, the different ages of sampling may indicate that the tissue may not be at the same stage of development. From studies on *Citrillus lanatus*, Botha and Botha (1991) showed that the PFP subunit concentration varies in the same tissue at different stages of development. This suggests that PFP-β expression, which has been shown to correspond with PFP activity in sugarcane (Whittaker and Botha, 1999) may be different in internode 7 tissue of N19 sampled at different times.

Recently, the PFP-β gene in sugarcane was used as the source of a culm-specific promoter to drive transgene expression in the mature culm (Groenewald, 1998). Comparison of the cDNA sequence obtained by Groenewald and the transcriptionally expressed sequence characterised here (PFP-B1/B8$_{500bp}$) confirms that one sequence is being expressed in sugarcane. The PFP-β region was selected as the source of the promoter due to the high activity of PFP in the mature culm (internode 6 and 7) (Whittaker and Botha, 1997). However, Northern hybridisation analysis indicates that the PFP-β gene is more highly expressed in the younger tissue (leafroll) compared to the mature culm (internode 7) where expression was undetectable.

Despite the discrepancy, the current Northern analysis indicates that PFP-β is highly regulated during maturation. This has been observed previously in sugarcane (Whittaker and Botha, 1999) and is in accordance with the suggestion that PFP activity is in part modulated by coarse regulation of transcriptional and / or translational factors which has been deduced from studies in seeds of cucumber (Botha and Botha, 1991), castor bean (Blakeley *et al.*, 1992) and barley (Nielson, 1994).

CHAPTER 5

General Discussion and Conclusion

PFP has been recognised as a key regulatory enzyme in sucrose metabolism and it may therefore provide a target for genetic manipulation to increase sucrose levels in sugarcane. The PFP-β gene, which encodes the catalytic subunit of the enzyme, offered a potential model for the study of allelic variation in sugarcane, a complex polyploid.

Variation at the PFP-β genic region was investigated using PCR with specific primers. Amplification of genomic DNA from two hybrid commercial varieties revealed two PFP-β amplicons differing in size by approximately 150bp. Sequence analysis of the two amplicons (designated PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ respectively) indicated that a higher degree of variation existed within the putative intron compared to the exon regions. Minor sequence variations, which may be attributed to the comparison of single pass sequences, were detected within the two amplicon populations. However, further sequencing of individual clones of the respective PFP-β variants may be necessary to resolve the variation within the amplicon populations with more precision. This may reveal distinctly different sequences within the PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ fragment populations with higher homology to each other than that presented here. The results of PCR, cloning and sequencing indicated that two major variants of the PFP-β gene exist in sugarcane. It was hypothesised that PFP-β may be a single locus gene in sugarcane and that these variants may be allelic.

This hypothesis was tested by investigating the nature of the PFP-β gene within the genome. Southern hybridisation analysis, using the PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ amplicons as probes, revealed a simple hybridisation pattern. Recently, Harris *et al.* (2000) used Southern hybridisation analysis to characterise two potato nitrate reductase cDNA fragments. Their results indicated a single nitrate reductase gene, and mapping studies revealed a single nitrate reductase locus in the potato genome. Thus, they concluded that the two cDNA fragments, which were 99.7%

identical at the amino acid level, were alleles. In the present study, it is speculated that the variants (PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$ amplicons) represent separate loci, which may have arisen as a result of duplication. From the results of PCR, it may be speculated that the two variants do not segregate since both amplicons (PFP-B1/B8$_{1250bp}$ and PFP-B1/B8$_{1100bp}$) are consistently produced from genomic DNA from both commercial and ancestral varieties. Also, the Southern hybridisation pattern for the two PFP-β variants is simple. Thus, it is proposed that the variants may represent separate loci, which are closely linked on the same chromosome. As shown by Harris *et al.* (2000), future restriction fragment length polymorphism (RFLP) mapping studies on a sugarcane mapping population may serve to confirm the latter hypothesis.

Only the variant characterised by the PFP-B1/B8$_{1250bp}$ appears to be expressed in commercial sugarcane as well as in the ancestral *Saccharum officinarum* and *Saccharum spontaneum* species representatives. According to Wendel (2000), duplicate genes may retain their original or similar function, undergo changes in protein function or regulation, or one copy may become silenced through mutational or epigenetic means. Working with present data, it can be speculated that the other variant characterised by the PFP-B1/B8$_{1100bp}$ fragment arose as a result of gene duplication, but accumulated mutations in the coding regions so that the gene product was silenced, resulting in a pseudogene. Given that this variant is also not expressed in the ancestral species representatives, it can be suggested that the variant is representative of an ancestral pseudogene which evolved long before the evolution of the *Saccharum* species ancestors.

Mutations within the PFP-β variants were not identified since only exons 10, 11 and 12 were compared by sequence analysis. Future analysis of the open reading frames of the entire PFP-β gene variants may be valuable in classifying the variant, characterised by the PFP-B1/B8$_{1100bp}$ fragment, as a pseudogene. Numerous plant pseudogenes have been characterised by sequence analysis. Gottlieb and Ford (1997) analysed 18 exons of the PgiC2 gene in *Clarkia mildrediae*, which revealed numerous mutations in the coding regions that suggested that the gene was a pseudogene. Longhurst *et al.* (1994) identified multiple alcohol dehydrogenase

related sequences in tomato which were cloned and shown to have 94% similarity to each other and 77% similarity with the tomato ADH 2 gene over a 1000bp region. The homologous regions included introns and exons but the equivalent exons contained frame shifts, deletions and stop codons. Thus the two sequences were considered pseudogenes. Suzuki *et al.* (1997) characterised three clustered members of the self –incompatibility (S) multigene family in *Brassica*, two of which were considered to be pseudogenes because deletions causing frameshifts were identified in the sequences.

