Sugar Tech
https://doi.org/10.1007/s12355-019-00788-9

RESEARCH ARTICLE

SSR-Based Genetic Identity of Sugarcane Clones and its Potential Application in Breeding and Variety Extension

Zhoutao Wang[1] · Yongbao Pan[2] · Jun Luo[1] · Qian You[1] · Liping Xu[1] · Hua Zhang[1] · Youxiong Que[1]

Received: 17 September 2019 / Accepted: 14 December 2019

Abstract Sugarcane variety regional test and integrated demonstration play an important role in identifying new varieties of high yield, disease resistance and wide adaptability. In this study, 15 SSR primer pairs were used to assess the genetic diversity among 68 sugarcane clones involved in three cycles of national regional tests and four cycles of integrated demonstrations in China. In total, 141 DNA fragments of 100 to 350 bp in length were identified, of which 139 fragments (98.58%) were polymorphic. Clustering analysis of UPGMA algorithm based on the Nei genetic similarity coefficient divided the 68 sugarcane clones into five groups. Group I only had one clone YT00-318 due to its high heterogeneity. Groups II, III, VI and V contained 4, 6, 5 and 52 sugarcane clones, respectively. A small subgroup A in group V was identified at the genetic similarity coefficient 0.890 that contained ROC22, FN07-3206, FN40, GT09-12 and LC07-150. ROC22 was a check variety with wide adaptability, high cane yield, high sugar and several other excellent characteristics; the other four clones might have a high potential of release by sharing the same excellent traits. Principal component analysis showed that the 68 sugarcane clones within quadrants I, II, III and IV showed a high homogeneity, and no series of clones obviously gathered together. The SSR fingerprint information of the 68 sugarcane clones has been drawn into a SSR fingerprint map for the identification of sugarcane clones in Chinese sugarcane breeding programs.

Keywords Sugarcane · Regional test · Integrated demonstration · SSR marker · Genetic distance · Variety extension · Breeding

Introduction

Since sugarcane is the most important sugar crop and also is a potential energy crop, the demand of sugarcane production and research continues to increase (Luo et al. 2015; Que et al. 2014). Factors such as population growth, production cost, climate change, reduced cultivation land area and aging population have put food and energy security into a serious challenge to the mankind (Menhas et al. 2016; Kumar 2016; Matsumoto 2015). One of the feasible ways to tackle this challenge is to continuously improve the yield and quality of crop varieties.

From the production and development process in rice and corn, we have learned that every major grain yield improvement has been due to the exploration and utilization of key germplasm resources. In Yunnan province of China, more than 2000 sugarcane germplasm were collected and are being maintained at the National Sugarcane Germplasm Resource Nursery of China. However, there are more germplasm resources in the USA and India, with more than 4000 and 6000 clones, respectively (Chen et al. 2011). Such large germplasm resources contain huge genetic potential, and a thorough evaluation of these germplasm resources provides a prerequisite basis for sugarcane breeding. Morphological, cellular, biochemical

✉ Jun Luo
sisluojun@126.com

✉ Youxiong Que
queyouxiong@126.com

[1] Key Laboratory of Sugarcane Biology and Genetic Breeding, Ministry of Agriculture, Fujian Agriculture and Forestry University, ShangXiadian Road Num. 15, Cangshan District, Fuzhou City 350002, Fujian Province, China

[2] Sugarcane Research Unit, USDA-ARS, Houma, LA 70360, USA

Published online: 10 January 2020

and molecular markers have been widely used in crop germplasm identification in recent years. They provide an effective basis for the search, identification and classification of plant crop germplasm (Erskine and Muehlbauer 1991; Nayak et al. 2005; Li et al. 2015). Praveen et al. (2015) reported a Sugarcane Germplasm Database (SGDB) in 2015, and all sugarcane germplasm in this database are characterized by biochemical, cytological, morphological and agronomic traits including disease and insect resistance. The database can improve the screen efficiency of hybrid parents greatly.

Unlike DNA markers, non-DNA markers have some limitations in their application, especially the assessment of population genetic diversity. For example, the number of non-DNA markers may be limited, or the expression of non-DNA markers may be influenced by environments and crop development stage (Ran et al. 2010). Up to date, several types of DNA markers are available, including random amplified polymorphic DNA (RAPD), amplified fragment length polymorphisms (AFLP), single nucleotide polymorphisms (SNP), simple sequence repeats (SSR) (Gardner et al. 2013; Andrade et al. 2015; Pan 2006) and so on. These DNA markers have provided accurate, economic and efficient ways to assess the genetic diversity and phylogenetic relationship of germplasm resources and to construct genetic linkage maps (Ganie et al. 2015). Among them, SSR markers, which are also called microsatellites, contain simple sequence repeats of dinucleotides, trinucleotides, tetranucleotides, or more (the number of nucleotides generally of one to six and the times of repetition of 15 to 60). SSR markers have shown a good level of polymorphism and distribute throughout the eukaryotic genomes (Tautz 1989; Edwards et al. 1991). With the advantages of simplicity, repeatability, co-dominance and richness, SSR markers have been widely used in plants up to now, including genetic diversity analysis (Baert-Desurmont et al. 2016; Amar et al. 2011), gene tagging (Cuming 2016; Dubey et al. 2009), mapping (Andru et al. 2011; Tan et al. 2013) and pedigree analysis (Dreisigacker et al. 2004; Romero-Severson et al. 2001). In addition, SSRs may also involve several biological functions, including the regulation of the transcription factors binding and enhancer functions (Martin et al. 2005; Rockman and Wray 2002), the regulation of mRNA stability (Chen et al. 2007), nucleosome positioning (Gymrek et al. 2016) and so on. Pan (2006) evaluated 221 SSR primer pairs developed by the International Sugarcane Microsatellite Consortium on five US sugarcane clones and found 67 primer pairs (30%) are highly polymorphic with PIC values ranging from 56 to 80%, which provided a basis for the SSR evaluation of sugarcane genetic diversity.

The objective of China's Sugarcane National Regional Tests and Integrated Demonstrations is to identify the best varieties with high and stable yields and good quality. Reliable reference data from the tests and demonstrations will provide reference to promote good varieties and determine suitable planting regions (Luo et al. 2014). Most of the sugarcane varieties cultivated in the word today can be traced back to only a few common ancestors (Deren 1995). Partially because of the genetic bottleneck effect, the rate of genetic gains through sugarcane crossing has been slow (Edmé et al. 2005). The national regional tests and integrated demonstrations of sugarcane varieties can provide quantitative data on certain characters, which are valuable to the breeders. Besides, when the breeders select crossing parents from the local germplasm collection, it would be helpful to know the genetic relationship among clones of the germplasm collection and predict the promotion potential of new varieties (You et al. 2016). In the present study, SSR markers were used to assess the genetic diversity among 68 Chinese sugarcane clones involved in three cycles of regional tests and four cycles of integrated demonstrations, which were subjected to SSR fingerprint collection for the first time. A SSR fingerprint map of the 68 sugarcane clones was constructed, which provided high-density information useful to identify clones quickly and conveniently estimate genetic differences among clones and so on.

