GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

Genetic variability among the chloroplast genomes of sugarcane (*Saccharum* spp) and its wild progenitor species *Saccharum spontaneum* L.

J.-R. Zhu[1]*, H. Zhou[2]*, Y.-B. Pan[3] and X. Lu[1]

[1]Yunnan Key Laboratory of Sugarcane Genetic Improvement, Sugarcane Research Institute, Yunnan Academy of Agricultural Sciences, Kaiyuan, Yunnan, China
[2]Guangxi Laboratory of Sugarcane Biotechnology and Genetic Improvement, China Ministry of Agriculture, Sugarcane Research Institute, Guangxi Academy of Agricultural Sciences, Nanning, Guangxi, China
[3]United States Department of Agriculture, Agricultural Research Service, Sugarcane Research Laboratory, Houma, LA, USA

*These authors contributed equally to this study.
Corresponding author: Y.-B. Pan
E-mail: yongbao.pan@ars.usda.gov

Genet. Mol. Res. 13 (2): 3037-3047 (2014)
Received September 23, 2013
Accepted October 2, 2013
Published January 24, 2014
DOI http://dx.doi.org/10.4238/2014.January.24.3

ABSTRACT. A striking characteristic of modern sugarcane is that all sugarcane cultivars (*Saccharum* spp) share a common cytoplasm from *S. officinarum*. To explore the potential value of *S. spontaneum* cytoplasm, new *Saccharum* hybrids with an *S. spontaneum* cytoplasm were developed at the United States Department of Agriculture-Agricultural Research Service, Sugarcane Research Laboratory, through a combination of conventional and molecular breeding approaches. In this study, we analyzed the genetic variability among the chloroplast genomes of four sugarcane cultivars, eight *S. spontaneum* clones, and three F_1 progeny containing an *S. spontaneum* cytoplasm. Based on the complete chloroplast

genome sequence information of two sugarcane cultivars (NCo 310 and SP 80-3280) and five related grass species (barley, maize, rice, sorghum, and wheat), 19 polymerase chain reaction primer pairs were designed targeting various chloroplast DNA (cpDNA) segments with a total length varying from 4781 to 4791 bp. Ten of the 19 cpDNA segments were polymorphic, harboring 14 mutation sites [a 15-nt insertion/deletion (indel), a 5-nt indel, two poly (T) tracts, and 10 single nucleotide polymorphisms]. We demonstrate for the first time that the chloroplast genome of *S. spontaneum* was maternally inherited. Comparative sequence homology analyses clustered sugarcane cultivars into a distinctive group away from *S. spontaneum* and its progeny. Three mutation sites with a consistent, yet species-specific, nucleotide composition were found, namely, an A/C transversion and two indels. The genetic variability among cpDNA of sugarcane cultivars and *S. spontaneum* will be useful information to determine the maternal origin in the *Saccharum* genus.

Key words: Sugarcane; *Saccharum spontaneum*; Chloroplast genome; Maternal inheritance

INTRODUCTION

Sugarcane cultivars (*Saccharum* spp) are aneupolyploid, interspecific hybrids of *S. officinarum* and *S. spontaneum*. Ming et al. (1998) hypothesized that only two species, *S. robustum* and *S. spontaneum*, are the progenitors of modern sugarcane, that *S. officinarum* is derived from *S. robustum*, and that *S. barberi* and *S. sinense* are cultivated forms of interspecific hybrids between *S. spontaneum* and *S. officinarum*. Nonetheless, almost all sugarcane cultivars grown in the world today are derived from a few common ancestral clones and therefore share a limited genetic base (Arceneaux, 1967; Tew, 1987, 2003). A common objective of many breeding programs is to expand the genetic base of sugarcane cultivars by introducing agriculturally desirable traits from related wild species, in particular, *S. spontaneum* (Legendre and Breaux, 1983; Burner and Legendre, 1993). *S. spontaneum* clones are the most widely used germplasm in the basic (introgression) breeding program at the United States Department of Agriculture-Agricultural Research Service (USDA-ARS), Sugarcane Research Laboratory (SRL), because of their cold tolerance, ratoon ability, resistance to diseases and insects, and vigor (Dunckelman and Breaux, 1970; Dunckelman and Legendre, 1982). Thirty-three *S. spontaneum* clones that had been maintained at the SRL were classified into eight distinctive groups based on the genetic variability revealed by random amplified polymorphic DNA (RAPD) markers (Pan et al., 2004a). The grouping was found to be independent of both geographic origin and morphology.

Genes that control plant growth and development are located on either the nucleus (the nuclear genome) or the cytoplasm (the cytoplasmic genome). Several studies have demonstrated that the nuclear genomes of sugarcane cultivars consist of 100-130 chromosomes (D'Hont et al., 1996; Piperidis et al., 2010). About 15-27.5% of these chro-

mosomes are derived from *S. spontaneum*, of which 10-23% are entire chromosomes of *S. spontaneum* and 8-13% are recombinant chromosomes between *S. officinarum* and *S. spontaneum*. However, the cytoplasmic genomes of all sugarcane cultivars grown in the world today are derived only from *S. officinarum* (Melloto-Passarin et al., 2004). Since 1997, the sugarcane breeders at the USDA-ARS, SRL, have attempted to explore the potential value of the cytoplasmic genome of *S. spontaneum* by crossing *S. spontaneum* clones as maternal parents with elite cultivars. The resulting F_1 progeny were identified first by clone-specific DNA markers prior to transplanting into the field for evaluation and selection. Several new lines of *Saccharum* hybrids with an *S. spontaneum* cytoplasm were selected for further improvement (Pan et al., 2004b, 2006). One hybrid, Ho 02-113, has been thoroughly tested and released as the first biofuel cultivar that contains the cytoplasm of SES 234, an *S. spontaneum* clone (Hale et al., 2012).