Two pathogenesis related (PR) protein λ- genomic clones from tobacco were shown to be pseudogenes by sequence analysis which revealed mutations in the highly conserved PR-1 protein coding region (Pfitzner *et al.*, 1991). In addition, expression studies, which involved the fusion of the upstream sequences of the clones into the *E.coli* β–glucoronidase (GUS) reporter gene and the insertion of the protein coding region of one of the clones into transgenic tobacco plants, resulted in no detectable transcript or protein product, which indicated that the clones were not active *in planta*.

Although evidence from expression analysis suggests that the gene characterised by the smaller amplicon (PFP-B1/B8$_{1100bp}$) is a pseudogene, rigorous sequencing of individual PFP-B1/B8$_{1100bp}$ clones may reveal variable sequences some of which may be more homologous to the expressed sequence. Also, further investigation of expression of this gene must be considered under various conditions of environmental stresses, which is likely to affect PFP activity and in turn expression of the β subunit variants.

Based on the relationship previously shown to exist between PFP activity and sucrose, the level of PFP expression at different stages of sucrose accumulation was investigated using Northern hybridisation analysis. PFP-β is differentially expressed at various stages of sucrose accumulation in commercial sugarcane with higher levels of expression in the more immature, low-sucrose containing tissue. These results, together with the findings of Whittaker and Botha (1999), who reported that PFP activity, which was apparently controlled by the expression of the β-subunit of the enzyme, is inversely correlated with sucrose content, suggest that it may be

possible to increase sucrose levels in the more mature sucrose accumulating tissues of sugarcane by decreasing PFP-β expression and consequently PFP activity.

In summary, it was concluded that although variation does occur at the PFP-β genic region, these variants might not be allelic, rather separate genes, one of which appears to be a pseudogene. Evidence from the present study suggests that the gene variant characterised by the PFP-B1/B8$_{1250bp}$ amplicon should be targeted for future genetic manipulation of PFP-β in sugarcane.

REFERENCES CITED

Al-Janabi, SM, Honeycutt, RJ, McClelland, M and Sobral, BWS (1994). A genetic linkage map of *Saccharum spontaneum* (L) 'SES 208'. *Genetics* **134**: 12499-1260.

Al-Janabi, SM (1998). Genetics, phylogenetics and comparative genetics of *Saccharum* L., a polysomic polyploid Poales: Andropogoneae. *Biotechnology Annual Review* **4**: 285–319.

Altschul, SF (1998). Fundamentals of database searching In Brenner, S, Patterson, M, Handel, M (eds.) Trends Guide to Bioinformatics, Elsevier Science, pp. 7-9.

Altschul, SF, Gish, W, Miller, W, Myers, EW and Lipman, DJ (1990). Basic local alignment search tool. *Journal of Molecular Biology* **215**: 403-410.

Anonymous (1996). Economic and agricultural review, Central Economic and Advisory Service, Pretoria, pp. 3-15.

Anonymous (2000). South African Sugar Association (SASA) annual report (1999/2000), pp. 1-32.

ap Rees, T, Morrell, S, Edwards, J, Wilson, PM and Green, JH (1985). Pyrophosphate and the glycolysis of sucrose in higher plants. *In*: Heath, RL and Preiss, J (eds.) Regulation of carbon partitioning in photosynthetic tissue. American Society of Plant Physiologists, Rockville, MD, pp. 76-92.

Arumuganathan, K and Earle, ED (1991). Nuclear DNA content of some important plant species. *Plant Molecular Biology Reporter* **215**: 403-410.

Ashihara, H and Stupavaska, S (1984). Comparison of activities and properties of pyrophosphate- and adenosine triphosphate-dependent phosphofructokinases of black gram (Phaseolus mungo) seeds. *Journal of Plant Physiology* **116**: 241-252.

Bennett, MD and Smith JB (1991). Nuclear DNA amounts in angiosperms. *Philosophical Transactions of the Royal Society of London (Biological)* **334**: 309-345.

Biles, CL and Martyn, RD (1993). Peroxidase, polyphenoloxidase, and shikimate dehydrogenase isozymes in relation to tissue type, maturity and pathogen induction of watermelon seedlings. *Plant Physiology and Biochemistry* **31** (4): 499-506.

Birchler, JA and Newton, KJ (1981). Modulation of protein levels in chromosomal dosage series of maize: the biochemical basis of aneuploid syndromes. *Genetics* **99**: 247-266.

Black, CC, Mustardy, L, Sung, S-S, Komanik, PP, Xu, DP, Paz, N (1987). Regulation and roles for alternative pathways of hexose metabolism in plants. *Plant Physiology* **63**: 387-394.

Blakeley, SD (1997). The manipulation of resource allocation in plants *In*: Dennis, DT, Turpin, DH, Lefebvre, DD and Layzell, DB (eds.) Plant Metabolism, 2nd edition, Addison Wesley Longman, United Kingdom, pp. 580–591.

Blakeley, SD, Crews, L, Todd, JF and Dennis, DT (1992). Expression of the genes for the α- and β- subunits of the pyrophosphate dependent phosphofructokinase in germinating and developing seeds from *Ricinus communis*. *Plant Physiology* **99**: 1245-1250.

