

POLYMORPHISM IN THE PROMOTER REGION OF THE SUCROSE SYNTHASE-2 GENE OF *SACCHARUM* GENOTYPES

Sarah E. Lingle and John M. Dyer

USDA-ARS, Southern Regional Research Center, 1100 Robert E. Lee Blvd., New Orleans, LA 70124

ABSTRACT

Sucrose synthase (E.C. 2.4.1.13) is an important enzyme of sucrose metabolism in sugarcane (*Saccharum* spp. hybrids). The objectives of this study were to isolate and characterize the complete genomic sequence of the *Sus2* gene in sugarcane and to compare the promoter regions of the genes in diverse genotypes, in order to develop PCR-based molecular markers for this gene in sugarcane. We isolated and sequenced the entire gene encoding one of the forms of sucrose synthase, *Sus2*, in sugarcane. The entire sequence is 7771 base pairs (bp) long (GenBank accession AY118266) and includes a promoter region of 1900 bp, an open reading frame encoded by 16 exons, and 258 bp of 3' end sequence. It shares many of the features of the homologous maize (*Zea mays* L.) gene, *Shrunken1*. Southern blot analysis indicated that *Sus2* is a low copy number gene, despite the high degree of polyploidy in the genus *Saccharum*. We cloned and compared sequences of the promoter region of *Sus2* from 'Muntok Java,' a *S. officinarum* x *S. spontaneum* hybrid, and 'PIN 84-1,' a *S. spontaneum*. The promoter sequences were polymorphic between the genotypes, and both genotypes were heterozygous. The sequence differences were primarily due to the presence of different large insertion-deletions (indels) ranging in size from 233 to 247 bp. The potential to use the variable indels as PCR markers for *Sus2* was demonstrated. These may be useful molecular markers for identifying different *Sus2* alleles.

INTRODUCTION

Sucrose synthase (*Sus*, E.C. 2.4.1.13) is a major enzyme of sucrose metabolism in sugarcane (*Saccharum* spp. hybrids) (Botha and Black, 2000; Lingle, 1999; Lingle et al., 2001). In a mapping study of *Saccharum*, a probe from the maize (*Zea mays*) *Shrunken-1* (*Sh1*) gene identified a restriction fragment length polymorphism (RFLP) linked to the Brix character in sugarcane (Ming et al., 2001). We previously cloned and sequenced the cDNA for the *Sus2* gene from sugarcane (GenBank accession AF263384; Lingle and Dyer, 2001). This gene is homologous to the maize gene that produces the *Shrunken-1* phenotype. Enzyme and northern analysis of developing internodes of diverse genotypes indicate that this gene is differentially expressed among genotypes (Lingle et al., 2001). Differential gene expression may be the result of polymorphism within the promoter region of the gene.

While useful for gene mapping studies, RFLP analysis is not feasible for marker assisted selection in breeding programs. Polymerase chain reaction (PCR) markers would be more useful. This study was done to isolate and characterize the complete genomic sequence of the *Sus2* gene in sugarcane, and to compare the promoter regions of the genes in diverse genotypes in order to develop PCR-based molecular markers for this gene in sugarcane.

MATERIALS AND METHODS

Isolation and Sequencing of the *Sus2* Gene

Filters of the sugarcane bacterial artificial chromosome (BAC) library SHCRBa (Tomkins et al., 1999) were obtained from Clemson University Genomics Institute (CUGI, Clemson, South Carolina). This BAC library was constructed from genomic DNA of the hybrid 'R570'. Nonradioactive digoxigenin (DIG)-labeled probe of *Sus2* was synthesized by PCR from a sugarcane *Sus2* cDNA (Lingle and Dyer, 2001) using the PCR DIG Probe Synthesis Kit¹ (Roche Diagnostics Corporation, Indianapolis, IN) following the manufacturer's protocol. The probe was hybridized to the filters in EasyHyb (Roche Diagnostics Corporation) overnight at 42°C, then washed, blocked and visualized using chemiluminescent detection (CDP-Star) following the Roche protocol. Among the six filters there were 18 positive clones. Ten of these were chosen at random and obtained from CUGI. The presence of at least part of *Sus2* in the BAC clones was verified by PCR. One BAC clone, SHCRBa 117B10, was randomly selected for shotgun sequencing.