Asano et al. (2004) and Calsa et al. (2004) published the complete chloroplast genomic DNA sequences of two sugarcane cultivars, namely, NCo 310 (GenBank ID: NC006084) and SP 80-3280 (GenBank ID: NC005878). Both chloroplast genomes share an identical size of 141,182 bp, with an overall (A+T) content of 61.6%, and almost identical nucleotide sequences. Upon comparing the complete chloroplast genomic DNA sequences between NCo 310 and maize (GenBank ID: X86563), Takahashi et al. (2005) found 26 polymorphic regions. They designed polymerase chain reaction (PCR) primers to amplify these regions from the six *Saccharum* species and sequenced these amplified DNA products. Based on DNA sequence analysis, five *Saccharum* species, namely, *S. officinarum*, *S. robustum*, *S. sinense*, *S. barberi*, and *S. edule*, were more closely related to one another than to *S. spontaneum*. In addition, the DNA sequences of these 26 chloroplast genomic regions were identical between *S. sinense* and *S. barberi*. A higher level of intra-specific sequence variations was observed in *S. spontaneum*, whereas no or very low levels of sequence variability were found within each of the other five *Saccharum* species.

In this study, we investigated the genetic variability among the chloroplast genomes of two sugarcane cultivars, eight *S. spontaneum* clones representing each of the eight subgroups of *S. spontaneum* defined by Pan et al. (2004a) and three F_1 progeny containing an *S. spontaneum* cytoplasm. We also studied the inheritance of parental *S. spontaneum* chloroplast genome into its F_1 progeny.

MATERIAL AND METHODS

Plant material and nucleic acid extraction

Two sugarcane cultivars (CP 62-258 and LCP 85-384), eight *S. spontaneum* clones, two F_1 progeny (US 99-44 and US 99-47) of the *S. spontaneum* clone Djatiroto as the maternal parent, and one F_1 progeny (Ho 02-113) of the *S. spontaneum* clone SES 234 as the maternal parent were included in the study (Table 1). Plants of these clones were maintained in the Pathology Greenhouse at the SRL, Houma, LA, USA. The top visible dewlap leaf was the source of tissue collection. Total nucleic acids were extracted from about 0.2 g leaf tissue according to Pan et al. (2000). DNA concentrations were determined by NanoDrop™ 1000 (NanoDrop, Bethesda, MD, USA), and qualities were checked by agarose gel electrophoresis. The DNA samples were stored at -20°C until use.

Table 1. A list of *Saccharum* clones from the local collection at the USDA-ARS, MSA, Sugarcane Research Laboratory, Houma, Louisiana.

Name	Species	Origin[a]	Group[a]
Tainan	*Saccharum spontaneum*	Taiwan	I
S 66-84	*Saccharum spontaneum*	Taiwan	II
US 56-15-8	*Saccharum spontaneum*	Thailand	III
SES 234	*Saccharum spontaneum*	India	IV
IND 82-311	*Saccharum spontaneum*	India	V
Djatiroto	*Saccharum spontaneum*	Indonesia	VI
IND 81-144	*Saccharum spontaneum*	India	VII
SES 231	*Saccharum spontaneum*	India	VIII
CP 62-258	*Saccharum* spp	USA	
LCP 85-384	*Saccharum* spp	USA	
Ho 02-113[b]	*Saccharum* spp (F_1 of SES 234 x LCP 85-384)	USA	
US 99-44[c]	*Saccharum* spp (F_1 of Djatiroto x LCP 85-384)	USA	
US 99-47[c]	*Saccharum* spp (F_1 of Djatiroto x LCP 85-384)	USA	

[a]Pan et al. (2004a). [b]Pan et al. (2006); Hale et al. (2012). [c]Pan et al. (2004b).

Primer design, PCR amplification, and DNA sequencing

The complete chloroplast DNA (cpDNA) sequences of two sugarcane cultivars, namely, SP 80-3280 (GenBank ID: AE009947) (Calsa et al., 2004) and NCo 310 (GenBank ID: AP006714) (Asano et al., 2004), *Hordeum vulgare* (GenBank ID: EF115541), *Oryza sativa* (GenBank ID: AP006728), *Triticum aestivum* (GenBank ID: AB042240), *Zea mays* (GenBank ID: X86563), and *Sorghum bicolor* (GenBank ID: EF115542) were downloaded from the GenBank database (http://www.ncbi.nlm.nih.gov/genbank). The two cultivars share the same cpDNA size of exactly 141,182 bp and almost the same nucleotide sequence except for four single nucleotide polymorphisms (SNPs), two C/T and two G/T. All downloaded cpDNA sequences were aligned using the Multiple Sequence Alignment Program of DNAMAN® (Lynnon BioSoft, Vaudteuil, Quebec, Canada) to display nucleotide sequence variability. Based on the sequence variability, 19 PCR primer pairs were designed with DNAMAN® targeting the most polymorphic regions (Table 2). The primers were synthesized at Integrated DNA Technologies, Inc. (Coralville, IA, USA).