Blakeley, SD and Dennis, DT (1993). Molecular approaches to the manipulation of carbon allocation in plants. *Canadian Journal of Botany* **71**: 765-778.

Botha, A-M and Botha, FC (1991). Pyrophosphate dependent phosphofructokinase of *Citrillus lanatus*: molecular forms and expression of subunits. *Plant Physiology* **96**:1185-1192.

Botha, A-M and Botha FC (1993). Effect of the radicle and hormones on the subunit composition and molecular form of pyrophosphate dependent phosphofructokinase in the cotyledons of *Citrillus lanatus. Australian Journal of Plant Physiology* **20**: 265-273.

Botha, FC and Dennis, DT (1986). Isozymes of phosphoglyceromutase from the developing endosperm of *Ricinus communis*: Isolation and kinetic properties. *Archives of Biochemistry and Biophysics* **245**: 96-103.

Botha, FC, De Vries, C, Small, JGC (1986). Isolation and characterisation of pyrophosphate:D-fructose-6-phosphate 1-phosphotransferase from cucumber seeds. *Plant Cell Physiology* **27**: 1285-1295.

Botha, FC, De vries, C, Small, JGC (1988). Immunological characterisation of the pyrophosphate dependent fructose-6-phosphate phosphate phosphotransferase. *Plant Cell Physiology* **27**: 1285-1295.

Botha, FC, Whittaker, A, Vorster, DJ and Black, KG (1996). Sucrose accumulation rate, carbon partitioning and expression of key regulatory enzyme activities in sugarcane stem tissue *In*: Wilson JR, Hogarth, DM, Campbell, JA and Garside, AL (eds.) Sugarcane: Research Towards Efficient and Sustainable Production, CSIRO Division of Tropical Crops and Pastures, Brisbane, pp. 98-101.

Bremer, G (1961). Problems in breeding and cytology of sugarcane. I: A short history of sugarcane breeding. The original forms of *Saccharum. Euphytica* **10**: 59-78.

Brenner, SE (1998). Practical database searching *In*: Brenner, S, Patterson, M, Handel, M (eds.) Trends Guide to Bioinformatics, Elsevier Science pp. 9-14.

Brown, TA (1989). Genetics. A molecular approach, Von Nostrand Reinhold (International) Co. Ltd., Hong Kong, pp. 3-50.

Burner, DM and Legendre, BL (1994). Cytogenetic and fertility characteristics of elite sugarcane clones. *Sugar Cane* **1**: 6-10.

Butler, ET and Chamberlain, MJ (1982). Bacteriophage SP6-specific RNA polymerase. I: Isolation and characterisation of the enzyme. *Journal of Biological Chemistry* **257**: 5772.

Carlisle, SM, Blakeley, SD, Hemmingsen, SM, Trevanion, SJ, Hiyoshi, T, Kruger, NJ, Dennis, DT (1990). Pyrophosphate-dependent phosphofructokinase: conversion of protein sequences between α- and β- subunits and with the ATP-dependent phosphofructokinase. *Journal of Biological Chemistry* **265**: 18366-18371.

Carnal, NW and Black, CC (1979). Pyrophosphate-dependent 6-phosphofructokinase, a new glycolytic enzyme in pineapple leaves. *Biochemical and Biophysical Research Communications* **86** (1): 20-26.

Carnal, NW and Black, CC (1983). Phosphofructokinase activities in photosynthetic organisms. The occurrence of pyrophosphate-dependent 6-phosphofructokinase in plants and algae. *Plant physiology* **71**: 150-155.

Charlieu, J-P (1994). PCR as a technique used daily in molecular biology *In* Griffin, H and Griffin, A (eds.) PCR technology, current innovations, CRC Press, Inc., USA, pp.1-4.

Cheng, H-F and Tao, M (1990). Differential proteolysis of the subunits of pyrophosphate-dependent-6-phosphofructo-1-phosphotransferase. *Journal of Biological Chemistry* **265**: 2173-2177.

Chernys, JT and Zeevaart, TAD (2000). Characterisation of the 9-cis-epoxycarotenoid dioxygenase gene family and the regulation of abscisic acid biosynthesis in Avocado. *Plant Physiology* **124**: 343-353.

Chouery, PS, Taliercio, EW, Carlson, SJ and Ruan, Y-L (1998). Genetic evidence that the two isozymes of sucrose synthase present in developing maize endosperm are critical, one for cell wall integrity and the other for starch biosynthesis. *Molecular and General Genetics* **259**: 88-96.

Cseke, C, Weeden, NF, Buchanan, BB, Uyeda, K (1982). A special fructose bisphosphate functions as a cytoplasmic regulatory metabolite in green leaves. *Proceedings of the National Academy of Science*, USA **79**: 4322-4326.

Dancer, JE and ap Rees, T (1989). Relationship between pyrophosphate: fructose-6-phosphate-1-transferase, sucrose breakdown and respiration. *Journal of Plant Physiology* **135**: 197-206.

Daniels, J and Roach, BT (1987). Taxonomy and evolution. *In*: Heinz, DJ (ed.) Sugarcane improvement through breeding. Elsevier, Amsterdam, pp. 7-84.

Da Silva, J. A. G. and Sobral, B. W. S (1995). Genetics of polyploids. *In*: Sobral, B. W. S. (Ed.), The impact of plant molecular genetics. Birkhauser, Boston, pp. 3-30.

Davanloo, P, Rosenberg, AH, Dunn, JJ, Studier, FW (1984). Cloning and expression of the gene for bacteriophage T7 RNA polymerase. *Proceedings of the National Academy of Science* **81**: 2035.