Materials and Methods

Plant Material and DNA Extraction

The 68 sugarcane clones involved in three cycles of national regional tests during 2014–2017 and four cycles of integrated demonstrations during 2010–2014 were grown at a field nursery of the Key Laboratory of Sugarcane Biology and Genetic Breeding, Ministry of Agriculture, Fujian Agriculture and Forestry University in 2015 (Table 1). At the sugarcane trefoil stage, we collected the leaf tissue from each clone for genomic DNA extraction by an optimized CTAB method (Yao et al. 2005). The concentration and quality of all DNA samples were detected by Synergy^{TM} HT Multi-Mode Microplate Reader, and all DNA samples were diluted to 50 ng/μL.

SSR-PCR Amplification and Detection

Based on previous research results (You et al. 2016; Wang et al. 2018), 15 primer pairs were selected for this study. All forward primers were labeled with fluorescence dye 6-FAM before SSR-PCR amplification, and sugarcane gDNA samples were sent to Beijing MicroRead Genetics Co., Ltd. (Beijing, China) for SSR amplification and capillary electrophoresis. PCR amplifications were conducted

Table 1 A list of 68 sugarcane clones involved in three cycles of national regional tests and four cycles of integrated demonstrations

No.	Clones	Parents	Breeding institutions[a]	Series[b]
1	FN07-2020	YT91-976/LCP85-384	SRI, FAFU, Fujian Province	FN
2	FN09-12206	CP65-357/YC97-40	SRI, FAFU, Fujian Province	FN
3	FN09-2201	ROC22/GT00-122	SRI, FAFU, Fujian Province	FN
4	FN09-4095	YT93-159/YZ91-790	SRI, FAFU, Fujian Province	FN
5	FN09-7111	GT96-44/ROC11	SRI, FAFU, Fujian Province	FN
6	FN1110	ROC20/YT91-976	SRI, FAFU, Fujian Province	FN
7	FN11-2105	CZ89-103/YR05-770	SRI, FAFU, Fujian Province	FN
8	FN15	CP72-1210/YN73-204	SRI, FAFU, Fujian Province	FN
9	FN38	YT83-257/YT83-271	SRI, FAFU, Fujian Province	FN
10	FN39	YT91-976/CP84-1198	SRI, FAFU, Fujian Province	FN
11	FN40	FN93-3406/YT91-976	SRI, FAFU, Fujian Province	FN
12	FN07-3206	90-1211/77-979	SRI, FAFU, Fujian Province	FN
13	MT02-205	YC90-3/ROC10	SRI, AAS, Fujian Province	MT
14	MT06-1405	MT92-649/ROC10	SRI, AAS, Fujian Province	MT
15	MT07-2005	YC73-512/ROC22	SRI, AAS, Fujian Province	MT
16	MT09-104	GT90-420/ROC10	SRI, AAS, Fujian Province	MT
17	GT02-351	CP80-1827/ROC10	SRI, AAS, Guangxi Province	GT
18	GT02-467	ROC23/CP84-1198	SRI, AAS, Guangxi Province	GT
19	GT02-901	ROC23/CP84-1198	SRI, AAS, Guangxi Province	GT
20	GT06-1492	CP72-1210/ZZ92-126	SRI, AAS, Guangxi Province	GT
21	GT06-2081	GT00-122/YC97-47	SRI, AAS, Guangxi Province	GT
22	GT08-1180	ROC26/ROC22	SRI, AAS, Guangxi Province	GT
23	GT08-1533	ZZ90-76/GT94-116	SRI, AAS, Guangxi Province	GT
24	GT09-12	ROC24/YN79-780	SRI, AAS, Guangxi Province	GT
25	GT29	YC94-46/ROC22	SRI, AAS, Guangxi Province	GT
26	GT30	YT91-976/ROC11	SRI, AAS, Guangxi Province	GT
27	GT31	YT85-177/CP81-1254	SRI, AAS, Guangxi Province	GT
28	GT32	YT91-976/ROC1	SRI, AAS, Guangxi Province	GT
29	GT40	YN86-295/CP84-1198	SRI, AAS, Guangxi Province	GT
30	GT44	ROC1/GT92-66	SRI, AAS, Guangxi Province	GT
31	GT97-69	YT73-204/ROC1	SRI, AAS, Guangxi Province	GT
32	LC03-1137	HoCP93/746/ROC22	SRC, Liucheng County, Guangxi Province	LC
33	LC03-182	HoCP93-746/ROC22	SRC, Liucheng County, Guangxi Province	LC
34	LC05-129	CP81-1254/ROC22	SRC, Liucheng County, Guangxi Province	LC
35	LC05-136	CP81-1254/ROC22	SRC, Liucheng County, Guangxi Province	LC
36	LC07-150	YT85-177/ROC22	SRC, Liucheng County, Guangxi Province	LC
37	LC07-500	YT92-1287/CP72-1210	SRC, Liucheng County, Guangxi Province	LC
38	LC07-506	YT85-177/ROC22	SRC, Liucheng County, Guangxi Province	LC
39	YZ01-1413	YT85-177/ROC10	SRI, AAS, Yunnan Province	YZ
40	YZ03-103	YZ91-976/CP85-1432	SRI, AAS, Yunnan Province	YZ
41	YZ03-194	ROC25/Y97-20	SRI, AAS, Yunnan Province	YZ
42	YZ05-49	YC90-56/ROC23	SRI, AAS, Yunnan Province	YZ
43	YZ08-1095	CP84-1198/K5	SRI, AAS, Yunnan Province	YZ
44	YZ08-1609	YZ94-343/YT00-236	SRI, AAS, Yunnan Province	YZ
45	YZ08-2060	YT93-159/Q121	SRI, AAS, Yunnan Province	YZ
46	YZ09-1028	YR05-178/MT86-2121	SRI, AAS, Yunnan Province	YZ
47	YZ09-1601	CP94-1110/CT89-103	SRI, AAS, Yunnan Province	YZ
48	YZ99-596	Co419/YC85-881	SRI, AAS, Yunnan Province	YZ

Table 1

49	YZ99-91	ROC10/YC84-153	SRI, AAS, Yunnan Province	YZ
50	YR07-1433	YR99-155/L75-20	SRI, AAS, Yunnan Province	YZ
51	YR09-315	CL69-52YR05-285	SRI, AAS, Yunnan Province	YZ
52	YR10-187	ROC20/YR05-282	SRI, AAS, Yunnan Province	YZ
53	YR10-701	YR08-18/YR05-701	SRI, AAS, Yunnan Province	YZ
54	DZ07-36	GT92-66/CP67-412	SRI, Dehong Prefecture, Yunnan Province	DZ
55	DZ09-78	GT94-119/ROC10	SRI, Dehong Prefecture, Yunnan Province	DZ
56	DZ09-84	GT94-119/ROC10	SRI, Dehong Prefecture, Yunnan Province	DZ
57	YG43	YT93-213/YT93-159	SIRI, Guangdong Academy of Science	YT
58	YG46	YT00-236/GT96-211	SIRI, Guangdong Academy of Science	YT
59	YG47	YN73-204/ROC22	SIRI, Guangdong Academy of Science	YT
60	YG48	HoCP95/YT97-76	SIRI, Guangdong Academy of Science	YT
61	YG50	YT96-86/YT99-66	SIRI, Guangdong Academy of Science	YT
62	HZ22	YT93-159/ROC22	SIRI, Guangdong Academy of Science	YT
63	YT00-236	YN73-204/CP72-1210	SIRI, Guangdong Academy of Science	YT
64	YT00-318	YN73-204/CP86-1633	SIRI, Guangdong Academy of Science	YT
65	YT55	YN73-204/CP72-1210	SIRI, Guangdong Academy of Science	YT
66	YT96-86	YT85-177/ZZ74-141	SIRI, Guangdong Academy of Science	YT
67	GZ07-538	ROC10/CP57-614	SRI, Jiangxi Province	–
68	ROC22		Sugar Processing Research Institute, Taiwan	–