The PCR volume was 20 µL, containing approximately 10 ng total nucleic acids, 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 3.0 mM $MgCl_2$, 0.2 mM dNTPs, 0.3 µL of each forward and reverse primers (10 mM), 4.0 µL 5X Q-solution, and 0.5 U HotStart *Taq* DNA polymerase (Qiagen, Valencia, CA, USA). The PCR amplifications were conducted on a GeneAmp PCR System 9700 (Applied Biosystems, Foster City, CA, USA) following the cycling program of 94°C for 15 min; 40 cycles of 94°C for 45 s, annealing for 45 s, and 72°C for 45 s; and 72°C for 10 min. Annealing temperatures varied with primer pairs used (Table 2). Amplified PCR products were analyzed on 2% binary gels consisting of 0.7% agarose (Sigma-Aldrich, St. Louis, MO, USA), 0.65% Synergel (Diversified Biotech, New York, NY, USA), and 0.25 µg/mL ethidium bromide in 0.5X Tris, borate, and ethylenediaminetetraacetic acid (TBE) buffer. Gel electrophoreses were conducted at 4 V/cm for 1 h. Gel images were documented on a Gel Logic 200 Imaging System (Eastman Kodak, Rochester, NY, USA). A gel piece containing the target PCR product was sampled using a 10-µL pipette tip for a second round of PCR amplification following the same protocols described above except that a 30-µL PCR volume was used. The target PCR products from the second-round amplification were subjected to direct nucleotide sequencing by MCLAB (http://www.mclab.com).

Table 2. Description of chloroplast DNA segment, position in reference sequence of NCo 310 (AP006714), polymorphic nature, GenBank accession number, PCR primer sequence, and annealing temperature.

Segment No.	Position in NCo 310	P or N[a]	Accession No.	Primers sequence (5'→3')	Tm (°C)
1	4796-4978	P	N/A[b]	TTGGGATAGATGTAGATAA	50
				TGTTCGAGATGAGATAAA	48
2	11704-12091	P	JQ664306-	AATTCCGTGGGTGGCGTAGC	64
			JQ664319	TGGGCTGTGGAGAGATGGCTGA	70
3	13260-13477	P	JQ664320-	CCCCTTCCTATCTTATCC	54
			JQ664333	GCACTTAGCAGTCCGTAT	54
4	20925-21195	P	JQ664334-	TTCCCGCTTCTTTCTATTC	54
			JQ664347	CTGTTCGGACCTGGAGTTT	58
5	29336-29559	N	JQ664348-	TTGACGATCCACGATACAG	56
			JQ664361	CATACTCATCTTCTAGGGTT	56
6	32025-32412	N	JQ664362-	GCGTCGGAAATGAGAGATA	56
			JQ664375	AATGAACTCCCGCTTCTAT	54
7	37950-38317	P	JQ664376-	ACAGGCACAAGCTTATCGCCAAA	68
			JQ664389	CCCGTAATCGTCGACCCCTTGC	72
8	38876-39141	N	JQ664390-	CTCTGTCCTATCCATTAGA	54
			JQ664403	CCTCTTTACCATTCTGTAT	52
9	46553-46650	P	N/A	GGATACACGACAGAAGGAA	56
				GAGATGGTGCGATTTGACT	56
10	54468-54771	P	JQ664404	TGCCAGAGGATATCCCTTTCATCCG	76
			JQ664417	CGAGGCTGTCAATTGGATTCCCCC	76
11	57044-57381	P	JQ664418	GCATAGATTGCTGTCAACA	54
			JQ664431	GATTAGGCGTAAATGAAACC	58
12	67882-68211	P	JQ664432-	CGAAAGCATACCACTCATA	54
			JQ664445	TGGATTTACGAAAGGGTTG	54
13	81180-81256	P	N/A	TAATGGTTAGGTTTGAATC	50
				GGAAGAAGAATAAGAAAGAT	52
14	81605-81780	N	N/A	CAATACGAAGCAATAGGTTG	56
				TTCCGCCCAGTGATAGAAT	56
15	86547-86736	N	N/A	TTGCGAGCCAAGATAAGAC	56
	(137684-137495)			ATAAGAAGCGACCCACTTT	54
16	132169-132308	N	N/A	GCGTTTCATTTGCTTCTCT	54
	(92062-91923)			GGTTCTTCTCTCCATCGGA	58
17	123585-123902	N	JQ664460-	CGATTTGGTACCGCTCGCGCA	68
	(100646-100329)		JQ664473	TCCCAAGCAGTGGGAGGGGA	66
18	112959-113245	N	JQ664474-	GTTTGTGAGAAAGCGTGAA	54
			JQ664487	CAACTCGTATCAATCAATCC	56
19	118381-118607	N	JQ664488-	TTGATGGGTGTTGAGGACC	58
	(105843-105642)		JQ664501	TTTGCTGGCTTATTTGGCA	54

[a]P = polymorphic; N = non-polymorphic. [b]N/A: Sequences less than 200 bp were not accepted by the GenBank.