BAC DNA was shotgun subcloned using the TOPO Shotgun Subcloning Kit (Invitrogen Corporation, Carlsbad, CA). Plasmids isolated from subclones were sequenced using BigDye Terminator and BigDye Primer kits (Applied Biosystems, Foster City, CA) and ABI Prism 377 DNA Sequencer (Applied Biosystems). Sequence homology was determined by BLAST searches (Altschul et al., 1997) of nucleotide sequence databases at the National Center for Biotechnology Information (NCBI).

Once we had sequenced DNA fragments from the BAC clone that were homologous to the sugarcane *Sus2* cDNA or *Sus2* genes from maize or other related grasses, the sequence information was used to design sets of primers to amplify overlapping parts of the gene from 'Muntok Java', a sucrose-storing hybrid of *S. officinarum* and *S. spontaneum*. The sequences of those fragments were assembled to give the entire gene sequence.

Sequence Variation Within and Between Genotypes

Gene polymorphisms are often found in the promoter region. We designed primers to amplify the 5' end of the gene. The primers, 5'-TTTGGGTATTTGGCTTCTGG-3' and 5'-TGGGATGAGAGGAGAAGGTG-3', were designed to span the region from about 2000 base pairs upstream of the transcription start site to the start of the third exon. Genomic DNA isolated from Muntok Java and 'Pin84-1', a low sucrose *S. spontaneum* genotype, and BAC DNA from SHCRa 117B10 were amplified using the primers and Platinum PCR Supermix (Invitrogen.) The PCR reaction produced fragments about 3500 base pairs long (data not shown). PCR products were ligated into vector pCRII-TOPO using the TOPO TA Cloning Kit (Invitrogen), and transformed into chemically competent TOP10 cells. Plasmid DNA was isolated from four randomly chosen positive colonies from each genotype and the BAC clone. All PCR products were sequenced as described above. For comparison among PCR products, sequences were aligned using the AlignX program of Vector NTI Suite 7.0 (InforMax, Inc., Bethesda, MD).

¹Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U. S. Department of Agriculture.

Southern Blot

We used Southern blot to determine the complexity of the *Sus2* gene in the two genotypes. Genomic DNA (25 µg) from PIN 84-1 and Muntok Java was cut with EcoRI, HindIII, and XbaI, and with different combinations of these enzymes. Eight µg digested DNA was separated on 0.9% agarose gel in TBE, then transferred to positively charged nylon membrane by downward capillary blotting. Blots were hybridized overnight at 42°C in Easy Hyb (Roche) with DIG-labeled DNA probe synthesized by PCR of a linearized DNA clone using primers designed in Exon 2. Hybridized blots were washed, blocked and visualized as described above.

PCR Markers

Primers were designed to generate PCR markers for the *Sus2* gene based on two different promoter indels, designated A and B. Sense primers were designed specific to the indel sequences. These were 5'-TCGGGACGAATCTGTTGAG-3' and 5'-GATTCGATGTGATGGCAAGCAC-3', respectively. The antisense primer was designed in a conserved region downstream of the insertion point for those indels, 5'-GCATACAAAGGACAATAATAAAAGA-3'. PCR reactions were run using these primer sets on genomic DNA from PIN 84-1, Muntok Java, five progeny from a cross of PIN 84-1 x Muntok Java, and five additional genotypes: 'Yellow Caledonia' (*S. officinarum*), 'Molokai' (*S. robustum*), 'China' (*S. sinense*), 'Newra' (*S. barberi*), and 'US 56-15-8' (*S. spontaneum*). PCR products were separated on 1.5 % (w/v) agarose gels and visualized by ethidium bromide staining.

RESULTS AND DISCUSSION

The sequence of the *sucrose synthase-2* (*Sus2*) gene from sugarcane (GenBank accession AY118266) is 7771 base pairs long (Fig. 1). Sugarcane *Sus2* contains 16 exons that are 99 % identical to the *scsusy2* cDNA from sugarcane (accession AF263384). The BLAST search of NCBI's nonredundant database produced significant alignment with the maize *Sh1* sucrose synthase gene (*ZMSUCS2*; accession X02382), with an identity of about 90 %. The gene shares many features of the maize gene, including a long (1050 bp) intron in the 5' untranslated region of the pre-mRNA. Also similar to the maize gene, there is evidence of two polyadenylation sites. This was determined by sequencing the 3' end of the cDNA from several genotypes (data not shown). In addition, there is a short simple sequence (TAAAA) repeated three times in intron 14 that may be polymorphic among other sugarcane genotypes (data not shown).