[a]Breeding institutions: *SRI*, Sugarcane Research Institute; *FAFU*, Fujian Agriculture and Forestry University; *AAS*, Academy of Agricultural Sciences; *SRC*, Sugarcane Research Center; *SIRI*, Sugarcane Industry Research Institute

[b]Series: variety series symbols. *FN*, the series that includes all sugarcane clones bred by the SRI, FAFU, Fujian Province; *MT*, the series that includes all sugarcane clones bred by the SRI, AAS, Fujian Province; *GT*, the series that includes all sugarcane clones bred by the SRI, AAS, Guangxi Province; *LC*, the series that includes all sugarcane clones bred by the SRC, Liucheng County, Guangxi Province; *YZ*, the series that includes all sugarcane clones bred by the SRI, AAS, Yunnan Province; *DZ*, the series that includes all sugarcane clones bred by the SRI, Dehong Prefecture, Yunnan Province; *YT*, the series that includes all sugarcane clones bred by the SIRI, Guangdong Academy of Science

in a 20 μL reaction volume, containing about 25 ng of gDNA, 8.0 μL 2.5 × buffer V, 3 μM of each primer, 1 U rTaq. SSR amplification products were separated through capillary electrophoresis. Then, 1.0 μL tenfold-diluted amplified products with different length differences and fluorescent labeling were mixed with the internal standard of the standard molecular weight sample (0.5 μL ROX-500 size) and placed in the same lane in capillary electrophoresison analyzer ABI 3730XLDNA (Applied Biosystems inc., Foster City, CA, USA) to produce Genescan files, which were processed by GeneMapper® V3.0 (Applied Biosystems inc., Foster City, CA, USA) software to reveal and size calibration of fluorescence-labeled fingerprints. The GeneMapper® parameters were: PreRun Module: GS PR36A-2, 400, Plate Check Module: Plant Check A; Collect time: 2.5 h; Run Module: GS Run 36A-2400; Lanes: 64.

Data Collection and Statistical Analysis

The results of capillary electrophoresis were stored in a PDF file, in which each absorption peak represents an amplified fragment with a calibrated size. First, we need to identify a locus, and at least one band on this locus belongs to a specific clone. Then, we sort all the bands by matrix format, and every band was scored as "1" for the presence of the band, while "0" in the absence of the band (Ran et al. 2010). All 0/1 information is recorded manually in Excel spreadsheet. NTSYS-pc 2.10e software was used to calculate genetic similarity coefficient (GS) and generate clustering figure. Powermarker v3.25 software was used to calculate the genetic diversity parameters PIC (Tian and Wang 2007; Huang et al. 2010). The percentage of polymorphic bands (PPB) was calculated using the following formula:

$$\mathrm{PPB} = \frac{\mathrm{NPB}}{\mathrm{NTB}} \times 100\%, \tag{1}$$

where NTB represents the number of total bands; NPB represents the number of polymorphic bands. The PIC (polymorphism information content) of primer pairs was calculated according to (http://statgen.ncsu.edu/powermarker/downloads/Manual.pdf):

$$\mathrm{PIC}_l = 1 - \sum_{u=1}^{k} p_{lu}^2 - \sum_{u=1}^{k-1} \sum_{v=u+1}^{k} 2p_{lu}^2 p_{lv}^2, \tag{2}$$

where PIC represents the genetic diversity parameters that was calculated from the 68 sugarcane clones, and u and v represent the frequencies of uth and vth alleles, respectively. The k represents the number of alleles.

Results

PIC Values and Number of SSR Alleles Amplified

Based on previous studies of You et al. (2016) and Wang et al. (2018), 141 amplified alleles were selected according to the amplification quality and size range of 100 to 350 bp. The number of polymorphic alleles was 139 (98.58%, Table 2). Except primer pairs SEP8 and SMC569CS, the PPB of other primer pairs were 100%, and the average PPB of all primer pairs reached 98.58%. In addition, the PIC values were also at a very high level with an average of 0.933.

Genetic Distance-Based Clustering Analysis

The cluster analysis of UPGMA algorithm based on the Nei genetic similarity coefficient is shown in Fig. 1. We divided the 68 sugarcane clones into five groups (I, II, III, IV and VI) according to the clustering results. Among the five groups, groups I (YT00-318) and II (YR09-315, YG48, FN09-7111 and FN09-4095) have greater heterogeneity, followed by III, IV and V, respectively. Moreover, a small subgroup A contained ROC22 at the genetic similarity coefficient of 0.890. Except for ROC22, the subgroup also contained clones FN07-3206, FN40, GT09-12 and LC07-150. Due to the fact of ROC22 has wide adaptability, high yield, high sugar and other excellent characteristics, those clones in the same subgroup might be more likely to have these excellent characteristics and higher promotion potential. In addition, we found that ROC22 as the main sugarcane variety cultivated in China was clustered with most sugarcane clones in group V, in which the heterogeneity is low. The principal reason is that the main grown cultivars are often used as hybrid parents, and this is shown in Table 1.

Principal Component Analysis (PCA)

In two-dimensional principal component analysis (Fig. 2), we analyzed the similarity of the genetic background of different series of clones. Only, we found that the GT series (blue dots) was mainly distributed in the first, second, and third quadrants, and the YT series (yellow dots) were mainly distributed in the first and fourth quadrants. Overall, all series of clones had more or less intertwined, which indicates that the genetic background of the 68 sugarcane clones is very close, or every breeding institution has no preference when choosing hybrid breeding parents.