Dean, C and Schmidt, R (1995). Plant Genomes: A current molecular description, *Annual Review of Plant Physiology and Plant Molecular Biology* **46**: 395-418.

Dedonar, A, Rethy, R, Fredericq, H, Montagu, MV, Krebbers, E (1993). Arabidopsis rbcS genes are differentially regulated by light. *Plant Physiology* **101**: 801-808.

Dellaporta, SL, Wood, J and Hicks, JB (1983). A plant mini preparation: Version II. *Plant Molecular Biology Reporter* **1**: 19-21.

Denhardt, DT (1966). A membrane-filter technique for the detection of complimentary DNA. *Biochemical and Biophysics Research Communication*. **23**: 641-646.

Dennis, DT and Greyson, MF (1987). Fructose 6-phosphate metabolism in plants. *Physiology Plantarum* **69**: 395-404.

Dennis, ES, Gerlach, WL, Pryor, AJ, Bennetzen, JL, Inglis, A, Llewellyn, D, Sacha, MM, Ferl, RJ and Peacock, WJ (1984). Molecular analysis of the alcohol dehydrogenase (*Adh*1) gene of maize. *Nucleic Acids Research* **12**: 3983-4000.

Dennis, ES, Sachs, MM, Gerlach, WL, Finnegan, EJ and Peacock, WJ (1985). Molecular analysis of the maize alcohol dehydrogenase 2 (Adh2) gene of maize. *Nucleic Acids Research* **13**: 727-743.

D' Hont, A, Grivet, L, Feldmann, SR, Berding, N and Glaszmann, JC (1996). Characterization of the double genome structure of modern sugarcane cultivars (*Saccharum* spp.) by molecular cytogenetics. *Molecular and General Genetics* **250**: 405-413.

Don, RH, Cox, PT, Wainright, BJ, Baker, K and Mattick, JS (1991). 'Touchdown' PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Research* **19**: 4008.

Doolite, WF and Sapienza, C (1980). Selfish genes, the phenotype paradigm, and genome evolution. *Nature* **284**: 601-603.

Duff, SMG, Moorhead, GBG, Lefebvre, DD and Plaxton, WS (1989). Phosphate starvation inducible 'bypasses' of adenylate and phosphate dependent glycolytic enzymes in *Brassica nigra* suspension cells. *Plant Physiology* **90**: 1275-1278.

Dvorak, J and Dubcovsky, J (1996). Genome analysis of polyploid species employing variation in repeated nucleotide sequences *In* Jauhar, PP (ed.) <u>Methods of genome analysis in plants</u>, CRC Press, USA, pp.133-145.

Erkillä, MJ, Leah, R, Ahokas, H and Cameron-Mills, V (1998). Allele dependent Barley grain B-amylase activity. *Plant Physiology* **117**: 679-685.

Erlich, HA (1989). <u>PCR technology, principles and applications for DNA amplification</u>, Stockton Press, USA, pp. 1-16.

Fincham, JRS (1983). <u>Genetics</u>, John Wright and Sons, Ltd., Great Britain, pp. 124.

Finnegan, PM, Whelan, J, Millar, AH, Zhang, Q, Smith, MK, Wiskich, JT and Day, DA (1997). Differential expression of the multigene family encoding the soybean mitochondrial alternative oxidase. *Plant Physiology* **114**: 455-466.

Flavell, RB (1980). The molecular characterisation and organisation of plant chromosomal DNA sequences. *Annual Review of Plant Physiology* **31**: 569-576.

Flavell, RB (1986). Repetitive DNA and chromosomal evolution in plants. *Philosophical Transactions of the Royal Society of London (Biological)* **312**: 227-242.

Flavell RB and Moore, G (1996). Plant Genome constituents and their organisation *In*: Foster, GD and Twell, D (eds.) <u>Plant Gene Isolation, Principles and Practices</u>, John Wiley and Sons, UK, pp. 1-20.

Futuyama, DJ (1986). <u>Evolutionary biology</u> 2nd ed., Sinauer Associates Inc. Publishers, Sunderland, USA, pp. 45–50.

Gengenbach, BG, VanDee, KL, Egli, E, Hildebrandt, KM, Yun, SJ, Lutz, SM, Marshall, LC, Wyse, DL, Somers, DA (1999). Genetic relationships of alleles for tolerance to sethoxydim herbicide in maize. *Crop Science* **39**: 812-818.

Gottlieb, LD and Ford, VS (1997). A recently silenced duplicate PgiC locus in Clarkia. *Molecular Biology and Evolution* **14**: 125-132.

Gottlob-McHugh, SG, Sangwan, RS, Blakeley, SD, Vanlerberghe, GC, Ko, K, Turpin, DH, Plaxton, WC, Miki, BL and Dennis, DT (1992). Normal growth of transgenic tobacco plants in the absence of cytosolic pyruvate kinase. *Plant Physiology* **100**: 820-825.

Grandbastien, M-A (1992). Retroelements in higher plants. *Trends in Genetics* **8**: 103-108.

Grant, V (1981). Plant Speciation, Columbia University Press, New York, pp. 283- 487.

Grivet, L, D' Hont, A, Roques, D, Feldmann, P, Launaud, C and Glaszmann, JC (1996). RFLP mapping in cultivated sugarcane (*Saccharum* spp.): Genome organisation in a highly polyploid and aneuploid interspecific hybrid. *Genetics* **142**: 987-1000.