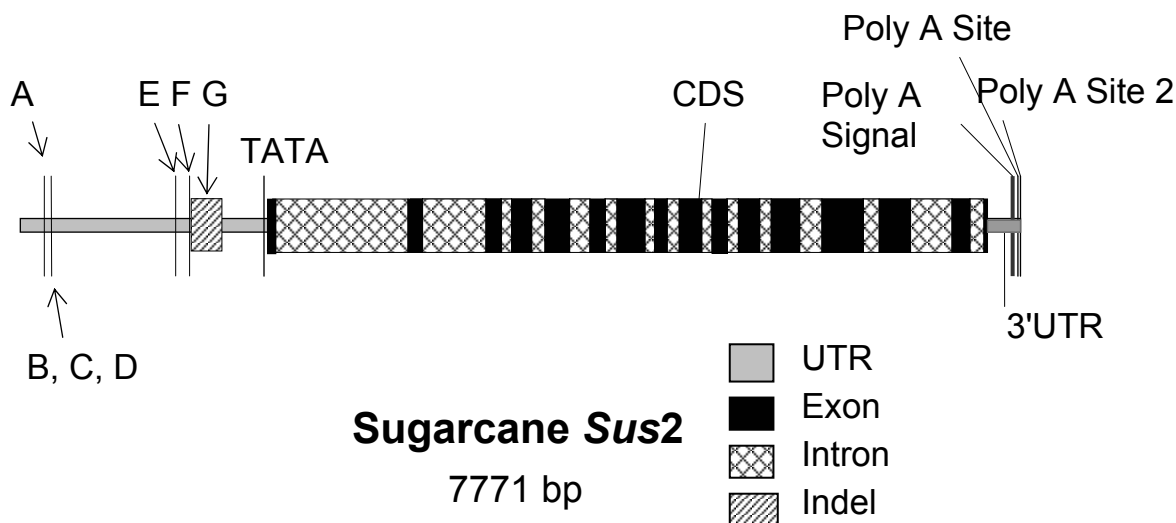


Figure 1. Schematic representation of the *sucrose synthase-2* (*Sus2*) gene in sugarcane. The repetitive element indel arbitrarily designated 'G' is present in the sequence deposited in GenBank (accession AY118266). Letters A to F indicate the relative locations of other indels present in one or more genotypes.

To determine sequence variation in the promoter region of *Sus2* between and within genotypes, we compared the sequence of four PCR products from the 5' region of the gene from Muntok Java and from PIN 84-1. Each of the four PCR products from Muntok Java was different, while there were two different sequences from PIN 84-1. Thus, both genotypes are heterozygous. The differences were mostly upstream of the transcription start point. Among all the PCR products there was 96 % identity in the sequence downstream of the transcription start point. The exons were 98 % identical, while the introns were 96 % identical. In maize, *Sh1* alleles are 98 % identical, while *Sh1* is only 56 % identical to *Sus1*. Therefore, we concluded that PCR products we sequenced from sugarcane were alleles of *Sus2* rather than other members of the *Sus* gene family. The sequences have been deposited in GenBank (Accessions AY670698, AY670699, AY670700, AY670701 and AY670702).

Upstream of the transcription start point, there were seven indels of 233 to 247 bp that were present in one or more of the PCR products from the two genotypes. These were arbitrarily designated A to G (Fig. 2). The program RepeatMasker (Smit, AFA & Green, P RepeatMasker at <http://ftp.genome.washington.edu/RM/RepeatMasker.html>) identified each of these indels as similar to repetitive elements. Each of the indels had 62 to 70 % homology to one or more internal sections of the Candystripe (*Cs1*) transposon of Sorghum (Chopra et al., 1999). Thus, the indels appear to be remnants of a transposon. The homology of the indels to *Cs1* and to each other is shown in Table 1.

