agronomy

MDPI

Editorial

Development and Integration of an SSR-Based Molecular Identity Database into Sugarcane Breeding Program

Yong-Bao Pan

United States Department of Agriculture-Agricultural Research Service, Southeast Area, Sugarcane Research Unit, 5883 USDA Road, Houma, LA 70360, USA; yongbao.pan@ars.usda.gov; Tel.: +1-985-853-3165

Academic Editor: Peter Langridge
Received: 14 March 2016; Accepted: 20 April 2016; Published: 25 April 2016

Abstract: Sugarcane breeding is very difficult and it takes 12 to 14 years to develop a new cultivar for commercial production. This is because sugarcane varieties are highly polyploid, inter-specific hybrids with 100 to 130 chromosomes that may vary across geographical areas. Other obstacles/constraints include the small size of flowers that may not synchronize but may self-pollinate, difficulty in distinguishing hybrids from self progenies, extreme (G × E) interactive effect, and potential variety mis-identification during vegetative propagation and varietal exchange. To help cane breeders circumvent these constraints, a simple sequence repeats (SSR)-based molecular identity database has been developed at the United States Department of Agriculture-Agricultural Research Service, Sugarcane Research Unit in Houma, LA. Since 2005, approximately 2000 molecular identities have been constructed for clones of sugarcane and related *Saccharum* species that cover geographical areas including Argentina, Australia, Bangladesh, China, Colombia, India, Mexico, Pakistan, South Africa, Thailand, USA (Louisiana, Florida, Texas, and Hawaii), and Venezuela. The molecular identity database is updated annually and has been utilized to: (1) provide molecular descriptors to newly registered cultivars; (2) identify in a timely fashion any mislabeled or unidentifiable clones from cross parents and field evaluation plots; (3) develop de novo clones of energy cane with *S. spontaneum* cytoplasm; (4) provide clone-specific fingerprint information for assessing cross quality and paternity of polycross; (5) determine genetic relatedness of parental clones; (6) select F_1 hybrids from (elite × wild) or (wild × elite) crosses; and (7) investigate the inheritance of SSR markers in sugarcane. The integration of the molecular identity database into the sugarcane breeding program may improve the overall efficacy of cultivar development and commercialization.

Keywords: sugarcane breeding; SSR; molecular identity database

Generally speaking, there are probably nine key issues that affect both the productivity and the sustainability of sugarcane agriculture and integrated industry. These issues are land, fertility, water, variety, planting density, crop protection, cultural practices, harvesting and processing, and recently, computer information technology [1]. To all sugarcane farmers, it remains of top-most concern to grow the right cultivars. While it is the duty of conventional breeders to develop desirable sugarcane cultivars, biotechnologists can contribute greatly to the variety development process (crossing, selection, and evaluation) through the development and application of molecular breeding tools. Conventional sugarcane breeding is probably the most difficult job of any crop, due to the fact that sugarcane cultivars (*Saccharum* spp. hybrids) are highly polyploidy inter-specific hybrids containing 100 to 130 chromosomes [2,3]. The number of chromosomes may vary across geographical areas. Other obstacles/constraints include small flower size, the development of the flower which may not synchronize between crossing parents, the likelihood of self-pollination, the difficulty in

visually distinguishing F_1 hybrids from self progenies, the extreme genotype × environment or G × E interactive effect, and potential variety mis-identification during vegetative propagation and varietal exchange, *etc.* [3,4]. It takes 12 to 14 years to develop a new sugarcane variety upon selection and evaluation against about 20 traits that include high tonnage, high sugar yield, early maturity, low fiber, harvest-ability, cold tolerance, ratooning ability, and resistance to a number of disease and insect pests [5].