Groenewald, S (1998). Promoters for sugarcane transformation: isolation of specific sequences and evaluation of *rol C*. PhD Thesis, University of Natal, Pietermaritzburg.

Guo, M, Davis, D, Birchler, JA (1996). Dosage effects on gene expression in maize ploidy series. *Genetics* **142**: 1349-1355.

Hahn, WJ and Grifco, FT (1996). Molecular markers in plant conservation genetics In Sobral, WS (ed.) The impact of plant molecular genetics, Birkhauser Boston, USA, pp. 113-115.

Hajirezaei, M, Sonnewald, S, Viola, R, Carlisle, S, Dennis, DT and Stitt, M (1994). Transgenic potato plants with strongly decreased expression of pyrophosphate

phosphotransferase show no visible phenotype and only minor changes in metabolic fluxes in their tubers. *Planta* **192**: 16-30.

Harris, N, Foster, JM, Kumar, A, Davies, HV, Gebhardt, C, and Wray, JL (2000). Two cDNAs representing alleles of the nitrate reductase gene of potato (*Solanum tuberosum* L. cv. Desiree): sequence analysis, genomic organisation and expression. *Journal of Experimental Botany* **51**: 1017-1026.

Harvey, A (1998). A study of allelic variation in sugarcane using the PFP-β region as a model. BTech Thesis, Natal Technikon, Durban.

Harvey, A, Huckett, BH and Botha, FC (1999). Allelic variation in the gene encoding pyrophosphate dependent phosphofructokinase (PFP) in sugarcane. *Proceedings of the South African Sugar Technologists Association* **73**: 134-136.

Heslop-Harrison, JS and Schwarzacher, T (1996). Genomic southern and in situ hybridisation for plant genome analysis *In* Jauhar, PP (ed.) Methods of genome analysis in plants, CRC Press, USA, pp.163-179.

Holland, PM, Abramson, RD, Watson, R and Gelfand, DH (1991). Detection of specific polymerase chain reaction product by utilising the %'-3' exonuclease activity of *Thermus aquaticus* DNA polymerase. *Proceedings of the National Academy of Sciences, USA* **88**: 7276-7280.

Honeycutt, RJ, Sobral, BWS, Keim, P and Irvine, JE (1992). A rapid DNA extraction method for sugarcane and its relatives. *Plant Molecular Biology Reporter* **10**: 66-72.

Hudspeth, RL and Grula, JW (1989). Structure and expression of the maize gene encoding the phosphoenolpyruvate carboxylase isozyme involved in C4 photosynthesis. *Plant Molecular Biology* **12**: 579-589.

Hulbert, SH, Ritcher, TE, Axtell, JD, Bennetzen, JL (1990). Genetic mapping and characterisation of sorghum and related crops by means of maize DNA probes. *Proceedings of the National Academy of Science, USA* **87**:4251-55.

Innis, MA, Myambo, KB, Gelfand, DH and Brow, MAD (1988). DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. *Proceedings of the National Academy of Sciences, USA* **85**: 9436-9440.

Intapruk, C, Yamamoto, K, Fujiyama, K, Takano, M and Shinmyo, A (1993). Cloning of cDNAs encoding two peroxidases of A*rabidopsis thaliana* and their organ specific expression. *Journal of Fermentation and Bioengineering* **75** (3): 166-172.

Jones, N, Oughman, H and Thomas, H (1997). Markers and mapping: we are all geneticists now. *New Phytologist* **137**: 165-177.

Kahl, G (ed.) (1995). Dictionary of Gene Technology, VCH Verlagsgessellschaft mbH, Weinham, Germany, 550p.

Kawaoka, A, Sato, S, Nakahara, Ko, Matsushima, N, Okada, N, Sekine, M, Shinmyo, A and Takano, M (1992). Expression and promoter activity of genes for isozymes of horseradish peroxidase. *Plant Cell Physiology* **33** (8): 1143-1150.

Kim WY, Kim CY, Cheong NE, Choi YO, Lee KO, Lee SH, Park JB, Nakano A, Bahk JD, Cho MJ and Lee, SY (1999). Characterisation of two fungal-elicitor-induced rice cDNAs encoding functional homologoues of the rab-specific GDP-dissociation inhibitor. *Planta* **210**: 143-149.

Kiss, T, Abel, S and Solymosy, F (1989). A plant pseudogene for U1 RNA. *Plant Molecular Biology* **12**: 709-711.

Kiss, T and Solymosy, F (1993). Molecular analysis of a U3 RNA gene locus in tomato: transcription signals, the coding region, expression in transgenic tobacco plants and tandemly repeated pseudogenes. *Nucleic Acids Research* **18**: 1941-1949.

Kombrink, E, Kruger, NJ and Beevers, H (1984). Kinetic properties of pyrophosphate: fructose-6-phosphate phosphotransferase from germinating castor bean endosperm. *Plant Physiology* **74**: 395-401.

Kowalczyk, S, Januszewska, B, Cymerska, E and Maslowski, P (1984). The occurrence of inorganic pyrophosphate: D-fructose-6-phosphate 1-phosphotransferase in higher plants. I. Initial characterisation of partially purified enzyme from *Sanseviera trifasciata* leaves. *Physiologia Plantarum* **60**: 31-37.

Kruger, NJ and Dennis, DT (1987). Molecular properties of pyrophosphate: fructose-6-phosphate phosphotransferase from potato tuber. *Biochemistry and Biophysics* **256** (1): 273-279.

Law, CN, Suarez, E, Miller, TE and Worland, AJ (1998). The influence of the group 1 chromosomes of wheat on ear emergence times and their involvement with vernalization and day length. *Heredity* **80**: 83-91.