Lingle: Polymorphism in the Promotor Region in the Sucrose Synthase-2 Gene of *Saccharum* Genotypes

While the function of the indel sequences remains unknown, they contain potential cis-acting sequences that might influence gene expression. Searches of the Plant Cis-acting Regulatory DNA Elements (PLACE) database (Higo, et al. 1999) showed that each indel contains several DNA sequences that influence tissue- or environment-specific expression of other genes. Different indels contain different elements, which may influence the expression of particular alleles. For instance, indels E and F have putative dehydration-responsive elements, while the others do not. These differences may be especially significant in indels E, F and G, which are within 1000 bp of the transcription start site.

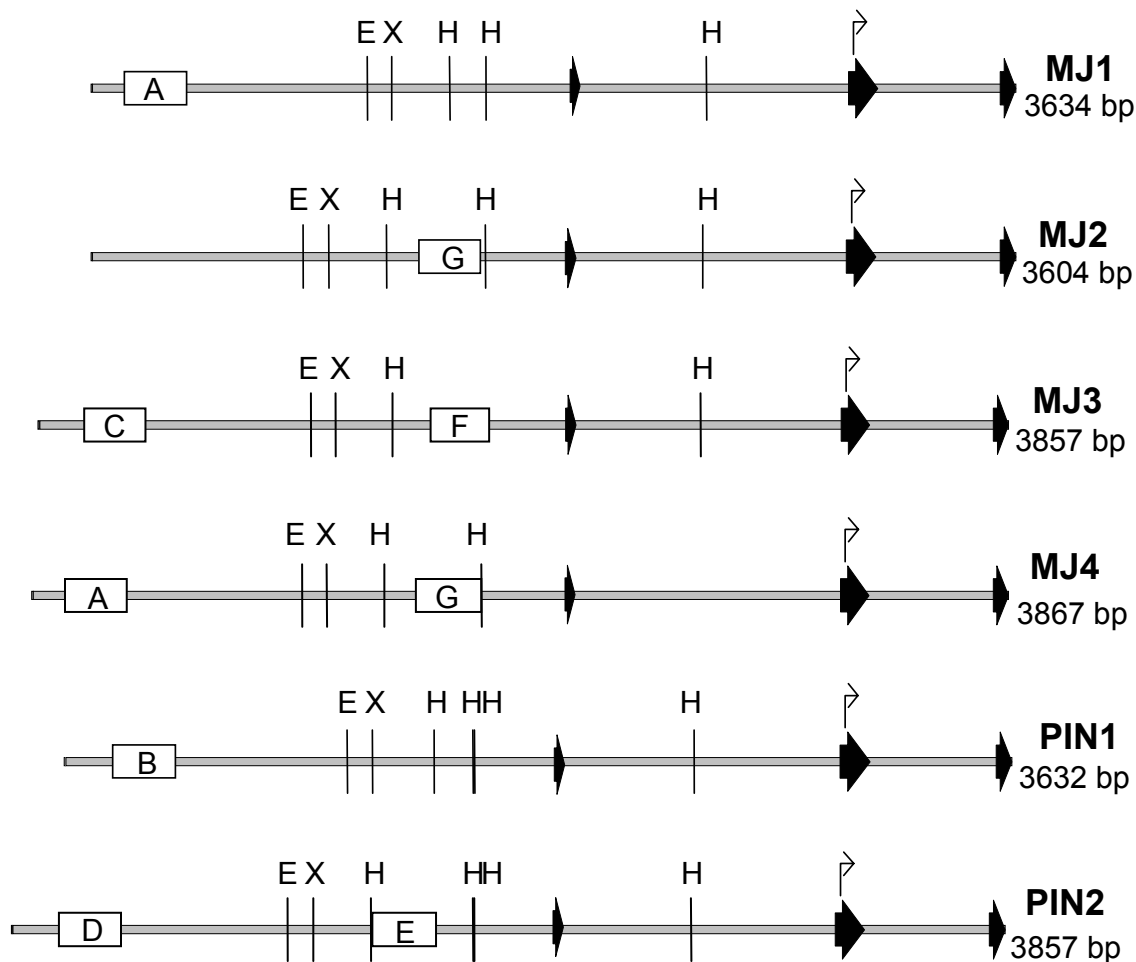


Figure 2. Schematic representation of the sequence of PCR products from the promoter region of *sucrose synthase-2* ‘Muntok Java’ (MJ) and ‘PIN84-1’ (PIN), indicating position of large indels (white boxes labeled A to G). Solid black arrows indicate exons. Bent arrows indicate the start codon of the coding sequence. Letters indicate restriction sites: E, *EcoRI*; H, *HindIII*; X, *XbaI*.