Applied biotechnology projects were initiated at the United States Department of Agriculture-Agricultural Research Service, Sugarcane Research Unit, Houma, Louisiana, USA in 1994, in research areas such as molecular evaluation of germplasm, development of species- and trait-specific DNA markers, genetic linkage mapping, microsatellite or simple sequence repeats (SSR) DNA marker-based molecular identity database, transgenic (GMO) sugarcane, and inheritance of molecular markers [1]. A sugarcane molecular identity database has been developed based on a panel of 21 polymorphic microsatellite (SSR) DNA markers (Table 1). These SSR markers were developed by the Sugarcane Microsatellite Consortium (SMC) supported by the International Consortium of Sugarcane Biotechnologists (ICSB) with 13 institution members. These include four institutions of Australia, namely, the Sugar Research of Australia (formerly the Bureau of Sugar Experiment Station), Centre of Plant Conservation Genetics, Commonwealth Science Industrial Research Organization, and the University of Queensland, the former Copersucar of Brazil, the Cenicaña of Colombia, the CIRAD of France, the Mauritius Sugar Industrial Research Institute, the Philippines Sugar Research Institute, South Africa Sugar Experiment Station, and three members from USA, namely, the American Sugar Cane League, the former Florida Sugar Cane League, and the former Hawaiian Sugar Planters' Association [6,7]. Unlike morphological traits that may vary due to (G × E) interactions, our research showed that the SSR DNA markers-based molecular identities are stable across years and geographic locations. The molecular identity of a sugarcane clone is defined by the presence (labeled as "A") or absence ("C") of 144 DNA distinctive fingerprints/fragments/alleles amplifiable from the clone's genomic DNA through PCR in a sequential order (Figure 1). The molecular identity of a sugarcane cultivar is unique and remains the same regardless of when or where the cultivar is grown. The quality and reliability of sugarcane molecular identities are ensured by a high throughput SSR genotyping platform, which utilizes leaf DNA samples, a liquid-handling robot, 384-well microplate, blue or green or yellow fluorescence-labeled PCR primers, red fluorescence-labeled DNA size markers, and a capillary electrophoresis (CE)-based DNA Sequencer [4]. Robust, yet distinctive, fluorescence peaks or SSR alleles are revealed from the CE files with genotyping software, either "GeneMapper" (Applied Biosystems, Inc. Foster City, CA, USA), or "GeneMarker" (SoftGenetics, LLC. College Station, PA, USA). Unlike agarose- or polyacrylamide-gel electrophoresis, during the CE process, each sample is run with 15 red fluorescence-labeled size standards in the range of 35 to 500 base pairs for accurate size calibration. Figure 2 shows one example, where eight polymorphic DNA fingerprints were amplified through PCR with a SSR primer pair SMC336BS from the 12 most recent Louisiana sugarcane cultivars. The sizes of these DNA fingerprints are 154, 166, 167, 169, 171, 175, 177, and 183 base pairs (bp), respectively, at a resolution power of just one base pair between 166 bp and 167 bp. Only two cultivars, L 01-299 and L 03-371, share four DNA fingerprints of 166, 169, 171, and 175 bp. Each of the other 10 cultivars has its unique DNA fingerprints.

Table 1. Name, repeat motif, nucleotide sequence of forward and reverse primer, and annealing temperature of 21 sugarcane microsatellite markers.