Lew, TL (1987) New varieties. *In*: Heinz, DJ (ed.) <u>Sugarcane improvement through breeding</u>, Elsevier, Amsterdam, pp. 559-594.

Lo, CC and Shong, S (1968). Collecting wild cane in Taiwan. *Proceedings of The International Society of Sugar Cane Technologists (ISSCT)*, **13**: 1047-1055.

Longhurst, T, Lee, E, Hinde, R, Brady, C and Speirs, J (1994). Structure of the tomato *Adh2* gene and *Adh2* pseudogenes, and a study of *Adh* 2 gene expression in fruit. *Plant Molecular Biology* **26**: 1073-1084.

Lucas, H, Moore, G, Murphy, G and Flavell, RB (1992). Inverted repeats in the long terminal repeats of the wheat retrotransposon Wis 2-1A. *Molecular Biology and Evolution* **1**: 716-728.

Lu, YH, D' Hont, A, Paulet, F, Grivet, L, Arnaud, M and Glaszmann, JC (1994). Molecular diversity and genome structure in modern sugarcane varieties. *Euphytica* **78**: 217-226.

Macdonald, FD and Preiss, J (1986). The subcellular location and characteristics of pyrophosphate-fructose-6-phosphate 1-phosphotransferase from suspension-cultures cells of soybean. *Planta* **167**:240-245.

Marshall, CR, Raff, EC, Raff, RA. (1994). Dollo's law and the death and resurrection of genes. *Proceedings of the National Academy of Science, USA* **91**: 12283-12287.

Mertens, E (1991). Pyrophosphate-dependent phosphofructokinase. An anaerobic glycolytic enzyme? *Federation of Biochemical Societies Letters* **285**: 1-5.

Moore, D (1999). Current protocols in molecular biology, John Wiley and Sons, Inc., USA, pp. 2.0.1-2.0.3.

Moore, G, Abbo, S, Cheung, W (1993). Key features of cereal genome organisation as revealed by the use of cytosine methylation sensitive restriction enzymes. *Genomics* **15**: 472-482.

Moore, PH (1995). Temporal and spatial regulation of sucrose accumulation in the sugarcane stem. *Australian Journal of Plant Physiology* **22**: 661-79.

Muller-Rober, B, Sonnewald, U and Willmitzer, L. (1992). Inhibition of ADP-glucose pyrophosphorylase in transgenic potatoes leads to sugar storing tubers and influences tuber formation and expression of tuber storage. *European Molecular Biology Organisation Journal* **11**: 1229-38.

Nagl, W (1976). Nuclear Organisation. *Annual Review of Plant Physiology* **27**: 39-69.

Nielson,TH (1994). Pyrophosphate:fructose-6-phosphate 1-phosphotransferase from barley seedlings. Isolation subunit composition and kinetic characterisation. *Physiologia Plantarum* **92**: 311-321.

Nilan, RA (1981). Induced gene and chromosome mutants. *In:* The Manipulation of Genetic Systems in Plant Breeding. A Royal Society Discussion, University Press, Cambridge, Great Britain, pp. 457-466.

Nowakowska, J (1988). Gene expression and oxalate oxidase activity of two germin isoforms induced by stress. *Acta Physiologiae Plantarum* **20**(1): 19-33.

'O Brien, WE, Bowien, S and Wood, HG (1975). Isolation and characterisation of a pyrophosphate dependent phosphofructokinase from *Propionibacterium shermanii. Journal of Biological Chemistry* **250**: 8690-8695.

Omann, FN, Beaulieu and Tyson, H (1993). cDNA sequence and tissue-specific expression of an anionic flax peroxidase. *Genome* **37**: 137-147.

Paterson, AH (1993). Molecular analysis of crop genomes. *Proceedings of the 10th Australian Plant Breeding Conferences*, 18-23 April, Gold Coast, pp. 233-236.

Paterson, AH (1996). Genome mapping in plants, RG Landes Company, Texas, USA, pp. 1-19.

Pfitzner, AJP, Beilman, A, Goodman, HM and Pitzner, UM (1991). Molecular analysis of two PR-1 pseudogenes from tobacco. *Plant Molecular Biology* **16**: 129-139.

Plaxton WC (1996). The organisation and regulation of plant glycolysis. *Annual Review of Plant Physiology and Plant Molecular Biology* **47**: 185-214.

Powell, W, Thomas, WTB, Caliggari, PDS and Jinks, JL (1985). The effects of major genes on quantitatively varying characters in barley 1. The GP *ert* locus. *Heredity* **54**: 343-348.

Ra, SR and Saris, PEJ (1995). Characterisation of prokaryotic mRNAs by RT-PCR, *Bio Techniques* **18**:797-795.

Raven, PH, Evert, RF and Eichhorn, SE (1986). Biology of plants, 4th edition, Worth Publishers Inc., New York, pp. 145-147.

Reeves, RE, South, DJ, Blytt, HJ and Warren, LG (1974). Pyrophosphate: fructose-6-phosphate-1-phosphotransferase. An enzyme with the glycolytic function of 6-phosphofructokinase. *Journal of Biological Chemistry* **249**: 7737-7741.

Roach, BT and Daniels, J (1987). A review of the origin and improvement of sugarcane. *In*: Copersucar International Sugarcane Workshop, Copersucar Technology Center, Brasil pp. 1-31.