Table 2 Number of total bands, number of polymorphic bands, percentage of polymorphic bands and polymorphism information content of 15 primer pairs of SSR maker

Primers	NTB[a]	NPB[b]	PPB[c] (%)	PIC[d]
SEP6	6	6	100	0.975
SPE8	13	12	92.31	0.969
SEP17	8	8	100	0.965
SEP23	21	21	100	0.984
SEP59	10	10	100	0.969
SEP70	8	8	100	0.952
SEP84	5	5	100	0.851
SEP89	7	7	100	0.823
SMC334BS	8	8	100	0.962
SMC336BS	9	9	100	0.986
SMC286CS	9	9	100	0.923
SMC569CS	8	7	87.50	0.865
SMC119CG	10	10	100	0.838
SMC31CUQ	9	9	100	0.957
mSSCIR43	10	10	100	0.983
Total/average	141	139	98.58	0.933

[a]The number of total bands

[b]The number of polymorphic bands

[c]The percentage of polymorphic bands

[d]The polymorphism information content

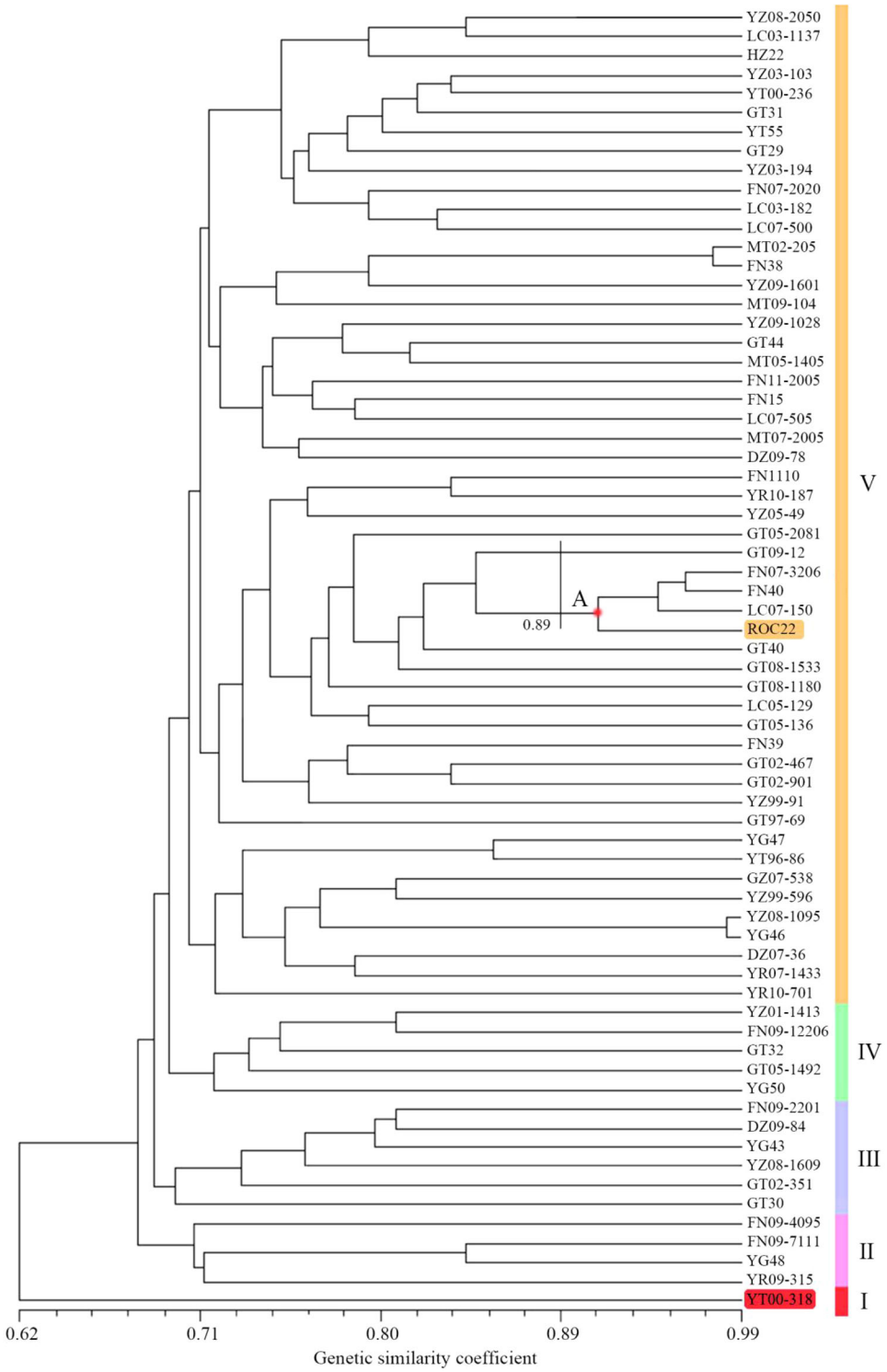

Fig. 1 Cluster analysis of UPGMA algorithm based on the Nei genetic similarity coefficient of the 68 sugarcane clones. The 68 sugarcane clones were divided into five subgroups I, II, III, IV and V. The heterogeneity of subgroup I is the highest, and followed by II, III, IV and V, respectively. The control clone ROC22 is in a small group A in subgroup V

A SSR Fingerprint Map with High Information Density

We also constructed a SSR fingerprint map of the 68 sugarcane clones (Fig. 3), which carries information of all polymorphic loci. On the right side of Fig. 3, there are 139 polymorphic loci, and the names of all clones are on the bottom of Fig. 3. In this SSR fingerprint map, each clone can be well separated by specific fingerprint combination.

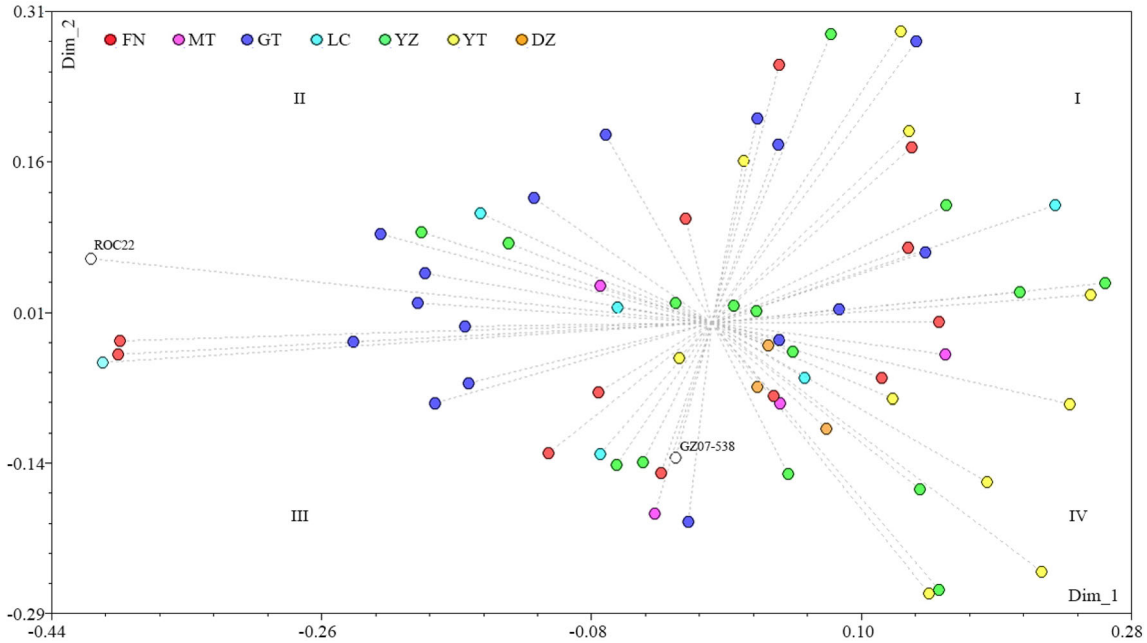

Fig. 2 Two-dimensional principal component analysis of the 68 sugarcane clones belonging to the seven series, namely FN (red), MT (carmine), GT (blue), LC (sky blue), YZ (green), YT (yellow) and DZ (orange) (color figure online)