DNA sequence analysis, annotation, and homology tree

Nucleotide sequence files of target PCR products were processed with the FinchTV software (http://www.geospiza.com/finchtv) to produce cpDNA sequence files. The cpDNA sequence files were annotated using Sequin® (Version 11.90, NCBI) before submission to GenBank. Multiple sequence alignments were conducted for the 13 samples plus NCo 310 and SP 80-3280 on each of the 19 cpDNA segments using the optimal alignment method of DNAMAN® to reveal sequence variations, including transitions, transversions, and insertion/deletions (indels). These mutation sites were recorded and compiled to form an artificial nucleotide sequence or haplotype for each clone, in which the letter "N" was used for sites with missing nucleotides. The resulting artificial nucleotide sequences of the 15 haplotypes were aligned using DNAMAN® with manual editing so that the same mutation sites were aligned perfectly to produce a pairwise homology coefficient matrix for the construction of homology trees.

RESULTS

Nucleotide sequences of 19 cpDNA segments

In total, the nucleotide sequences of 19 cpDNA segments were determined. The start and end positions of these cpDNA segments corresponding to those of NCo 310 are shown in Table 2. The total length of these segments varied from 4781 to 4791 bp. The (A+T) content of these sequences varied from 65.6 to 65.7% among the 13 samples, with a mean of 65.6% (Table 3). These values were greater than the (A+T) content of the complete chloroplast genome of sugarcane cultivars (61.6%). Nine of these 19 cpDNA segments shared the same nucleotide sequences across samples, including NCo 310 and SP 80-3280, and therefore were monomorphic. However, varying degrees of nucleotide sequence polymorphism were found among the other 10 segments (Table 2). Eight primer pairs, namely, 3, 4, 6, 9, 11, 12, 13, and 14, amplified cpDNA fragments with mono-base sequence repeats ($n \geq 8$). Poly (A_8) and poly (T_{11}) tracts were found in the fragments amplified by primer pair 3, poly ($T_{10\text{-}13}$) was found in the segments amplified by primer pair 4, poly (T_8) was found in the segments amplified by primer pair 6, poly (A_9) was found in the segments amplified by primer pair 9, poly (T_{10}) was found in the segments amplified by primer pair 11, poly ($T_{10\text{-}12}$) was found in the segments amplified by primer pair 12, poly (T_{13}) was found in the segments amplified by primer pair 13, and poly (T_8) and poly (T_9) were found in the segments amplified by primer pair 14. In addition, the poly ($T_{10\text{-}13}$) tract amplified by primer pair 4 and the poly ($T_{10\text{-}12}$) tract amplified by primer pair 12 were polymorphic among all samples.

Table 3. Description of the 14 mutation sites among the 19 chloroplast DNA segments amplified from two sugarcane cultivars, eight *Saccharum spontaneum* clones, and three F_1 progeny by primer pairs 1, 2, 3, 4, 7, 9, 10, 11, 12, and 13.

Accessions	Segment/primer pair[a]														(A+T) (%)
	1	2	2	2	3	4	4	7	9	10	11	12	12	13	
NCo 310[b]	C	C	A	C	G	T_{10}---	T	A	C	A	G	TTTTTATATTTTTAT	T_{10}TT	-----	65.7
SP 80-3280[b]	C	T	C	T	G	T_{10}---	T	C	C	T	G	TTTTTATATTTTTAT	T_{10}TT	-----	65.7
CP 62-258	C	T	A	C	A	T_{10}---	C	A	C	T	A	TTTTTATATTTTTAT	T_{10}TT	-----	65.7
LCP 85-384	C	T	A	C	A	T_{10}---	C	A	C	T	A	TTTTTATATTTTTAT	T_{10}TT	-----	65.7
Djatiroto	A	T	A	C	C	T_{10}TT-	C	A	A	T	A	----------	T_{10}T-	AAGTA	65.6
Ho 02-113	A	T	A	C	C	T_{10}TT-	C	A	C	T	A	----------	T_{10}--	AAGTA	65.6
IND 81-144	A	T	A	C	C	T_{10}TT-	C	A	C	T	A	----------	T_{10}T-	AAGTA	65.6
IND 82-311	A	T	A	C	C	T_{10}TT-	C	A	C	T	A	----------	T_{10}TT	AAGTA	65.6
S 66-84	A	T	A	C	C	T_{10}TTT	C	A	C	T	A	----------	T_{10}T-	AAGTA	65.6
SES 231	A	T	A	C	C	T_{10}TT-	C	A	C	T	A	----------	T_{10}T-	AAGTA	65.6
SES 234	A	T	A	C	C	T_{10}TT-	C	A	C	T	A	----------	T_{10}--	AAGTA	65.6
Tainan	A	T	A	C	C	T_{10}TT-	C	A	C	T	A	----------	T_{10}T-	AAGTA	65.6
US 56-15-8	A	T	A	C	C	T_{10}TT-	C	A	C	T	A	----------	T_{10}T-	AAGTA	65.6
US 99-44	A	T	A	C	C	T_{10}TT-	C	A	A	T	A	----------	T_{10}T-	AAGTA	65.6
US 99-47	A	T	A	C	C	T_{10}TT-	C	A	A	T	A	----------	T_{10}T-	AAGTA	65.6