Roose, ML and Gottlieb, LD (1980). Biochemical properties and level of expression of alcohol dehydrogenases in the allotetraploid plant *Tragopogan miscellus* and its diploid progenitors. *Biochemical Genetics* **18**: 1065-1085.

Rose, TM, Schultz, ER, Heinikoff, JG, Pietrokovski, S, McCallum, CM and Heinekoff, S (1998). Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. *Nucleic Acids Research* **26** (7): 1627-1635.

Rozario, C, Smyth, MW and Müller, M (1995). Primary sequence of a putative pyrophosphate linked phosphofructokinase gene of *Giardia lamblia*. *Biochimica et Biophysica Acta* **1260**: 218-222.

Sabularse DC and Anderson, RL (1981). D-fructose 2,6-bisphosphate: a naturally occuring activator for inorganic pyrophosphate:D-fructose-6-phosphate 1-phosphotransferase in plants. *Biochemical and Biophysical Research Communications* **103** (3): 848-855.

Saiki, RK, Scharf, S, Faloona, F, Mullis, KB, Horn, GT, Erlich, HA and Arnheim, N (1985). Enzymatic amplification of β-globin sequences and restriction site analysis for diagnosis of sickle cell anaemia. *Science* **239**: 487-491.

Sambrook, J, Fritsch, EF, Maniatis, T (1989). Molecular cloning: a laboratory manual, 2nd edition, Cold Spring Harbor Laboratory Press, USA.

Saponaro, C, Pogna, NE, Castagna, R, Pasquini, M, Cacciatori, P and Redaelli, R (1995). Allelic varation at the Gli-A1^m, Gli-A2^m and Glu-A1^m loci and breadmaking quality in diploid wheat *Triticum monococcum*. *Genetical Research, Camb.* **66**: 127-137.

Simmonds, NW (1979). Principles of Crop Improvement, Longman Group Ltd., London. pp. 1-26.

Sloop, KW and Rhodes, SJ (2000). Combined hot start and touchdown PCR protocols optimise amplification of difficult targets. *Biochemica* **1**: 35-36.

Smith, JM (1989). Evolutionary Genetics, Oxford University Press, New York, USA, pp. 53-54, 202-215.

Soltis, DE and Soltis, PS (1993). Molecular data and the dynamic nature of polyploidy. *Critical Reviews in Plant Sciences* **12** (3): 243-273.

Soltis, PS, Doyle, JJ and Soltis, DE (1992). Molecular data and polyploid evolution in plants. *In*: Soltis, PS, Soltis, DE and Doyle, JJ (eds.) Molecular systematics of plants, Chapman and Hall, New York, pp. 177-210.

Sonnewald, U (1992). Expression of *E.coli* inorganic pyrophosphatase in transgenic plants alters photoassimilate partitioning. *Plant Journal* **2**: 571-581.

Souaze, F, Ntodou-Thome, A, Tran, CY, Rostene, W and Forqez, P (1996). Quantitative RT-PCR: limits and accuracy, *Bio Techniques* **21**: 280-285.

Sowokinos, JR, Thomas, C, Burrell, M (1997). Pyrophosphorylases in potato. V. Allelic polymorphism of UDP-Glucose pyrophosphorylase in potato cultivars and its association with tuber resistance to sweetening in the cold. *Plant Physiology* **113**: 511-517.

Sreenivasan, TV Ahloowalia, BS Heinz, DJ (1987). Cytogenetics. *In*: Heinz, DJ (ed.) Sugarcane improvement through breeding, Elsevier Publishers, Amsterdam, pp. 211-253.

Stark, DM, Timmerman, KP, Barry, GF, Preiss, J and Kishore, GM (1992). Regulation of the amount of starch in plant tissues by ADP glucose pyrophosphorylase. *Science*. **258**: 287-92.

Stevenson, GC (1965). Genetics and breeding of sugarcane, Longmans, Green and Co. Ltd., London, pp. 1-38.

Stitt, M (1989). Product inhibition of potato tuber pyrophosphate: fructose-6-phosphate phosphotransferase by phosphate and pyrophosphate. *Plant Physiology* **89**: 628-633.

Stitt, M (1990). Fructose-2,6-bisphosphate as a regulatory molecule in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* **41**: 153-185.

Suh, MC, Cho, HS, Kim, YS, Liu, JR and Lee, H-S (1998). Multiple genes encoding serine/ threonine protein phosphatases and their differential expression in *Nicotiana tabacum*. *Plant Molecular Biology* **36**: 315-322.

Suzuki, G, Watanabe, M, Kai, N, Matsuda, N, Toriyama, K, Takayama, S, Isogi, A and Hinata, K (1997). Three members of the S multigene family are linked to the S locus of *Brassica. Molecular and General Genetics* **256**: 257-264.

Swiderski, MR, Zaborowska, Z and Legocki, AB (2000). Identification of new nodulin cDNAs from yellow lupine by differential display. *Plant Science* **151** (1): 75-83.

Tanksley, SD and Pichersky, E (1988). Organisation and evolution of sequences in the plant nuclear genome. *In*: Gottlieb, LD and Jain, SK (eds.) Plant Evolutionary Biology, London, Chapman and Hall Ltd., Great Britain, pp. 55-79.

Theodorou, ME, Cornel, FA, Duff, SMG, Plaxton, WC (1992). Phosphate starvation inducible synthesis of the alpha-subunit of the pyrophosphate-dependent phosphofructokinase in black mustard suspension cells. *Journal of Biological Chemistry* **267**: 21901-21905.