SSR Name	Repeat Motif	Forward Primer Sequence (5′ to 3′) Reverse Primer Sequence (5′ to 3′)	Annealing Temp (°C)
SMC119CG	$(TTG)_{12}$	TTCATCTCTAGCCTACCCCAA AGCAGCCATTTACCCAGGA	58
SMC1604SA	$(TGC)_7$	AGGGAAAAGGTAGCCTTGG TTCCAACAGACTTGGGTGG	58
SMC18SA	$(CGA)_{10}$	ATTCGGCTCGACCTCGGGAT AGTCGAAAGGTATAATAGTGTTAC	62
SMC24DUQ	$(TG)_{13}$	CGCAACGACATATACACTTCGG CGACATCACGGAGCAATCAGT	64
SMC278CS	$(TG)_{19}$ $(AG)_{25}$	TTCTAGTGCCAATCCATCTCAGA CATGCCAACTTCCAAACAGACT	64
SMC31CUQ	$(TC)_{10}$ $(AC)_{22}$	CATGCCAACTTCCAATACAGACT AGTGCCAATCCATCTCAGAGA	62
SMC334BS	$(TG)_{36}$	CAATTCTGACCGTGCAAAGAT CGATGAGCTTGATTGCGAATG	60
SMC336BS	$(TG)_{23}(AG)_{19}$	ATTCTAGTGCCAATCCATCTCA CATGCCAACTTCCAAACAGAC	62
SMC36BUQ	$(TTG)_7$	GGGTTTCATCTC TAGCCTACC TCAGTAGCAGAGTCAGACGCTT	64
SMC486CG	$(CA)_{34}$	GAAATTGCCTCCCAGGATTA CCAACTTGAGAATTGAGATTCG	58
SMC569CS	$(TG)_{37}$	GCGATGGTTCCTATGCAACTT TTCGTGGCTGAGATTCACACTA	62
SMC7CUQ	$(CA)_{10}$ $(C)_4$	GCCAAAGCAAGGGTCACTAGA AGCTCTATCAGTTGAAACCGA	60
SMC597CS	$(AG)_{31}$	GCACACCACTCGAATAACGGAT AGTATATCGTCCCTGGCATTCA	64
SMC703BS	$(CA)_{12}$	GCCTTTCTCCAAACCAATTAGT GTTGTTTATGGAATGGTGAGGA	62
SMC851MS	$(AG)_{29}$	ACTAAAATGGCAAGGGTGGT CGTGAGCCCACATATCATGC	58
mSSCIR66	$(GT)_{43}GC$ $(GT)_6$	AGGTGATTTAGCAGCATA CACAAATAAACCCAATGA	48
mSSCIR3	$(GT)_{28}$	ATAGCTCCCACACCAAATGC GGACTACTCCACAATGATGC	60
SMC1751CL	$(TGC)_7$	GCCATGCCCATGCTAAAGAT ACGTTGGTCCCGGAACCG	60
SMC22DUQ	$(CAG)_5C$ $(AGG)_5$	CCATTCGACGAAAGCGTCCT CAAGCGTTGTGCTGCCGAGT	62
mSSCIR43	$(GT)_3(AT)_2(GT)_{29}$	ATTCAACGATTT TCACGAG AACCTAGCAATTTACAAGAG	52
mSSCIR74	$(CGC)_9$	GCGCAAGCCACACTGAGA ACGCAACGCAAAACAACG	54

I	SMC119CG					SMC1604SA						SMC18SA					SMC24DUQ								
	106	112	118	128	131	109	112	115	118	121	124	137	140	144	147	150	126	128	131	135	137	142			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22			
	5					6						5					6								
II	SMC278CS											SMC31CUQ													
	140	153	166	168	170	174	176	178	182	138	150	160	162	163	165	167	171	173	177	179					
	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42					
	9											11													
III	SMC334BS					SMC336BS											SMC36BUQ								
	146	149	151	161	163	164	141	154	164	166	167	169	171	173	175	177	183	112	118	121					
	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62					
	6					11											3								
IV	SMC486CG					SMC569CS						SMC7CUQ													
	224	227	237	239	241	167	170	210	219	222		158	162	164	166	168	170								
	63	64	65	66	67	68	69	70	71	72		73	74	75	76	77	78								
	5					5						6													
V	SMC597CS											SMC703BS													
	144	148	154	157	159	161	163	164	165	168	174	201	206	208	210	212	214	216	220	222					
	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98					
	11											9													
VI	SMC851MS						mSSCIR66					mSSCIR3													
	128	130	132	134	136	141	127	130	132	134		141	145	171	173	175	177	178	180	182	187				
	99	100	101	102	103	104	105	106	107	108		109	110	111	112	113	114	115	116	117	118				
	6						4					10													
VII	SMC1751CL						SMC22DUQ							mSSCIR43											
	140	144	147	151	154	125	148	151	154	157	160	163	206	209	233	235	237	239	248	250	252				
	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139				
	5					7							9												
VIII	mSSCIR74																								
	217	220	223	226	229																				
	140	141	142	143	144																				
	5																								

Figure 1. Formulation of 144 microsatellite (SSR) DNA fingerprints-based sugarcane molecular identity. These fingerprints are amplifiable from sugarcane genomic DNA through PCR using one of the 21 simple sequence repeats (SSR) markers listed in Table 1. Within each section (I through VIII shown on the left), name of the SSR marker (row 1), allele size (base pairs) (row 2), sequential order (row 3), and number of allele per marker (row 4) are shown. The presence or absence of each of the 144 SSR DNA fingerprints in a sugarcane cultivar, when combined, constitutes its molecular identity.