[a]Chloroplast genomic location of the mutations: primer pair 1, (C/A), *rps16*/intron; primer pair 2, (C/T), *psbC*; primer pair 2, (A/C), intergenic; primer pair 2, (C/T), *trnS*; primer pair 3, (A/G/C), intergenic region of *trnG* and *trnfM*; primer pair 4, [poly ($T_{10\text{-}13}$), T/C], intergenic region of *trnC* and *rpoB*; primer pair 7, (A/C), *atpA*; primer pair 9, (A/C), *ycf3*/intron; primer pair 10, (A/T), *trnM*; primer pair 11, (A/G), intergenic region of *atpB* and *rbcL*; primer pair 12, [(15-nt indel, poly ($T_{10\text{-}12}$)], intergenic region of *rpl33* and *rps18*, and primer pair 13, (5-nt indel), *rpl16*/intron. [b]Reference cpDNA sequences of the NCo 310 or SP 80-3280.

Characteristics of mutations

In total, 14 mutation sites were found on the 10 polymorphic cpDNA segments, includ-

ing a 15-nt indel (TTTTTATATTTTTAT), a 5-nt indel (AAGTA), two poly (T) polymorphisms (T_{10-12} and T_{10-13}), and 10 SNPs (Table 3). These mutations were located in the following chloroplast genomic sites: the (C/A) SNP from primer pair 1 within intron of *rps16*; the two (C/T) SNPs from primer pair 2 within *psbC* and *trnS*, respectively; the (A/C) SNP from primer pair 2 within intergenic region of *psbC* and *trnS*; the (A/G/C) SNP from primer pair 3 within intergenic region of *trnG* and *trnfM*; the (T/C) SNPs and poly (T_{10-13}) from primer pair 4 within intergenic region of *trnC* and *rpoB*; the (A/C) SNP from primer pair 7 within *atpA*; the (A/C) SNP from primer pair 9 within intron of *ycf3*; the (A/T) SNP from primer pair 10 within *trnM*; the (A/G) SNP from primer pair 11 within intergenic region of *atpB* and *rbcL*; the 15-nt indel and poly (T_{10-12}) from primer pair 12 within intergenic region of *rpl33* and *rps18*; and the 5-nt indel from primer pair 13 within intron of *rpl16* (refer to GenBank ID AE009947 and AP006714).

Of the 10 SNPs, four were transitions [one A/G (A to G or G to A) and three C/T], five were transversions (one A/T and four A/C), and one was a mix type (A/C/G). The 15-nt indel was absent in the cpDNA segments amplified by primer pair 12 from all *S. spontaneum* clones and the three interspecific F_1 progeny. It was present in corresponding cpDNA segments of all the four sugarcane cultivars (NCo 310, SP 80-3280, CP 62-258, and LCP 85-384). The opposite was true for the 5-nt indel amplified by primer pair 13, i.e., it was absent in the cultivars but present in all the *S. spontaneum* clones and the three interspecific F_1 progeny. In addition, all sugarcane cultivars shared the poly (T_{10}) and poly (T_{12}) in cpDNA segments amplified by primer pairs 4 and 12, respectively, whereas polymorphic poly (T_{10-13}) nucleotide sequences were found amongst *S. spontaneum* clones and the three F_1 progeny.

Maternal inheritance of cpDNA

Ho 02-113, an F_1 progeny of SES 234 and LCP 85-384, exhibited the cpDNA haplotype [(A)(T)(A)(C)(C)(TTN)(C)(A)(C)(T)(A)(NNNNNNNNNNNNNNN)(NN)(AAGTA)]. This cpDNA haplotype was identical to that of the maternal parent, SES 234. In contrast, the cpDNA haplotype of Ho 02-113 only shared eight common SNP sites (underlined) with the cpDNA haplotype of its paternal parent LCP 85-384, i.e. [(C)(T)(A)(C)(A)(NNN)(C)(A)(C)(T)(A)(TT TTTATATTTTTAT)(TT)(NNNNN)]. The 15-nt indel sequence (TTTTTATATTTTTAT) of LCP 85-384 was absent in both Ho 02-113 and SES 234. On the other hand, the 5-nt indel sequence (AAGTA) found in both Ho 02-113 and SES 234, and it was absent in LCP 85-384.

Similarly, the cpDNA haplotypes of US 97-44 and US 97-47, two F_1 progeny of Djatiroto and LCP 85-384, were the same, i.e. [(A)(T)(A)(C)(C)(TTN)(C)(A)(A)(T)(A)(NNNNNNNNNNNNNNN)(TN)(AAGTA)]. These haplotypes were identical to that of the maternal parent, Djatiroto. Again, the 15-nt indel nucleotides (TTTTTATATTTTTAT) found in LCP 85-384 were absent in US 97-44, US 97-47, and Djatiroto, whereas the 5-nt indel sequence (AAGTA) found in US 97-44, US 97-47, and Djatiroto was absent in LCP 85-384 (Table 3).

