

# ALLELIC VARIATION IN THE GENE ENCODING PYROPHOSPHATE DEPENDENT PHOSPHOFRUCTOKINASE (PFP) IN SUGARCANE

AR HARVEY<sup>1</sup>, BI HUCKETT<sup>1,2</sup> AND FC BOTHA<sup>1,2</sup>

<sup>1</sup>SASA Experiment Station, Private Bag X02, Mount Edgecombe, 4300, South Africa

<sup>2</sup>Institute for Plant Biotechnology, University of Stellenbosch, Private Bag X1, Matieland, 7602, South Africa

## Abstract

The enzyme pyrophosphate dependent phosphofructokinase (PFP), considered to play an important role in sucrose metabolism in plants, comprises two protein subunits encoded by separate genes. Part of the  $\beta$ -subunit gene was used as a model in investigating allelic variation in the polyploid genetic environment of sugarcane. The gene fragment was amplified by the polymerase chain reaction (PCR) from various total genomic DNA samples and a number of individual DNA molecules cloned randomly from each source. Sequence analysis of clones has demonstrated that two strikingly different alleles are present in ancestral *Saccharum* species as well as modern hybrid varieties.

## Introduction

Sugarcane is a highly heterozygous and genetically cumbersome crop (Aljanabi, 1998). This is at least in part because it is highly polyploid, having a suspected base chromosome number of 10 and between 110 and 130 chromosomes in the 2n state, contributed by two major ancestral species, *Saccharum officinarum* and *Saccharum spontaneum* (D'Hont *et al.*, 1996). The number of alleles of important genes and the degree to which those alleles vary in the complex genetic environment of sugarcane is of considerable interest because of the potential for exploitation of genetic variation through breeding, and manipulation of gene function via genetic engineering. In addition, such information might throw light on the general speciation and evolution of polyploids, about which little is known (Soltis *et al.*, 1992). The working hypothesis for this study was that allelic variation can best be established by targeting a single copy gene (i.e. a gene with one copy per haploid genome), the argument being that such a gene should be present in multiple allelic copies in sugarcane. The gene encoding the enzyme pyrophosphate dependent phosphofructokinase (PFP) is such a gene (Blakeley *et al.*, 1992). PFP is considered to play an important role in sucrose metabolism in plants (Dancer and ap Rees, 1989). Part of the gene encoding the  $\beta$ -subunit of the enzyme was adopted as the study model. Using genomic DNA from different sugarcane sources as template, the gene fragment was amplified by the polymerase chain reaction (PCR) with primers designed to complement conserved regions of the gene. Amplified fragments were ligated into plasmid vectors and individually cloned in *E. coli* for comparative sequencing analysis.

## Materials and Methods

### Primers

Forward primer B1 was designed from a region of high homology identified by comparison of known PFP gene sequences from castor bean and potato while reverse primer B8 was designed from a sugarcane PFP sequence (Sarita Groenewald; PhD thesis, University of Natal, 1998, and personal communication):

B1 (20bp) (5' to 3') ATI-GAT-TTC-ATI-CCI-GAG-GT

B8 (21bp) (5' to 3') CCA-TCA-GTG-ATG-TCA-ATG-CTG

The predicted product of B1/B8 amplification is a region of the PFP gene representing exons 10, 11 and 12 and their flanking introns, approximately 1200bp in size in all.