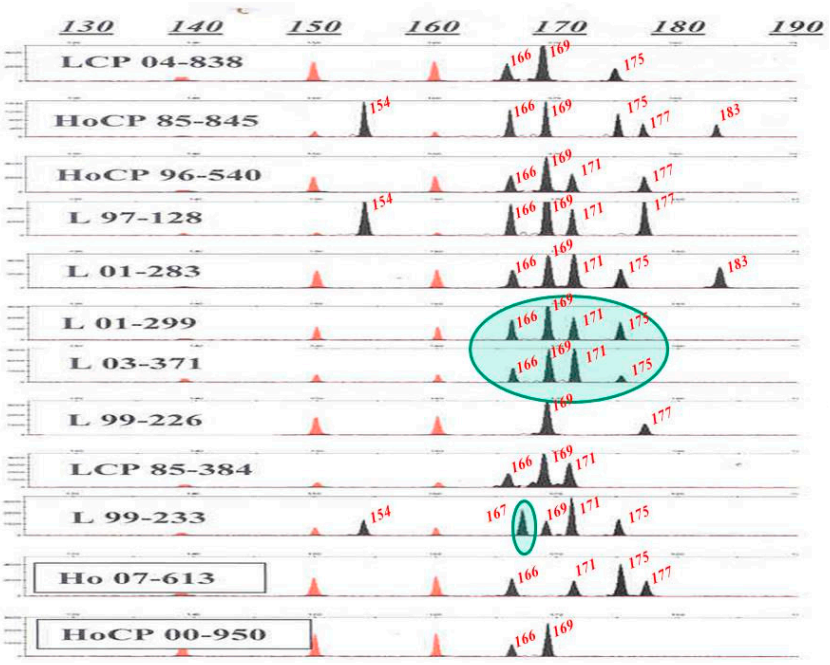

Figure 2. Capillary electrophoregrams of eight microsatellite (SSR) DNA fingerprints (filled in black color) from 12 Louisiana sugarcane cultivars, namely (from top), LCP 04-838, HoCP 85-845, HoCP 96-540, L 97-128, L 01-283, L 01-299, L 03-371, L 99-226, LCP 85-384, L 99-233, HoCP 07-613, and HoCP 00-950, and three DNA size markers, 139, 150, and 160 base pairs (filled in red color). The fingerprints of cultivars are amplified from the genomic DNA of leaf tissue through PCR primed by SMC336BS primer pair. The values shown on top are size marks and the values shown on left represent relative fluorescence intensity strength or the relative yield of amplified DNA fragment. The size of these eight SSR DNA fingerprints is 154, 166, 167, 169, 171, 175, 177, and 183 bp, respectively.

Since 2005, SSR-based clone-specific molecular identities have been constructed for over 2000 clones of sugarcane cultivars and/or related *Saccharum* species (*S. officinarum*, *S. spontaneum*, *S. robustum*, *S. barberi*, *S. sinense*, and *S. edule*) [3]. Fingerprinted sugarcane cultivars cover many geographical areas including Argentina, Australia, Bangladesh, China, Colombia, India, Mexico, Pakistan, South Africa, Thailand, the USA (Louisiana, Florida, Texas, and Hawaii), and Venezuela. These molecular identities have been successfully utilized to promote the efficiency of the conventional sugarcane breeding program at the United State Department of Agriculture-Agricultural Research Service, Sugarcane Research Unit:

Firstly, the geneticists and breeders are able to include molecular descriptors to registration articles on both sugarcane and energy cane (*) cultivars; for example, HoCP 96-540 [8], Ho 95-988 [9], Ho 00-950 [10], HoCP 91-552* [11], Ho 00-961* [12], and Ho 02-113* [13] were all registered with a molecular descriptor.

Secondly, the geneticists and breeders are able to identify mis-labeled clones in a timely fashion, and to remove mis-labeled clones from the crossing carts or field evaluation plots [14].