Homology between sugarcane cultivars and *S. spontaneum*

A homology tree is shown in Figure 1 in which the 13 samples and the two reference cultivars (NCo 310 and SP 80-3280) were clustered into two distinctive groups at a 58% homology level. Group I included the eight *S. spontaneum* clones and the three F_1 progeny. This result coincided with the report by Pan et al. (2004a), in which 33 *S. spontaneum* clones

were clustered into a main group exclusively from the cultivar group based on the genetic variability revealed by RAPD markers. Group II included all the sugarcane cultivars, namely, NCo 310, SP 80-3280, CP 62-258, and LCP 85-384. In addition, six identical mutation sites were shared by the cpDNA segments of all the sugarcane cultivars in Group II, which were amplified by primer pairs 1 (C/A transversion), 4 [poly (T_{10})], 9 (C/A), 12 [15-nt indel (TTTTTATATTTTTAT) and poly (T_{12})], and 13 [5-nt indel (N_5)], respectively. For *S. spontaneum* and F_1 progeny samples in Group I, 11 identical mutation sites were also detected, including those amplified by primer pairs 1 (A/C), 2 (T/C, A/C, C/T), 3 (C/A/G), 4 (C/T), 7 (A/C), 10 (T/A), 11 (A/G), 12 [15-nt indel (N_{15})], and 13 [5-nt indel (AAGTA)] (Table 3).

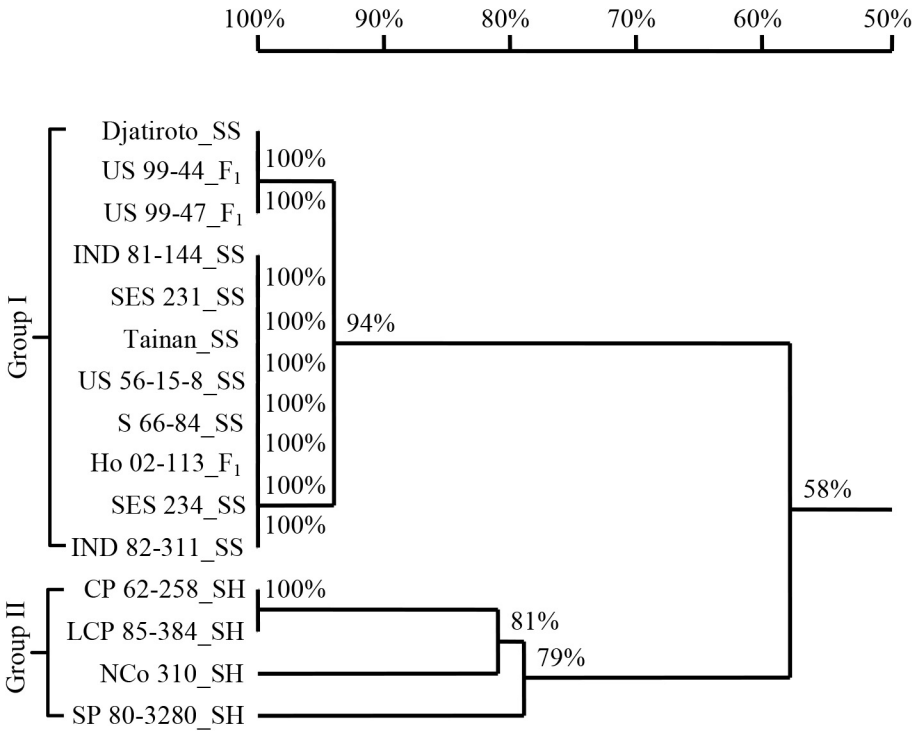

Figure 1. Homology tree among four sugarcane cultivars (SH) (CP 62-258 and LCP 85-384 from this study; NCo 310 and SP 80-3280 from GenBank database), eight *Saccharum spontaneum* (SS) clones, and three F_1 progeny (F_1) by the DNAMAN® software (Lynnon BioSoft, Vaudteuil, Quebec, Canada) based on the multiple sequence editor program.

DISCUSSION

cpDNA sequence variations were widely used to investigate interspecific relationships among angiosperms and other plants decades ago (Palmer et al., 1988; Clegg et al., 1991). In general, the (A+T) content of the chloroplast genome is greater than that of the nuclear genome, and most chloroplast microsatellites are composed of single-nucleotide repeats of A or T (Weising and Gardner, 1999). The overall (A+T) contents of all Poaceae plants are

very similar, for example, 61.6% for sugarcane (Asano et al., 2004), 61.0% for rice (GenBank ID: AP006728), 61.7% for wheat (GenBank ID: AB042240), 61.5% for maize (GenBank ID: X86563), 61.7% for barley (GenBank ID: EF115541), and 61.5% for sorghum (GenBank ID: EF115542). In this study, the average (A+T) content of the 19 cpDNA segments was 65.6%, which was greater than the (A+T) content of the whole cpDNA (61.6%) (Asano et al., 2004). In addition, there was a 15-bp insertion in all sugarcane cultivars that was composed of only A or T, rendering the (A+T) content of sugarcane cultivars (65.7%) slightly higher than that of *S. spontaneum* clones (65.6%).