### Gene amplification, cloning and sequencing

Genomic DNA was extracted from *Saccharum officinarum* (Black Cheribon), *Saccharum spontaneum* (Coimbatore) and a commercial variety (N21) as described by Harvey and Hockett (1998). PCR reactions contained, in a final volume of 50 $\mu$ l: 10mM Tris-HCl, 50mM KCl, 1,5mM MgCl<sub>2</sub>, 50 $\mu$ g gelatine, 0,33mM deoxynucleoside triphosphates (dNTPs), 0,4 $\mu$ M primers B1 and B8, 1U Taq DNA polymerase (Boehringer Mannheim) and 75ng genomic DNA, at a final pH of 8,3. Thermal treatment was 94°C for 1,5 min (1 cycle), 94°C for 45 sec, 45°C for 30 sec and 72°C for 2 min (10 cycles) and 94°C for 30 sec, 40°C for 30 sec and 72°C for 2 min (30 cycles) followed by 72°C for 2 min then 35°C for 5 min. PCR products were purified by electrophoresis through 1% (w/v) agarose gels in TBE buffer (45mM tris-borate, 1mM EDTA, pH 8,0) followed by excision of bands visualised by conventional ethidium bromide staining and gel removal by spun column (GenElute, Supelco). PCR fragments were ligated into the pGEM<sup>®</sup>-T Easy plasmid vector (Promega) and cloned in *E. coli* strain JM109 according to manufacturer's instructions. Cloned plasmids were extracted from bacterial cells and purified using column separation kits (Nucleobond AX20, Macherey-Nagel) while estimations of quality and quantity were made both spectrophotometrically at 260nm and by visual observation following gel electrophoresis. Plasmid inserts were used with either B1 or B8 primer in cycle sequencing reactions according to standard dye terminator chemistry methodologies (Big Dye Kit, Perkin Elmer). Reaction product analysis was accomplished by capillary electrophoresis using a Perkin Elmer ABI Prism 310 Genetic Analyzer.

*Sequence analysis*

Sequences were edited manually to remove plasmid vector components. PFP homology was measured by nucleotide comparison of individual sequences to genes registered in international databases using the Basic Local Alignment Search Tool (BLAST) algorithm (Altschul *et al.*, 1990). Sequence overlaps, alignments and other data manipulations were accomplished using Perkin Elmer software Sequence Navigator 1.0.1 and/or the software programme DNASIS for Windows v2.1.

**Results and Discussion**

Two PCR amplification products of approximately 1250bp and 1100bp in size were obtained consistently from the genomic DNAs of all three sources: Black Cheribon, Coimbatore and N21. The fragments were designated PFP.B1/B8<sub>1250</sub> and PFP.B1/B8<sub>1100</sub> respectively. In the case of Coimbatore, a third product of approximately 950bp was sometimes obtained. Only PFP.B1/B8<sub>1250</sub> and PFP.B1/B8<sub>1100</sub> were cloned and sequenced in this study. B1 generated sequences were used for BLAST homology searches. All scores obtained were above the significance threshold of 80 and ranged from 119 to 522 (mean = 411), indicating that all gene fragments obtained by PCR were identifiable as PFP. Where possible, forward and reverse sequences were integrated using Contig Manager, an overlap function of the software programme DNASIS, to provide full length frag-

ment sequences. Results of cloning, sequencing and BLAST homology searches of fragments PFP.B1/B8<sub>1250</sub> and PFP.B1/B8<sub>1100</sub> from all three germplasm types are summarised in Table 1.