Discussion

Due to ease of operation, high repeatability and polymorphism, SSR markers have been frequently used in genetic diversity analysis of sugarcane germplasm (Pan 2006; Perera et al. 2012; Payne 2013). The study of Liu et al. (2018) indicated that the estimated gene diversity in *Lolium perenne* L. was the highest using SSR markers, followed by SNP and DArT markers, which illustrated the superiority of SSR markers in the analysis of species genetic diversity. Similarly, SSR marker was more informative compared to ISSR marker in determination of gene diversity, polymorphic information content (PIC) and heterozygosity in an *Etlingera elatior* population with 57 members (Ismail et al. 2019). When using SSR maker, the selection of effective primer pairs is a key step for getting accurate and reproducible results. SSR markers with higher PIC values will have a relatively higher probability in detecting genetic variability (Chandra et al. 2014). Even though PIC values may change with different populations of test material, these values can be used to assess the potential usefulness of a molecular marker (Arkova et al. 2015; Pan 2006). The 15 SSR primer pairs used in the present study were selected from a large number of primer pairs with high PIC values (0.933 mean) and capability of producing of distinctive fragments, and they are very effective in detecting genetic variation among the 68 sugarcane clones. This is why they were selected as standard primers to establish the SSR fingerprint map, and that was also partly confirmed by the high accuracy in genetic distance analysis of sugarcane germplasm and classification of these germplasm.

It is known from Table 1 that clones LC03-1137 and LC03-182 are full siblings originated from the same male (ROC22) and female (HoCP93-746) parents. Clone HZ22 shared the same male parent (ROC22) with LC03-182 and LC03-1137. The female parent of HZ22 is YT93-159. Seemingly, the genetic distance between clones LC03-1137 and LC03-182 would be closer than the genetic distance between clones HZ22 and LC03-1137 or between clones HZ22 and LC03-182 according to the fact that clones LC03-1137 and LC03-182 share common parents. After all, the parents of YT93-159 are YN73-204 and CP72-1210 and are different from HoCP93-746, and as a result, YT93-159 and HoCP93-746 do not have very close genetic backgrounds so that their offspring crossed with the same variety, respectively, are also very closely related. However, the actual molecular maker detection results are contrary to the theoretical inference in accordance with the pedigree (Fig. 4): The genetic distance between clones HZ22 and LC03-1137 (0.7801) or between clones HZ22 and LC03-182 (0.8298) is closer than the genetic distance between clones LC03-1137 and LC03-182 (0.6950). This phenomenon requires a reasonable explanation.

From the genetics perspective, Fig. 4 shows this contradiction. When the genotypes of clones YT93-159, ROC22, and HoCP93-746 were supposed to be decaploid 6A4a10B5C5c, 10A8B2b5C5c and 8A2a2B8b5A5b, respectively, the genotypes of their offspring clones HZ22, LC03-1137 and LC03-182 may be 8A2a10B10C,

Fig. 3 An SSR fingerprint map of 68 sugarcane clones. The names of the 139 polymorphic loci are shown on the right side, and the names of all clones are shown on the bottom. All black boxes indicate the presence of the SSR loci, and the white boxes indicate the absence of the SSR loci

10A10B10c and 8A2a10b10C, respectively. This is a very reasonable explanation. But, even without considering variety mis-identity, or mix up, or accuracy of pedigree records, this genotypic difference can only be detected by

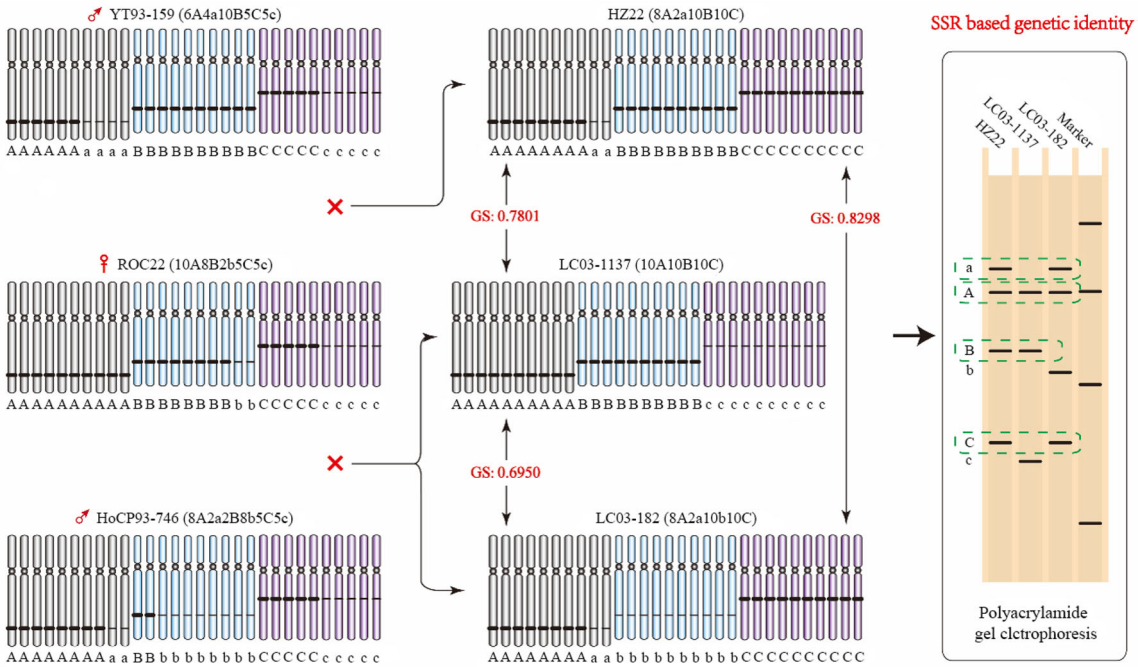

Fig. 4 The illustration why the genetic distance of the varieties identified by SSR markers is more accurate than the pedigree method

DNA molecular maker (such as SSR, AFLP, RFPD and ISSR) rather than pedigree or phenotype relationship. Therefore, even if the genetic distance between sugarcane varieties determined based on the pedigree is credible in most cases, it is also a requisite to combine the SSR molecular maker information. This argument is supported by Lima et al. (2002), who used 79 sugarcane varieties to compare genetic relationships assessed by pedigree relationship and genetic similarity coefficient based on DNA molecular maker, and they concluded that DNA molecular maker can provide more information about genetic similarity among varieties than pedigree.