When screening the complete sugarcane chloroplast genome sequence of 141,182 bp from the GenBank database for mononucleotide repeats with $n \geq 8$ (or $n \geq 10$), we found 57 (or 9) poly (A), 66 (or 22) poly (T), 3 (or 1) poly (G), and 2 (or 0) poly (C) tracts. Of the 19 amplified cpDNA segments from this study, we only found 8 (or 5) poly (T) and 2 (or 0) poly (A) tracts. Of the 65.6% average (A+T) content, 33.3% were A and 32.3% were T. Nonetheless, it appeared that the frequency of poly (T) was always greater than the frequency of poly (A). In general, more poly (T) tracts were present in more polymorphic segments. Takahashi et al. (2005) found many mutations within or adjacent to various types of sequence repeats and suggested that mutations would occur preferentially at sequences surrounded by sequence repeats. This was supported by our findings that almost all the 14 mutation sites found in this study either contained sequence repeats within these sites or were surrounded by sequence repeats.

The chloroplast genomes are maternally inherited in the majority of angiosperms, although there are a few exceptions (Lee et al., 1988; Chat et al., 1999). For example, reciprocal crosses between *Zea mays* and *Zea perennis* displayed strict maternal inheritance of chloroplast genomes (Conde et al., 1979), whereas cpDNA showed paternal inheritance in loblolly pine (Neale and Sederoff, 1989) and *Actinidia* (Testolin and Cipriani, 1997). In this study, we demonstrated for the first time that the cpDNA of *S. spontaneum* was maternally inherited. The three F_1 progeny of *S. spontaneum*, namely, Ho 02-113 (Pan et al., 2006; Hale et al., 2012), US 99-44, and US 99-47 (Pan et al., 2004a), shared identical cpDNA sequences to those of their maternal *S. spontaneum* parents, namely, SES 234 and Djatiroto.

The eight *S. spontaneum* clones involved in this study were selected to represent each of the eight distinctive *S. spontaneum* groups (Pan et al., 2004a). Although substantial degrees of genetic diversity were found among the eight clones based on RAPD fingerprints, the difference was minimal at the cpDNA sequence level. This indicated that the cpDNA sequence was conserved among *S. spontaneum* clones, albeit some degrees of cpDNA sequence variability were found between *S. spontaneum* and sugarcane cultivars. Of particular interest were the following observations. First, there were three species-specific mutation sites located on cpDNA segments amplified by primer pairs 1, 12, and 13. The primer pair 1-amplified cpDNA segment was C for cultivars but A for *S. spontaneum* and its F_1 progeny. For the primer pair 12-amplified cpDNA segment, the 15-nt sequence (TTTTTATATTTTTAT) was present in all sugarcane cultivars but was absent in *S. spontaneum* and its F_1 progeny. For the primer pair 13-amplified cpDNA segment, the 5-nt sequence (AAGTA) was present in *S. spontaneum* and its F_1 progeny but was absent in all sugarcane cultivars (Table 3). Second, the poly (T_{10}) site amplified by primer pair 4 was found exclusively in sugarcane cultivars. However, the site showed some length polymorphism ($T_{12\text{-}13}$) among *S. spontaneum* and its F_1 progeny. Third, the poly (T_{12}) site amplified by primer pair 12 was found in all sugarcane cultivars, but it showed some degree of length polymorphism ($T_{10\text{-}12}$) among *S. spontaneum* and its F_1 progeny

(Table 3). The sequence variations were confirmed by repeating the PCR amplification reactions by two different co-authors and re-sequencing of the resulting amplicons by the McLab.

In conclusion, 19 cpDNA segments were amplified through PCR from two sugarcane cultivars, eight *S. spontaneum* clones, and three F_1 progeny (*S. spontaneum* x *Saccharum* spp), and their nucleotide sequences were determined. Mutation survey and homology analysis showed 14 mutation sites among these clones, which included a 15-nt indel, a 5-nt indel, two poly (T) tracts, and 10 SNPs. The 15- and 5-nt indels and the (C/A) transversion were species-specific. Sugarcane cultivars formed a distinctive group from *S. spontaneum* and F_1 hybrids. We were able to demonstrate for the first time that the cpDNA of *S. spontaneum* was maternally inherited. The information of genetic variability among chloroplast genomes of sugarcane cultivars and *S. spontaneum* will be useful in the determination of maternal origin in the *Saccharum* genus.

ACKNOWLEDGMENTS

This study was conducted at the USDA-ARS, MSA, Sugarcane Research Laboratory under the USDA-ARS National Program "301" project "Genetic improvement of sugarcane by conventional and molecular approaches". Research partially funded by the American Sugar Cane League of the USA Inc., Thibodaux, LA. We also thank the State Administration of Foreign Experts Affairs of China Training Programs for providing living and international travel expenses for J.-R. Zhu and X. Lu and the Guangxi Natural Science Foundation (#2011GXNSFF018002) for providing international travel expense for H. Zhou.

REFERENCES

Arceneaux G (1967). Cultivated sugarcanes of the world and their botanical derivation. *Proc. Int. Soc. Sugarcane Technol.* 12: 844-854.

Asano T, Tsudzuki T, Takahashi S, Shimada H, et al. (2004). Complete nucleotide sequence of the sugarcane (*Saccharum officinarum*) chloroplast genome: a comparative analysis of four monocot chloroplast genomes. *DNA Res.* 11: 93-99.

Burner DM and Legendre BL (1993). Sugarcane genome amplification for the subtropics: a 20-year effort. *Sugar Cane* 3: 5-10.