Thirdly, molecular breeding gives the geneticists and breeders an option to hot water emasculate *S. spontaneum* plants by immersing the flowers in 50 °C circulating water bath for 5 min and cross the *S. spontaneum* plants as female with superior sugarcane cultivars as male parents. The geneticists and breeders then screen the resulting seedlings by cultivar-specific SSR fingerprints to identify true F_1 hybrids for field evaluation and selection and discard self-progeny and off-type progeny derived from stray pollens of unknown source [15]. Within 12 years, the USDA-ARS sugarcane geneticists and breeders were able to release the first energy cane cultivar Ho 02-113 [13] that contains a cytoplasm of SES234, a *S. spontaneum* clone.

Fourthly, the geneticists and breeders have been successful in using clone-specific SSR fingerprints [15] for several purposes including identifying true F_1 hybrids from several other (elite × wild) or (wild × elite) crosses [16], assessing the genetic relatedness of parental clones [17], determining the paternity of polycross progeny [18], and assessing cross quality.

Lastly, basic genetic studies are also being conducted at the USDA-ARS, Sugarcane Research Unit involving both pollens (gamete) and self- and cross-progenies (zygote) [19–21]. The geneticists and breeders found that the inheritance of SSR DNA fingerprints is in accordance with the Mendelian laws of segregation and independent assortment and that non-parental SSR fingerprints are encountered very rarely (only 1 out of 2392 PCR-based genotyping reactions). To ensure cross quality, breeders may enforce pollen control by trimming dehisced female flowers followed by hot water treatment and surrounding crosses with cubicles on all sides to prevent stray pollens. If pollen control is not enforced, the breeders may encounter non-parental SSR fingerprints more often.

In sugarcane breeding, crossing, evaluation, and selection are a revolving process with newly selected clones being assigned each year. These newly assigned sugarcane clones need to be fingerprinted using the same protocol. The resulting molecular identities are then added to the database. It is anticipated that before long, the sugarcane geneticists and breeders will be able to use the molecular identity database information to assess the reliability of sugarcane pedigree information recorded in their notebooks.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pan, Y.-B. Biotechnology: Impact on sugarcane agriculture and industry. *Sugar Tech* **2012**, *14*, 1–2. [CrossRef]
2. Wu, J.; Huang, Y.; Lin, Y.; Fu, C.; Liu, S.; Deng, Z.; Li, Q.; Huang, Z.; Chen, R.; Zhang, M. Unexpected Inheritance Pattern of *Erianthus arundinaceus* Chromosomes in the Intergeneric Progeny between *Saccharum* spp. and *Erianthus arundinaceus*. *PLoS ONE* **2014**, *9*, e110390. [CrossRef] [PubMed]
3. Pan, Y.-B. Databasing molecular identities of sugarcane (*Saccharum* spp.) clones constructed with microsatellite (SSR) DNA markers. *Am. J. Plant Sci.* **2010**, *1*, 87–94. [CrossRef]