Table 1. Homology of seven indels in the 5' promoter region of the *Sus2* gene from different sugarcane genotypes to *Sorghum Candystripe1* (*Cs1*) and each other. Negative signs indicate homology to the reverse complement.

Indel	Size	Percent homology to						
		Cs1	B	C	D	E	F	G
A	246	71.2	-71.5	-70.3	-69.6	-60.3	57.8	89.5
B	240	-70.8		86.6	87.8	69.0	-59.6	-73.6
C	244	-68.1			88.6	73.7	-58.9	-74.7
D	243	-68.5				71.7	-59.9	-75.5
E	247	-69.6					-62.8	-67.2
F	233	62.6						56.7
G	241	67.7						

The presence of multiple copies of the *Sus2* gene in sugarcane is not surprising in this complex polyploid. Of the genotypes used in this study, PIN84-1 is dodecaploid ($2n=96$), while Muntok Java has $2n=120$ chromosomes (Ming et al., 2001). Despite these large ploidy levels, a Southern blot revealed a simple pattern (Fig. 3). In all cases, only one or two bands were produced by each enzyme digest. The different size bands observed on the Southern blot were fully explained by the differences in indels in the promoter region.

To begin to dissect the contribution of *Sus2* to sugar accumulation in this complex genetic background, we developed PCR markers for the different *Sus2* alleles. Primer sets were designed based upon the indel sequences present in the *Sus2* promoter region, and these primers were used in PCR reactions with DNA extracted from PIN 84-1, Muntok Java, some of their progeny, and representatives of other *Saccharum* species. The primer set for indel A, which was present in two of the sequenced promoter regions from Muntok Java, produced a 230 bp product in Muntok Java, and one of the PIN 84-1 x Muntok Java progeny (Fig. 4). Primer set A also produced 230 bp products in Yellow Caledonia (*S. officinarum*), China (*S. sinense*), and Newra (*S. barberi*), but not in PIN 84-1, Molokai (*S. robustum*) or US 56-15-8 (*S. spontaneum*). The absence of the A230 product in other *S. spontaneum* genotypes has since been confirmed (J. Veremis, USDA-Houma, personal communication). Although this is not a complete survey, it is consistent with what is believed about relationships among *Saccharum* species. *Saccharum officinarum* and *S. spontaneum* are presumed to be the most diverged of the species, while *S. sinense* and *S. barberi* are hybrids of *S. officinarum* and *S. spontaneum* (Lu, et al. 1994; D'Hont et al., 1994). *Saccharum robustum* is a presumed ancestor of *S. officinarum*. We propose that the A230 product arose from an event in *S. officinarum* after that species separated from *S. robustum*, from which it is absent, and then was transmitted to *S. sinense* and *S. barberi* from the *S. officinarum* progenitors of those species.

Primer set B produced a 120 bp product in PIN 84-1, and four progeny, but not in any other genotype. However, a quick survey of other genotypes revealed the presence of this marker in many genotypes of all species (J. Veremis, personal communication). Because Insert

Lingle: Polymorphism in the Promotor Region in the Sucrose Synthase-2 Gene
of *Saccharum* Genotypes