Calsa JT, Carraro DM, Benatti MR, Barbosa AC, et al. (2004). Structural features and transcript-editing analysis of sugarcane (*Saccharum officinarum* L.) chloroplast genome. *Curr. Genet.* 46: 366-373.

Chat J, Chalak L and Petit RJ (1999). Strict paternal inheritance of chloroplast DNA and maternal inheritance of mitochondrial DNA in intraspecific crosses of kiwifruit. *Theor. Appl. Genet.* 99: 314-322.

Clegg MT, Learn GH and Golenberg EM (1991). Molecular evolution of chloroplast DNA. In: Evolution at the Molecular Level (Selander RK, Clark AG and Whittam TS, eds.). Sinauer Associates, Sunderland, 135-149.

Conde MF, Pring DR and Levings CS (1979). Maternal inheritance of organelle DNA's in *Zea mays-Zea perennis* reciprocal crosses. *J. Heredity* 70: 2-4.

D'Hont A, Grivet L, Feldmann P, Rao S, et al. (1996). Characterisation of the double genome structure of modern sugarcane cultivars (*Saccharum* spp.) by molecular cytogenetics. *Mol. Gen. Genet.* 250: 405-413.

Dunckelman PH and Breaux R (1970). New sugar cane breeding clones from Indian crosses evaluated at Houma, Louisiana, 1966-69. *Int. Sugar J.* 72: 43-44.

Dunckelman PH and Legendre BL (1982). Guide To Sugarcane Breeding in the Temperate Zone. USDA-ARS, ARM-S-22, New Orleans.

Hale AL, Dufrene EO, Tew TL, Cobill R, et al. (2012). Registration of 'Ho02-113' Sugarcane. *J. Plant Regis.* 7: 51-57.

Lee DJ, Blake TK and Smith SE (1988). Biparental inheritance of chloroplast DNA and the existence of heteroplasmic cells in alfalfa. *Theor. Appl. Genet.* 76: 545-549.

Legendre BL and Breaux RD (1983). The USDA Basic Sugarcane Breeding Program in Louisiana. Proceedings of the Inter-American Sugar Cane Seminar: Varieties and Breeding III, Louisiana, 96-98.

Melloto-Passarin DM, Calsar-Junior T and Carrer H (2004). Complete Chloroplast DNA Sequence of Sugarcane Reveals Similarity and Diversity Among *Saccharum* Species, Sorghum, Maize, Rice and Wheat. Plant Genome XIl Conference Book, 124. Available at [http://www.intlpag.org/2013/index.php/archives/pg-i-to-pag-xix-archives]. Accessed January 23, 2013.

Ming R, Liu SC, Lin YR, da Silva J, et al. (1998). Detailed alignment of *Saccharum* and *Sorghum* chromosomes: comparative organization of closely related diploid and polyploid genomes. *Genetics* 150: 1663-1682.

Neale DB and Sederoff RR (1989). Paternal inheritance of chloroplast DNA and maternal inheritance of mitochondrial DNA in loblolly pine. *Theor. Appl. Genet.* 77: 212-216.

Palmer JD, Jansen RK, Michaels HJ, Chase MW, et al. (1988). Chloroplast DNA variation and plant phylogeny. *Ann. MO. Bot. Garden* 75: 1180-1206.

Pan Y-B, Burner DM and Legendre BL (2000). An assessment of the phylogenetic relationship among sugarcane and related taxa based on the nucleotide sequence of 5S rRNA intergenic spacers. *Genetica* 108: 285-295.

Pan Y-B, Burner DM, Legendre BL and Grisham MP (2004a). An assessment of the genetic diversity within a collection of *Saccharum spontaneum* with RAPD-PCR. *Genet. Res. Crop Evol.* 51: 895-903.

Pan Y-B, Burner DM, Wei Q, Cordeiro GM, et al. (2004b). New *Saccharum* hybrids in *S. spontaneum* cytoplasm developed through a combination of conventional and molecular breeding approaches. *Plant Genet. Res.* 2: 131-139.

Pan Y-B, Tew TL, Schnell RJ and Richard EP (2006). Microsatellite DNA marker-assisted selection of *Saccharum spontaneum* cytoplasm-derived germplasm. *Sugar Tech.* 8: 23-29.

Piperidis G, Piperidis N and D'Hont A (2010). Molecular cytogenetic investigation of chromosome composition and transmission in sugarcane. *Mol. Genet. Genomics* 284: 65-73.

Takahashi S, Furukawa T, Asano T, Terajima Y, et al. (2005). Very close relationship of the chloroplast genomes among *Saccharum* species. *Theor. Appl. Genet.* 110: 1523-1529.

Testolin R and Cipriani G (1997). Paternal inheritance of chloroplast DNA and maternal inheritance of mitochondrial DNA in the genus *Actinidia*. *Theor. Appl. Genet.* 94: 897-903.

Tew TL (1987). New Varieties. In: Sugarcane Improvement through Breeding (Heinz DJ, ed.). Elsevier, New York, 559-594.

Tew TL (2003). World Sugarcane Variety Census - Year 2000. *Sugar Cane Intl. March/April* 12-18.

Weising K and Gardner RC (1999). A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. *Genome* 42: 9-19.