4. Pan, Y.-B.; Scheffler, B.S.; Richard, E.P., Jr. High throughput genotyping of commercial sugarcane clones with microsatellite (SSR) DNA markers. *Sugar Tech* **2007**, *9*, 176–181.
5. Bischoff, K.P.; Gravois, K.A. The development of new sugarcane varieties at the LSU Ag center. *J. Am. Sugar Cane Technol.* **2004**, *24*, 142–164.
6. Cordeiro, G.M.; Taylor, G.O.; Henry, R.J. Characterisation of microsatellite markers from sugarcane (*Saccharum* sp.), a highly polyploid species. *Plant Sci.* **2000**, *155*, 161–168. [CrossRef]
7. Pan, Y.-B. Highly polymorphic microsatellite DNA markers for sugarcane germplasm evaluation and variety identity testing. *Sugar Tech* **2006**, *8*, 246–256. [CrossRef]
8. Tew, T.L.; White, W.H.; Legendre, B.L.; Grisham, M.P.; Dufrene, E.O.; Garrison, D.D.; Veremis, J.C.; Pan, Y.-B.; Richard, E.P., Jr.; Miller, J.D. Registration of 'HoCP 96–540' Sugarcane. *Crop Sci.* **2004**, *45*, 785–786.
9. Tew, T.L.; Burner, D.M.; Legendre, B.L.; White, W.H.; Grisham, M.P.; Dufrene, E.O.; Garrison, D.D.; Veremis, J.C.; Pan, Y.-B.; Richard, E.P., Jr. Registration of 'Ho 95–988' Sugarcane. *Crop Sci.* **2005**, *45*, 1660–1661. [CrossRef]
10. Tew, T.L.; Dufrene, E.O., Jr.; Garrison, D.D.; White, W.H.; Grisham, M.P.; Pan, Y.-B.; Richard, E.P., Jr.; Legendre, B.L.; Miller, J.D. Registration of 'HoCP 00–950' sugarcane. *J. Plant Regist.* **2009**, *3*, 42–50. [CrossRef]
11. Tew, T.L.; Dufrene, E.O., Jr.; Cobill, R.M.; Garrison, D.D.; White, W.H.; Grisham, M.P.; Pan, Y.-B.; Legendre, B.L.; Richard, E.P., Jr.; Miller, J.D. Registration of 'HoCP 91–552' sugarcane. *J. Plant Regist.* **2011**, *5*, 181–190. [CrossRef]
12. White, W.H.; Cobill, R.M.; Tew, T.L.; Burner, D.M.; Grisham, M.P.; Dufrene, E.O., Jr.; Pan, Y.-B.; Richard, E.P., Jr.; Legendre, B.L. Registration of 'Ho 00–961' Sugarcane. *J. Plant Regist.* **2011**, *5*, 332–338. [CrossRef]
13. Hale, A.L.; Dufrene, E.O.; Tew, T.L.; Pan, Y.-B.; Viator, R.P.; White, P.M.; Veremis, J.C.; White, W.H.; Cobill, R.; Richard, E.P.; *et al.* Registration of 'Ho 02–113' Sugarcane. *J. Plant Regist.* **2011**, *7*, 51–57. [CrossRef]
14. Pan, Y.-B.; Miller, J.D.; Schnell, R.J.; Richard, E.P.; Wei, Q. Application of microsatellite and RAPD fingerprints in the Florida sugarcane variety program. *Sugar Cane Int.* **2003**, *189*, 19–28.
15. Chandra, A.; Grisham, M.P.; Pan, Y.-B. Allelic divergence and cultivar-specific SSR alleles revealed through fluorescence-labeled SSR markers based capillary electrophoregrams in highly polyploid species sugarcane. *Genome* **2014**, *57*, 363–372. [CrossRef] [PubMed]
16. Pan, Y.-B.; Tew, T.L.; Schnell, R.J.; Viator, R.P.; Richard, E.P., Jr.; Grisham, M.P.; White, W.H. Microsatellite DNA marker-assisted selection of *Saccharum spontaneum* cytoplasm-derived germplasm. *Sugar Tech* **2006**, *8*, 23–29. [CrossRef]
17. Pan, Y.-B.; Cordeiro, G.M.; Richard, E.P., Jr.; Henry, R.J. Molecular genotyping of sugarcane clones with microsatellite DNA markers. *Maydica* **2003**, *48*, 319–329.
18. Tew, T.L.; Pan, Y.-B. Microsatellite (simple sequence repeat) marker-based paternity analysis of a seven-parent sugarcane polycross. *Crop Sci.* **2010**, *50*, 1401–1408. [CrossRef]
19. Pan, Y.-B.; Liu, P.; Que, Y. Independently segregating simple sequence repeats (SSR) alleles in polyploid sugarcane. *Sugar Tech* **2015**, *17*, 235–242. [CrossRef]
20. Lu, X.; Zhou, Z.; Pan, Y.-B.; Chen, C.Y.; Zhu, J.; Chen, P.H.; Li, Y.-R.; Cai, Q.; Chen, R.K. Segregation analysis of microsatellite (SSR) markers in sugarcane polyploids. *Genet. Mol. Res.* **2015**, *14*, 18384–18395. [CrossRef] [PubMed]
21. Liu, P.; Chandra, A.; Que, Y.; Chen, P.-H.; Grisham, M.P.; White, W.H.; Dalley, C.D.; Tew, T.L.; Pan, Y.-B. Identification of QTLs controlling sucrose content based on an enriched genetic linkage map of sugarcane (*Saccharum* spp. hybrids) cultivar 'LCP 85–384'. *Euphytica* **2016**, *207*, 527–549. [CrossRef]