Regional test and integrated demonstration can evaluate sugarcane varieties from the angle of production characteristics (Wang et al. 2016). These characters include sugar yield, sugar content, disease resistance and plant height. These data are a very important reference for evaluation and promotion of sugarcane varieties, which make our study realistic and significant. What should be stressed here is that the sugarcane cultivars from Taiwan, for instance, ROC22, account for more than 80% in all planting areas in Mainland China in prior years, resulting in a short harvesting season with low average sugar yield and serious diseases and pests in many areas (Luo et al. 2014). This suggests that the germplasm simplification has been a serious threat to sucrose production. Before we extend new sugarcane varieties, their high yield, stability and regional adaptability calculated by AMMI and GGE-biplot model (Luo et al. 2015; Wang et al. 2016) are the first factors we need to consider, but it is better for us to consider their genetic background. If several new sugarcane varieties have highly similar genetic background (commercial varieties in particular), such a narrow genetic basis is difficult to adapt to the environmental fluctuations and these varieties may more likely have similar resistance or tolerance patterns to biotic or abiotic stress (Zhang et al. 2006). In addition, in evaluating the extension potential of new varieties, we should consider the genetic similarity between the newly bred varieties and the existing widely recognized varieties. The greater the similarity, the more likely the new varieties are to be close to the existing excellent varieties in adaptability, yield and sugar content, and the more likely they are to be accepted by the growers and increase its promotion opportunities (Wang et al. 2016). In this study, a small subgroup of Group V was identified at the genetic similarity coefficient 0.890 that contained ROC22, FN07-3206, FN40, GT09-12 and LC07-150. ROC22 was a check variety with wide adaptability, high cane yield, high sugar and several other excellent characteristics; the other four clones might have a high potential of release because of sharing the same excellent traits.

Scientific and efficient utilization of sugarcane germplasm resources will greatly increase the possibility of breeding new varieties of more excellence. Parental selection is the main and basic step of breeding. The effective identification of hybrid parents is the basic work of cross-breeding, which plays an important guiding role in the selection of parents (Benin et al. 2012; Yao et al. 2016). In breeding, the selection of sugarcane germplasm resources with high genetic differences can improve the

heterozygosity of the genotypes of the hybrid generation. The genetic distance assessment of crop germplasm resources is helpful for identifying and combining the best parents, producing offspring with the largest genetic variation and promoting the fine genes of different germplasm resources to penetrate into the new breeding lines, so as to achieve the goal of obtaining the ideal heterosis (Longin et al. 2011; Yao et al. 2016). Previous studies have shown that the genetic distance of parents is positively related to the heterosis of F_1, and the possibility of heterosis in new bred lines can be evaluated in advance (Wegary et al. 2013). For example, the results shown in Figs. 1 and 2 reveal the sugarcane clone similarity of genetic background from the angle of molecular marker and provided further reference information for promoting new sugarcane varieties. Besides, the present study also provides a reference for breeders to decide where to plant these clones. In Fig. 1, we found that YT00-318 has the highest heterogeneity among the 68 sugarcane clones and is of great value in breeding and variety extension.

Using a SSR-CE-based detection system, Chandra et al. (2014) detected 213 alleles in 24 sugarcane cultivars (12 each from India and the USA), and in addition to the CoLK9606 and CoS95255, the remaining varieties have their own unique SSR fingerprints. The reason for CoLk9606 and CoS95255 to share exactly the same SSR fingerprints is most probably due to the misplacement or mislabeling of clones (Chandra et al. 2014). However, this problem can also be solved by the SSR-CE-based detection system. For instance, Pan et al. (2003) have successfully identified clones that had been misidentified during a field trial, such as CP96-1602 or LCP85-384. Therefore, SSR fingerprinting is an effective tool to identify sugarcane clones. The SSR fingerprint map constructed in this paper contains high-density information (Fig. 3). It does not require complex storage facilities and a large amount of storage resources. For small germplasm populations, this information storage mode is very practical. We can quickly find the fingerprints of a sugarcane clone according to the order of clone names, and then compare it with other clones to estimate their genetic distance. We can also quickly confirm the identity of a clone based on its SSR loci information. Of course, there is another way to store SSR fingerprint information, that is, to build an SSR fingerprint database in a computer. It is no doubt that the fingerprint database can also be helpful to provide reference for the selection of breeding parent, protect the intellectual property rights of breeders, prevent fake seed cane flooding in the market and avoid disturbing the order of business. It is also a means of collection and utilization of germplasm resources (Pan 2010; Gao et al. 2012). Pan (2010) pioneered the first sugarcane molecular identity database with identities constructed since 2005. There are two characteristics about this database. First, the variety molecular identity was in the form of a nucleotide sequence, where "A" = presence of a specific SSR allele and "C" = absence of a specific SSR allele. Second, because of the needs for rigorous identification in his database, multiple samples of the same clone may be collected from up to four different locations, either in the same or in different years.

Conclusion

Genetic diversity assessment is an essential component of germplasm characterization and use. In this study, the genetic diversity among 68 valuable sugarcane clones involved in three cycles of national regional tests and four cycles of integrated demonstrations in China was assessed with 15 SSR primer pairs. A total of 141 DNA fragments were identified, of which 139 fragments (98.58%) were polymorphic. UPGMA algorithm-based clustering analysis placed the 68 sugarcane clones into five groups, from which several clones were found with a high level of heterogeneity, especially the clone YT00-318. Overall, however, the genetic background of these 68 clones was narrow and limited. Using the 139 polymorphic SSR fingerprints, a dense fingerprint map was drawn that can be used to quickly confirm the identity of each of the 68 clones.

Funding This work was funded by the Public Welfare Industrial Research Project of China (201003009-2), the 948 Program on the Introduction of International Advanced Agricultural Science and Technique of Department of Agriculture (2014-S18) and the China Agriculture Research System (CARS-17). The funders had no role in the study design, data collection and analysis, decision to publish or preparation of the manuscript.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Amar, M.H., M.K. Biswas, Z. Zhang, and W.W. Guo. 2011. Exploitation of SSR, SRAP and CAPS-SNP markers for genetic diversity of citrus germplasm collection. *Scientia Horticulturae* 128(3): 220–227. https://doi.org/10.1016/j.scienta.2011.01.021.

Andru, S., Y.B. Pan, S. Thongthawee, D.M. Burner, and C.A. Kimbeng. 2011. Genetic analysis of the sugarcane (*saccharum* spp.) cultivar 'LCP 85-384'. I. linkage mapping using AFLP, SSR, and TRAP markers. *Theoretical and Applied Genetics* 123(1): 77–93. https://doi.org/10.1007/s00122-011-1568-x.

Andrade, J.C.F.D., J. Terto, J.V. Silva, and C. Almeida. 2015. Expression profiles of sugarcane under drought conditions: variation in gene regulation. *Genetics and Molecular Biology*

38(4): 465–469. https://doi.org/10.1590/S1415-475738420140288.

Arkova, O.V., M.P. Ponomarenko, D.A. Rasskazov, I.A. Drachkova, T.V. Arshinova, and P.M. Ponomarenko. 2015. Obesity-related known and candidate SNP markers can significantly change affinity of tata-binding protein for human gene promoters. *BMC Genomics* 16(13 Supplement): S5. https://doi.org/10.1186/1471-2164-16-S13-S5.