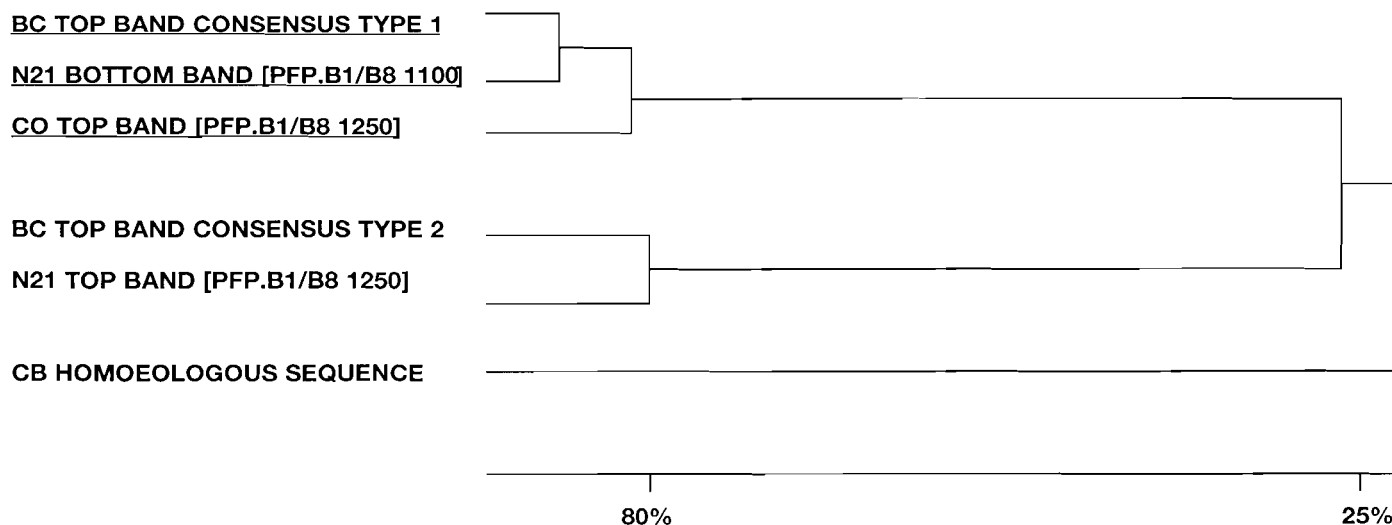
To investigate sequence differences between the two size species PFP.B1/B8<sub>1250</sub> and PFP.B1/B8<sub>1100</sub>, representative full length sequences obtained from commercial variety N21 (one sequence of each type, Table 1) were compared with each other in an initial analysis. Results showed that the size differences between the two fragments was due to sequence variations throughout their lengths rather than a single major insertion or deletion in one of the fragments. Homologous cDNA sequences of the castor bean PFP  $\beta$ -gene (exon regions 10, 11 and 12), assumed to be relatively conserved across species, were used to determine the positions of exon and intron regions in the sugarcane gene fragments. Exon regions of the two sugarcane variants showed 77-78% nucleotide homology while, as expected, intron regions showed greater variability with 48-59% nucleotide homology.

Having used N21 data to establish the existence of two variants of the gene, apparently associated with a size difference between the B1 and B8 primer sites, the data obtained for Black Cheribon and Coimbatore were analysed. For each of the germplasm sources, sequences clustered into one of two distinct sequence types, shown by consensus sequence comparisons to correspond closely to the two allelic variants

**Table 1. Sequencing results and PFP homology scores for cloned gene fragments obtained by PCR from genomic DNA of *Saccharum officinarum* (Black Cheribon), *Saccharum spontaneum* (Coimbatore) and the commercial variety N21.**

GERMPLASM SOURCE	FRAGMENT	CLONE NAME	B1 (FORWARD) GENERATED SEQUENCE LENGTH (bp)	B1-SEQUENCE BLAST SIGNIFICANCE SCORE FOR PFP	B8 (REVERSE) GENERATED SEQUENCE LENGTH (bp)	OVERLAP REGION (bp)	FULL, LENGTH SIZE (bp)	
BLACK CHERIBON	PFP.B1/B8 <sub>1250</sub>	BC.T#1	643	499	668	285	1026	
		BC.T#3	662	513	643	272	1033	
		BC.T#4	556	352	366	-	-	
		BC.T#5	650	322	670	66	1254	
		BC.T#6	370	277	606	-	-	
		BC.T#10	651	495	651	93	1209	
		BC.T#11	660	119	653	283	1030	
		BC.T#12	650	495	687	129	1208	
		BC.T#13	640	495	637	21	1256	
		PFP.B1/B8 <sub>1100</sub>	BC.B#17	530	522	392	-	-
			BC.B#18	532	522	432	-	-
			BC.B#21	475	504	390	-	-
			BC.B#24	485	331	653	91	1047
COIMBATORE	PFP.B1/B8 <sub>1250</sub>	CO.T#4	743	489	-	-	-	
		CO.T#7	522	250	532	24	1030	
		CO.T#8	461	197	465	-	-	
N21	PFP.B1/B8 <sub>1250</sub>	N21.T#10	777	495	715	273	1219	
		N21.B#9	735	522	686	395	1026	

BC, Black Cheribon; CO, Coimbatore; T, top gel fragment, 1250bp; B, bottom gel fragment, 1100bp. BLAST scores above 80 indicate significant homology to known PFP genes.