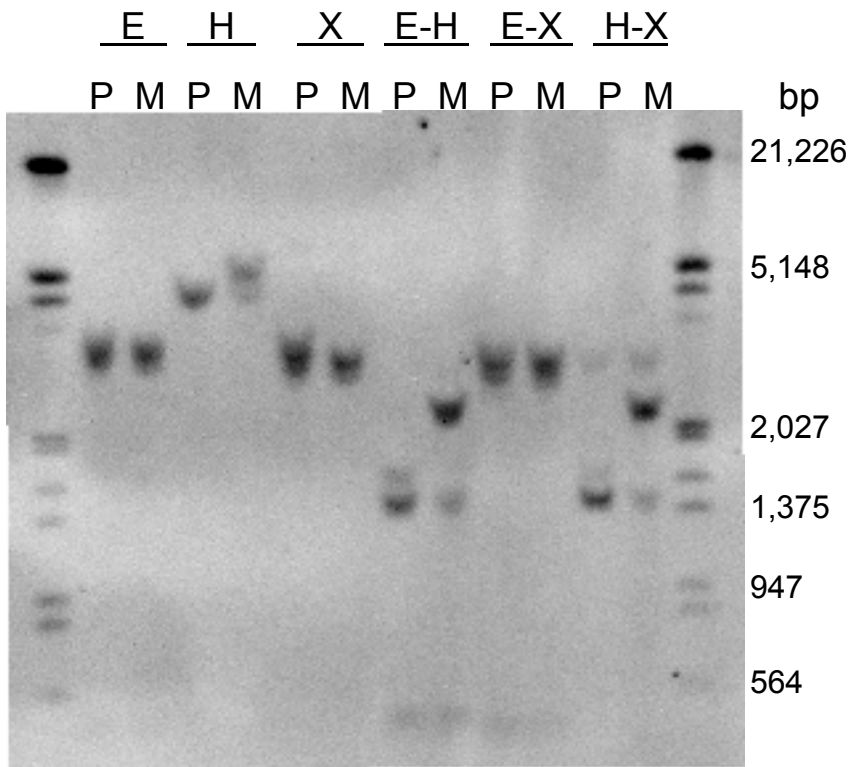


Figure 3. Southern blot of genomic DNA from *Saccharum* genotypes “PIN 84-1” (P) or “Muntok Java” (M) digested with EcoRI (E), HindIII (H), XbaI (X), and three enzyme combinations. Size markers (in base pairs) are indicated to the right.

B, C, and D are very similar, it is possible that the primer pairs for Insert B are not specific enough to reduce nonspecific amplification.

The primer sets, especially primer set A, produce markers that are polymorphic and inherited, which will make them useful for selecting for specific *Sus2* alleles in further studies, and for understanding the correlation between specific *Sus2* alleles and accumulation of sucrose in sugarcane.