Theodorou, ME, and Plaxton, WC (1993). Metabolic adaptations of plant respiration to nutritional phosphate deprivation, *Plant Physiology* **101**: 339-344.

Thoiron, S and Briat, JF (1999). Differential expression of maize sugar responsive genes in response to iron deficiency. *Plant Physiology Biochemistry* **37**: 759-766.

Tibbot, BK and Skadsen, RW (1996). Molecular cloning and characterisation of a gibberellin-inducible, putative α–glucosidase gene from barley. *Plant Molecular Biology* **30**: 229-241.

Tindall, KR and Kunkel TA (1988). Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Biochemistry* **27**: 6008-6013.

Todd, JF, Blakeley, SD and Dennis, DT (1995). Structure of the genes encoding the α- and β-subunits of castor pyrophosphate-dependent phosphofructokinase. *Gene* **152**: 181-186.

Treager, JW and Roberts, LM (1992). The lectin gene family of *Ricinus communis*: Cloning of a functional ricin gene and three lectin pseudogenes. *Plant Molecular Biology* **18**: 515-525.

van der Have (1979). Plant Breeding Perspectives, Sneep, J and Hendriksen, AJT (eds.), Centre for Agricultual Publishing and Documentation, Wageningen, Netherlands, pp. 322-330.

VanHal, NLW, Vorst, O, vanHowelingen, AMML, KOK, EJ, Peijnenburg, A, Ahoroni, A, vanTunen, AJ and Keijer, J (2000). The application of DNA microarrays in gene expression analysis. *Journal of Biotechnology* **78** (3): 271-280.

Van Schafingten, E, Lederer, B, Bartrons, R and Hers, H-G (1982). A kinetic study of pyrophosphate: fructose-6-phosphate phosphotransferase from potato tubers. Application to a microassay of fructose 2,6-bisphosphate. *European Journal of Biochemistry* **129**: 191-195.

Vaux, P, Guerineau, F, Waugh, R and Brown, JWS (1992). Characterisation and expression of U1 snRNA gene from potato. *Plant Molecular Biology* **19**: 959-971.

Walbot, V and Cullis, CA (1985). Rapid genomic change in higher plants. *Annual Review of Plant Physiology* **36**: 367-396.

Wendel, JF (2000). Genome evolution in polyploids, *Plant Molecular Biology* **42**: 225-249.

Whittaker, A (1997). Pyrophosphate dependent phosphofructokinase (PFP) activity and other aspects of sucrose metabolism in sugarcane internodal tissues. PhD Thesis, University of Natal, Durban.

Whittaker, A and Botha, FC (1997). Carbon partitioning during sucrose accumulation in sugarcane internodal tissue. *Plant Physiology* **115**: 1651-1659.

Whittaker, A and Botha, FC (1999). Pyrophosphate:D-fructose-6-phosphate 1-phosphotransferase activity patterns in relation to sucrose storage across sugarcane varieties. *Physiologia Plantarum* **107**: 379-386.

Wong, JH, Kang, T and Buchanan, BB (1988). A novel pyrophosphate fructose-6-phosphate 1–phosphotransferase from carrot roots. Relation to PFK from the same source. *Federation of European Biochemical Societies Letters* **238** (2): 405-410.

Wong, JH, Kiss, F, Wu, M-X, Buchanan, BB (1990). Pyrophosphate fructose-6-P 1-phosphotransferase from tomato fruit. *Plant Physiology* **94**: 499-506.

Worrell, AC, Bruneau, J-M, Summerfelt, K, Boersig, M and Voelker, TA (1991). Expression of a maize sucrose phosphate synthase in tomato alters leaf carbohydrate partitioning. *Plant Cell* **3**: 1121- 1130.

Wu, K, Rooney, MF, Ferl, RJ (1997). The Arabidopsis 14-3-3 multigene family. *Plant Physiology* **114**: 1421-1431.

Wu, M-X, Smyth, DA and Black, CC (1983). Fructose 2,6-bisphosphate and the regulation of pyrophosphate-dependent phosphofructokinase activity in germinating pea seeds. *Plant Physiology* **73**: 188-191.

Xiang, CC and Chen, YD (2000). cDNA microarray technology and its applications. *Biotechnology Advances* **18** (1): 35-46.

Xu, J, Avigne, WT, McCarthy, DR, Koch, KE (1996). A similar dichotomy of sugar modulation and developmental expression affects both paths of sucrose metabolism: evidence from a maize invertase gene family. *The Plant Cell* **8**: 1209-1220.

Yan, H-F and Tao, M (1984). Multiple forms of pyrophosphate: D-fructose 6-phosphate 1-phosphotransferase from wheat seedlings. *Journal of Biological Chemistry* **259**: 5087-5092.

Yanisch-Perron, C, Viera, J and Messing, J (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**: 103.

CURRICULUM VITAE

Higher Education qualifications:

1997: Completed Bachelor of Science degree at the University of Natal

      Majors: Cell Biology

                            Environmental Biology

1998: Obtained a Bachelor of Science (Honours) degree at the University of Natal

      Courses: Current topics in Molecular biology

                            Ecotoxicology

                            Plant Biotechnology

                            An internship programme at SASEX (Title of Thesis: The development of reliable PCR-based markers linked to smut disease and eldana stalk borer resistance in sugarcane.)

      Major Project title: The evaluation of computer games as a viable education tool.

1999: Enrolled for a Master of Science degree at the University of Stellenbosch under the supervision of Dr BI Huckett and Prof. FC Botha.

2001: Microarray Technician at the University of Cape Town, Molecular and Cellular Biology Department.