**Figure 1.** Homology tree derived from sugarcane PFP gene fragments and the equivalent sequence from castor bean. All contributing sequences represent full length B1/B8 PFP fragments. Consensus sequences types 1 and 2 were each based on three full length fragments and were obtained in Sequence Navigator. The tree was derived in DNASIS. BC, Black Cheribon; CO, Coimbatore; CB, castor bean (*Ricinus communis*). Scale shows relative homology values in %.

identified in N21. However, correlation between sequence type and size species of origin was not tight, suggesting that contamination of the PFP.B1/B8<sub>1250</sub> and PFP.B1/B8<sub>1100</sub> fragment populations had occurred at the point of gel band excision. This is not unlikely, as it is recognised that DNA fragments such as these, differing in size by only 150bp, may appear separate visually but not be resolved perfectly by agarose gel electrophoresis. Future preparative separation would be accomplished more precisely in polyacrylamide gels.

Relationships between the various sugarcane PFP gene fragment sequences and the homoeologous sequence from castor bean are illustrated in Figure 1. Two distinct allelic clusters are identifiable among the sugarcane sequences. It is notable that the degree to which they differ from each other is almost as great as the degree to which each differs from castor bean, a dicotyledonous plant not closely related to sugarcane. Although the analysis is based on limited numbers of cloned fragments and only single sequence runs, the result is sufficiently striking to suggest that the variations do reflect significant allelic differences in the genome. The high level of distinction and the fact that both gene forms occur in the ancestral germplasm of sugarcane suggest that the alleles diverged early in the evolution of the genus *Saccharum*.

Confirmation of this work and further investigations into PFP gene variants and their expression is being undertaken.

#### Acknowledgements

The authors thank Deborah Carson of the Biotechnology

department at the SASA Experiment Station and Dr Sarita Groenewald of the Institute for Plant Biotechnology, University of Stellenbosch, for practical advice, helpful discussions and appraisal of the manuscript.

#### REFERENCES

- Aljanabi, SM (1998). Genetics, phylogenetics and comparative genetics of *Saccharum L.*, a polysomic polyploid Poales: Andropogoneae. *Biotech Ann Rev* 4: 285-319.
- Altschul, SF, Gish, W, Miller, W, Myers, EW and Lipman, DJ (1990). Basic Local Alignment Search Tool. *J Mol Biol* 215: 403-410.
- Blakeley, SD, Crews, L, Todd, JF and Dennis, DT (1992). Expression of the genes for the  $\alpha$ - and  $\beta$ -subunits of pyrophosphate-dependent phosphofructokinase in germinating and developing seeds from *Ricinus communis*. *Plant Phys* 99: 1245-1250.
- Dancer, JE and ap Rees, T (1989). Relationship between pyrophosphate:fructose-6-phosphate 1- transferase, sucrose breakdown and respiration. *J Plant Phys* 135: 197-206.
- D'Hont, A, Grivet, L, Feldmann, P, Rao, S, Berding, N and Glaszmann, JC (1996). Characterisation of the double genome structure of modern sugarcane cultivars (*Saccharum* spp.) by molecular cytogenetics. *Mol Gen Genet* 250: 405-413.
- Harvey, M and Hockett, BI (1998). Preliminary genetic analysis of the sugarcane population AA40. *Proc S Afr Sugar Technol Ass* 72: 316-317.
- Soltis, PS, Doyle, JJ, Soltis, DE (1992). Molecular data and polyploid evolution in plants. In: Soltis, PS, Soltis, DE, Doyle, JJ (Eds) *Molecular Systematics of Plants*. New York: Chapman and Hall: 177-210.