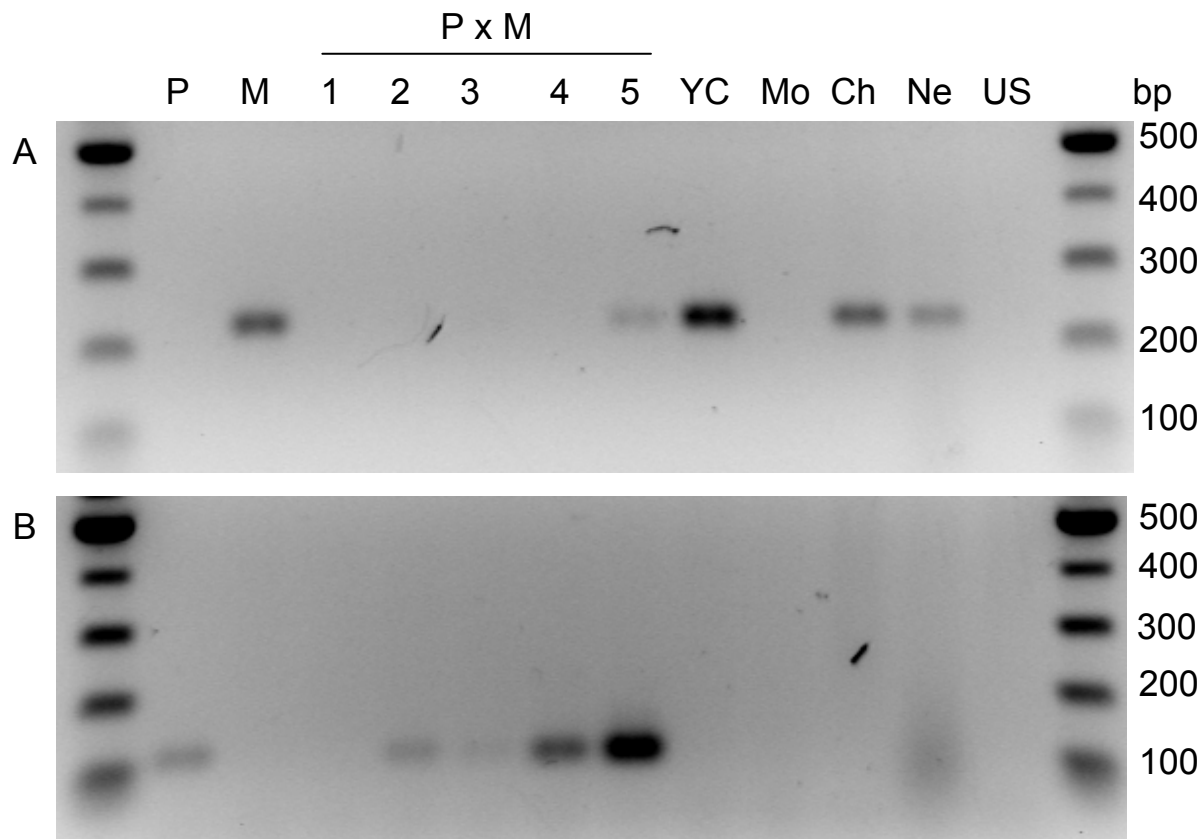


Figure 4. Ethidium-bromide stained agarose gel showing amplification of Indel A-specific (A) and Indel B-specific (B) products from genomic DNA of *Saccharum* genotypes ‘PIN 84-1’ (P), ‘Muntok Java’ (M), five PIN 84-1 x MJ Progeny (1-5), ‘Yellow Caledonia’ (YC; *S. officinarum*), ‘Molokai’ (Mo; *S. robustum*), ‘China’ (Ch; *S. sinense*), ‘Newra’ (Ne, *S. barberi*), and ‘US 56-15-8’ (US; *S. spontaneum*). Size markers are indicated to the right.

REFERENCES

1. Altschul, S.F., T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhan, W. Miller, and D.J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389-3402.
2. Botha, F.C., and K.G. Black. 2000. Sucrose phosphate synthase and sucrose synthase activity during maturation of internodal tissue in sugarcane. *Australian J. Plant Physiol.* 27:81-85.
3. Chopra, S., V. Brendel, J.B. Zhang, J.D. Axtell, and T. Peterson. 1999. Molecular characterization of a mutable pigmentation phenotype and isolation of the first active transposable element from *Sorghum bicolor*. *Proceedings of the National Academy of Sciences of the United States of America* 96:15330-15335.
4. D'Hont, A., Y.H. Lu, D. González de León, L. Grivet, P. Feldmann, C. Lanuad, and J.C. Glaszmann. 1994. A molecular approach to unraveling the genetics of sugarcane, a complex polyploid of the Andropogoneae tribe. *Genome* 37:222-230.
5. Higo, K., Y. Ugawa, M. Iwamoto, and K. Korenaga. 1999. Plant cis-acting regulatory DNA elements (PLACE) database. *Nucleic Acids Res.* 27:297-300.
6. Lingle, S.E. 1999. Sugar metabolism during growth and development in sugarcane internodes. *Crop Sci.* 39:480-486.
7. Lingle, S.E., A.B. Allen, and M.I. Valdez-Garza. 2001. Comparison of sucrose metabolism and gene expression in two diverse *Saccharum* genotypes. *Proc. Int. Soc. Sugar Cane Technol.* 24:323-326.
8. Lingle, S.E., and J.M. Dyer. 2001. Cloning and expression of sucrose synthase-1 cDNA from sugarcane. *J. Plant Physiol.* 158:129-131.
9. Lu, Y.H., A. D'Hont., D.I.T. Walker., P.S. Rao, P. Feldmann, and J.C. Glaszmann. 1994. Relationships among ancestral species of sugarcane revealed with RFLP using single copy maize nuclear probes. *Euphytica* 78:7-18.
10. Ming, R., S.C. Liu, P.H. Moore, J.E. Irvine, and A.H. Paterson. 2001. QTL analysis in a complex autopolyploid: genetic control of sugar content in sugarcane. *Genome Res.* 11:2075-2084.
11. Tomkins, J.P., Y. Yu, H. Miller-Smith, D.A. Frisch, S.S. Woo, and R.A. Wing. 1999. A bacterial artificial chromosome library for sugarcane. *Theoret. Appl. Gen.* 99:419-424.