Baert-Desurmont, S., F. Charbonnier, E. Houivet, L. Ippolito, J. Mauillon, and M. Bougeard. 2016. Clinical relevance of 8q23, 15q13 and 18q21 SNP genotyping to evaluate colorectal cancer risk. *European Journal of Human Genetics* 24: 99–105. https://doi.org/10.1038/ejhg.2015.72.

Benin, G., G. Matei, D.O.A. Costa, G.O. Silva, T.R. Hagemann, and D.S.C. Lemes. 2012. Relationships between four measures of genetic distance and breeding behavior in spring wheat. *Genetics and Molecular Research* 11(3): 2390–2400. https://doi.org/10.4238/2012.June.15.3.

Chandra, A., M.P. Grisham, and Y.B. Pan. 2014. Allelic divergence and cultivar-specific SSR alleles revealed by capillary electrophoresis using fluorescence-labeled SSR markers in sugarcane. *Genome* 57(6): 363–372. https://doi.org/10.1139/gen-2014-0072.

Chen, T.M., P.L. Kuo, C.H. Hsu, S.J. Tsai, M.J. Chen, C.W. Lin, and H.S. Sun. 2007. Microsatellite in the 3' untranslated region of human fibroblast growth factor 9 (*FGF9*) gene exhibits pleiotropic effect on modulating FGF9 protein expression. *Human Mutation* 28(1): 98. https://doi.org/10.1002/humu.9471.

Chen, R.K., L.P. Xu, and Y.Q. Lin. 2011. *Modern sugarcane genetic breeding*. Beijing: China Agriculture Press.

Cuming, A.C. 2016. Gene tagging in physcomitrella patens: An addition to the genetic toolbox. *New Phytologist* 21: 543–545. https://doi.org/10.1111/nph.14201.

Deren, C.W. 1995. Genetic base of U.S. mainland sugarcane. *Crop Science* 35(4): 1195–1199. https://doi.org/10.2135/cropsci1995.0011183X003500040047x.

Dreisigacker, S., P. Zhang, M.L. Warburton, M. Van Ginkel, D. Hoisington, and M. Bohn. 2004. SSR and pedigree analyses of genetic diversity among cimmyt wheat lines targeted to different megaenvironments. *Crop Science* 44(2): 381–388. https://doi.org/10.2135/cropsci2004.3810.

Dubey, L., B.M. Prasanna, and B. Ramesh. 2009. Analysis of drought tolerant and susceptible maize genotypes using SSR markers tagging candidate genes and consensus QTLs for drought tolerance. *Indian Society of Genetics and Plant Breeding* 69(4): 344–351. https://doi.org/10.1002/humu.21099.

Edmé, S.J., J.D. Miller, B. Glaz, P.Y.P. Tai, and J.C. Comstock. 2005. Genetic contribution to yield gains in the Florida sugarcane industry across 33 years. *Crop Science* 45(1): 92–97. https://doi.org/10.2135/cropsci2005.0092.

Edwards, A., A. Civitello, H.A. Hammond, and C.T. Caskey. 1991. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *American Journal of Human Genetics* 49: 746–756. https://doi.org/10.1016/0378-1119(92)90168-O.

Erskine, W., and F.J. Muehlbauer. 1991. Allozyme and morphological variability, outcrossing rate and core collection formation in lentil germplasm. *Theoretical and Applied Genetics* 83(1): 119–125. https://doi.org/10.1007/bf00229234.

Ganie, S.H., P. Upadhyay, S. Das, and M.P. Sharma. 2015. Authentication of medicinal plants by DNA markers. *Plant Gene* 4: 83–99. https://doi.org/10.1016/j.plgene.2015.10.002.

Gao, P., H. Ma, F. Luan, and H. Song. 2012. DNA fingerprinting of Chinese melon provides evidentiary support of seed quality appraisal. *PLoS ONE* 7(12): e52431. https://doi.org/10.1371/journal.pone.0052431.

Gardner, S.N., J.B. Thissen, K.S. Mcloughlin, T. Slezak, and C.J. Jaing. 2013. Optimizing SNP microarray probe design for high accuracy microbial genotyping. *Journal of Microbiological Methods* 94(3): 303–310. https://doi.org/10.1016/j.mimet.2013.07.006.

Gymrek, M., T. Willems, A. Guilmatre, H. Zeng, B. Markus, S. Georgiev, M.J. Daly, A.L. Price, J.K. Pritchard, A.J. Sharp, and Y. Erlich. 2016. Abundant contribution of short tandem repeats to gene expression variation in humans. *Nature Genetics* 48: 22–29. https://doi.org/10.1038/ng.3461.

Huang, M., F.M. Xie, L.Y. Chen, X.Q. Zhao, L. Jojee, and D. Madonna. 2010. Comparative analysis of genetic diversity and structure in rice using ILP and SSR markers. *Rice Science* 2010(4): 14–25.

Ismail, N.A., M.Y. Rafii, T.M.M. Mahmud, M.M. Hanafi, and G. Miah. 2019. Genetic diversity of torch ginger *(Etlingera elatior)* germplasm revealed by ISSR and SSR markers. *BioMed Research International* 2019: 1–14. https://doi.org/10.1155/2019/5904804.

Kumar, M. 2016. Impact of climate change on crop yield and role of model for achieving food security. *Environmental Monitoring and Assessment* 188(8): 1–14. https://doi.org/10.1007/s10661-016-5472-3.

Li, H., S.Y. Ca, J. Niu, P. Yuan, D. Zhao, and F. Zhang. 2015. The types and application of molecular markers in the study of pomegranate germplasm resources. *Acta Horticulturae* 1089: 127–132. https://doi.org/10.17660/ActaHortic.2015.1089.14.

Lima, M.L.A., A.A.F. Garcia, K.M. Oliveira, S. Matsuoka, H. Arizono, C.L. de Souza Jr, and A.P. de Souza. 2002. Analysis of genetic similarity detected by AFLP and coefficient of parentage among genotypes of sugar cane (*saccharum* spp.). *Theoretical and Applied Genetics* 104(1): 30–38. https://doi.org/10.1007/s001220200003.

Liu, S., U. Feuerstein, W. Luesink, S. Schulze, T. Asp, B. Studer, H.C. Becker, and K.J. Dehmer. 2018. DArT, SNP, and SSR analyses of genetic diversity in *Lolium perenne* L using bulk sampling. *BMC Genetics* 19(1): 10. https://doi.org/10.1186/s12863-017-0589-0.

Longin, F.H., W.X. Liu, N. Ranc, and J.C. Reif. 2011. Association of progeny variance and genetic distances among parents and implications for the design of elite maize breeding programs. *Maydica* 56(3): 227–231. https://doi.org/10.1007/s10343-010-0235-5.

Luo, J., Y.B. Pan, L.P. Xu, M.P. Grisham, H. Zhang, and Y.X. Que. 2015. Rational regional distribution of sugarcane cultivars in china. *Scientific Reports* 5(1): 15721. https://doi.org/10.1038/srep15721.

Luo, J., Y.B. Pan, L.P. Xu, H. Zhang, Z.N. Yuan, Z.H. Deng, R.K. Chen, and Y.X. Que. 2014. Cultivar evaluation and essential test locations identification for sugarcane breeding in China. *Scientific World Journal* 2014: 302753. https://doi.org/10.1155/2014/302753.

Martin, P., K. Makepeace, S.A. Hill, D.W. Hood, and E.R. Moxon. 2005. Microsatellite instability regulates transcription factor binding and gene expression. *Proceedings of the National Academy of Sciences of the United States of America* 102(10): 3800–3804. https://doi.org/10.1073/pnas.0406805102.

Matsumoto, K. 2015. Energy structure and energy security under climate mitigation scenarios in China. *PLoS ONE* 10(12): e0144884. https://doi.org/10.1007/978-1-4471-5286-6_4.

Menhas, R., S. Umer, and G. Shabbir. 2016. Climate change and its impact on food and nutrition security in Pakistan. *Iranian Journal of Public Health* 45(4): 549–550.

Nayak, S., P.K. Naik, L. Acharya, A.K. Mukherjee, P.C. Panda, and P. Das. 2005. Assessment of genetic diversity among 16 promising cultivars of ginger using cytological and molecular

markers. *Zeitschrift fur Naturforschung C-A Journal of Biosciences* 60(5–6): 485–492. https://doi.org/10.1515/znc-2005-5-618.

Pan, Y.B., J.D. Miller, R.J. Schnell, E.P. Richard, Jr., and Q. Wei. 2003. Application of microsatellite and RAPD fingerprints in the Florida sugarcane variety program. *International Sugar Journal* 19–28.

Pan, Y.B. 2006. Highly polymorphic microsatellite DNA markers for sugarcane germplasm evaluation and variety identity testing. *Sugar Tech* 8(4): 246–256. https://doi.org/10.1007/bf02943564.

Pan, Y.B. 2010. Databasing molecular identities of sugarcane (*saccharum* spp.) clones constructed with microsatellite (SSR) DNA markers. *American Journal of Plant Sciences* 1(2): 87–94. https://doi.org/10.4236/ajps.2010.12011.

Payne, W.J. 2013. Genetic and physical characterisation of the locus controlling columnar habit in apple (*Malus* × *domestica* Borkh.). *Molecular Breeding* 31(2): 429–440. https://doi.org/10.1007/s11032-012-9800-1.

Perera, M.F., M.E. Arias, D. Costilla, A.C. Luque, M.B. García, C. Dı´az Romero, J. Racedo, S. Ostengo, M.P. Filippone, M.I. Cuenya, and A.P. Castagnaro. 2012. Genetic diversity assessment and genotype identification in sugarcane based on DNA markers and morphological traits. *Euphytica* 185(3): 491–510. https://doi.org/10.1007/s10681-012-0661-9.

Praveen, K., M. Hemanth Kumar, A. Umamaheshwari, D.M. Reddy, P. Sudhakar, M. Munikumar, D. Pradhan, and N. Sabita. 2015. SGDB: A sugarcane germplasm database. *Sugar Tech* 17(2): 150–155. https://doi.org/10.1007/s12355-014-0307-4.

Que, Y.X., Y.B. Pan, Y.H. Lu, C. Yang, Y.T. Yang, N. Huang, and L.P. Xu. 2014. Genetic analysis of diversity within a chinese local sugarcane germplasm based on start codon targeted polymorphism. *BioMed Research International* 2014(5): 7. https://doi.org/10.1155/2014/468375.

Ran, L., W.L. Wu, Z. Yang, and Y.H. Li. 2010. SSR marker and its application to crop genetics and breeding. *Genomics and Applied Biology* 29: 137–143. https://doi.org/10.3724/SP.J.1238.2010.00569.

Rockman, M.V. and G.A. Wray. 2002. Abundant raw material for cis-regulatory evolution in humans. *Molecular Biology and Evolution* 19(11): 1991–2004. https://doi.org/10.1093/oxfordjournals.molbev.a004023.

Romero-Severson, J., J.S.C. Smith, J. Ziegle, J. Hauser, L. Joe, and G. Hookstra. 2001. Pedigree analysis and haplotype sharing within diverse groups of *zea mays* L. inbreds. *Theoretical and Applied Genetics* 103(4): 567–574. https://doi.org/10.1007/pl00002911.

Tan, L.Q., L.Y. Wang, K. Wei, C.C. Zhang, L.Y. Wu, G.N. Qi, H. Cheng, Q. Zhang, Q.M. Cui, and J.B. Liang. 2013. Floral transcriptome sequencing for SSR marker development and linkage map construction in the tea plant (*Camellia sinensis*). *PLoS ONE* 8(11): e81611. https://doi.org/10.1371/journal.pone.0081611.

Tautz, D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research* 17(16): 6463–6471. https://doi.org/10.1093/nar/17.16.6463.

Tian, Y., and C. Wang. 2007. Genetic diversity analysis of *Brassica oleracea* L. by SSR. *Journal of Northeast Agricultural University (English Edition)* 14(3): 202–205.

Wang, Z.T., W.H. Su, Y.X. Que, L.P. Xu, H. Zhang, and J. Luo. 2016. Analysis of yield stability and test site representativeness of sugarcane trials using combined AMMI and HA-GGE biplot models. *Chinese Journal of Eco-Agriculture* 24(6): 790–800. https://doi.org/10.13930/j.cnki.cjea.151284.

Wang, Z.T., Q. You, S.W. Gao, C.F. Wang, Z. Li, J.J. Ma, Y.X. Que, L.P. Xu, and J. Luo. 2018. Identification of sugarcane varieties by AFLP and SSR markers and its application. *Acta Agronomica Sinica* 44(5): 723–736. https://doi.org/10.3724/SP.J.1006.2018.00723.

Wegary, D., B. Vivek, and M. Labuschagne. 2013. Association of parental genetic distance with heterosis and specific combining ability in quality protein maize. *Euphytica* 191(2): 205–216. https://doi.org/10.1007/s10681-012-0757-2.

Yao, J., H. Li, J. Ye, and L. Shi. 2016. Relationship between parental genetic distance and offspring's heterosis for early growth traits in liriodendron: Implication for parent pair selection in cross breeding. *New Forests* 47(1): 163–177. https://doi.org/10.1007/s11056-015-9508-2.

Yao, W., A.L. Yu, J.S. Xu, H. Zhou, M.Q. Zhang, and R.K. Chen. 2005. A simple and quick method for extracting sugarcane genomic DNA. *Journal of Agricultural Biotechnology* 13(1): 121–122.

You, Q., Y.B. Pan, L.P. Xu, S.W. Gao, Q.N. Wang, Y.C. Su, Y.Q. Yang, Q.B. Wu, D.G. Zhou, and Y.X. Que. 2016. Genetic diversity analysis of sugarcane germplasm based on fluorescence-labeled simple sequence repeat markers and a capillary electrophoresis-based genotyping platform. *Sugar Tech* 18(4): 380–390. https://doi.org/10.1007/s12355-015-0395-9.

Zhang, X., J. Yang, and R. Yang. 2006. Study of the relationship between genetic distance and heterosis in castor. *Scientia Agricultura Sinica* 39(3): 633–640.