

Assessing the Genetic Diversity and Population Structure of World Germplasm Collection of *Erianthus arundinaceus* (Retz.) Jeswiet Using Sequence-Related Amplified Polymorphic Markers

R. Valarmathi¹ · H. K. Mahadevaswamy¹ · V. Ulaganathan¹ · C. Appunu¹ · S. Karthigeyan¹ · Adhini S. Pazhany¹

Received: 5 February 2021 / Accepted: 23 July 2021 / Published online: 5 August 2021
© Society for Sugar Research & Promotion 2021

Abstract Genetic diversity of world collection of *E. arundinaceus* germplasm clones was analyzed using SRAP marker polymorphic survey. The marker survey with thirty combinations of SRAP marker generated distinct DNA profiles with fragments ranging in size from 120 to 1000 bp in all the germplasm lines. The selected primer pairs yielded a total of 306 scorable bands and 237 polymorphic bands. The mean percentage of polymorphic bands (PPB) among all the population was 75.9%. The maximum genetic distance (0.49) was measured between the Indonesian and Indian collections, and the minimum distance of 0.040 was measured between the population of Fiji and Philippines, and within the population of Philippines. The phylogenetic tree derived based on the dissimilarity index clustered 208 accessions into ten groups, and the clusters correlated with geographical origin of the clones. From the structure output using the SRAP data set, the best K value or the uppermost level of structure obtained was nine. The Mean Fixation Index (Fst) was 0.79 with a range from 0.88 to 0.70. Cluster 9 showed the highest Fst of 0.88, and cluster 1 showed the lowest Fst of 0.70, respectively. The structure analysis with high Fst and low gene flow (Nm) values indicated a lower gene flow and significant genetic differentiation among the populations. Proportion of individuals sharing > 60% membership within cluster ranged from 50 to 95.7. Blast analysis of selected fragment sequences showed homology to functional genes indicating the amplification of coding regions by the SRAP markers.

Keywords Germplasm · *Erianthus* · Diversity · Polymorphism · SRAP marker

Introduction

Sugarcane is an economically important crop accounts for more than 80% of world's sugar production. India is among the topmost leading country in the world in sugarcane production. The world's total sugarcane production was estimated at 140,169 lakh tonnes in 2019, and India's share with an average of 400 lakh tonnes contributed 24.47% of world's total production (ISMA 2020). India's high sugarcane production is partially contributed by the high-yielding varieties developed by premier ICAR-Sugarcane Breeding Institute which covers more than 75% of area in the country. ICAR-Sugarcane Breeding Institute maintains one of the largest sugarcane germplasm collections, which is recognised as 'World Collection' by the International Society of Sugar Cane Technologists (ISSCT). The others being at World Collection of Sugarcane and Related Grasses (WCSRG) in Miami, Florida, USA, National Nursery of Sugarcane Germplasm Resources (NNSGR), China, Centre of Sugarcane Technology (CTC) in Camamu, Brazil (Jackson and Henry 2011; Cai et al. 2012). These field genebanks harbours a huge number of accessions of *Saccharum*, and its closely related taxa such as, *Erianthus*, *Sclerostachya*, *Narenga* and *Miscanthus* and most of them are yet to be fully exploited. The world collections at ICAR-SBI represent major clones collected from India, Indonesia, Fiji, New Guinea, Pakistan, Burma and Philippines except Chinese collections.

The genus *Erianthus*, a member of *Saccharum* complex, belongs to the primary gene pool in sugarcane breeding. *Erianthus* is reported to be originated in the Indo-Burma

✉ R. Valarmathi
valarmathi.r@icar.gov.in

¹ Division of Crop Improvement, ICAR-Sugarcane Breeding Institute, Coimbatore 641007, India

Chinese region and then, spread to other parts of the world with majority of the species distributed in India (Mukherjee 1957). The genus has seven species; among which *Eri-
anthus arundinaceus* is the most widely distributed species reported from India, China, Burma, Thailand, Philippines, Malaysia, Indonesia and New Guinea (Dao et al. 2013). It is a cane forming species exhibiting vigorous growth, high biomass production and high tillering ability. With extensive root system, the species shows tremendous adaptability to wide ecological adaptation, excellent tolerance to various abiotic stresses and resistance to biotic stress (Nair et al. 2017; Wang et al. 2019). Hence, the species have been utilised by the breeders in sugarcane improvement programs. Successful hybridization has been carried out to generate intergeneric hybrids between *Saccharum* and *Eri-
anthus* (Cai et al. 2008; Nair et al. 2017). Recently *Eri-
anthus*, as a high biomass clone which grows well under marginal inputs, is utilized in developing high-fibre energy canes as bioenergy crop (Deng et al. 2002; Chang et al. 2012).

World germplasm collections of *E. arundinaceus* clones collected from all over India and from other countries maintained at ICAR-Sugarcane Breeding Institute, Coim-
batore, represents a diversified germplasm resource. In order to precisely select diversified parental line to be used in the breeding programs, a clear understanding on the population structure and genetic diversity is essential (Nair et al. 1999; Ming et al. 2006; Tew and Cobill 2008; Aitken et al. 2010; Cai et al. 2012; Zhang et al. 2013; Manechini et al. 2018; Mugehu et al. 2020). The information on genetic diversity and population structure will not only help in maintaining diversified germplasm bank but will also help in easy access of promising clones to be utilized in various genetic improvement applications.

Studies on understanding the genetic diversity of *E. arundinaceus* sugarcane germplasm used morphological traits which was later replaced by more authentic information obtained using molecular DNA markers. Based on the morphological traits, high genetic variations among the Chinese *E. arundinaceus* accession and a relatively low variation among the accessions of Indonesia have been reported (Berding and Koike 1980; Yang et al. 1997; Cai et al. 2002). Using DNA markers such as RAPD, SSR, RFLP and AFLP diversity studies of several *Eri-
anthus* accessions, similar result was obtained indicating a high molecular diversity among the Chinese population compared to the Indonesian *E. arundinaceus* population (Besse et al. 1997; Zhang et al. 2004, 2008, 2017; Cai et al. 2012; Tsuruta et al. 2012). A more recent study on the Chinese *Eri-
anthus* germplasm was carried out by Zhang et al. (2013) using a highly versatile and reproducible marker called sequence-related amplified polymorphism (SRAP). These markers are specifically designed to amplify coding

regions of the genome with ambiguous primers targeting GC-rich exons (forward primers) and AT-rich promoters, introns and spacers (reverse primers). SRAP markers are simple, inexpensive and more effective markers for producing genome-wide fragments and hence, serve as a better marker system for genetic diversity studies. The study using SRAP marker indicated a large genetic variation among the Chinese *Eri-
anthus* population. So far, only two studies have analyzed the genetic diversity among Indian *Eri-
anthus* germplasm accessions using RAPD and AFLP markers and reported to be more polymorphic compared to the *Eri-
anthus* clones of Indonesia (Nair and Mary 2006; Cai et al. 2012). However, only three to four samples were used in both the study to conclude the genetic diversity existing among the Indian population.

The current study aims at understanding the genetic diversity of world germplasm collection of *E. arundi-
naceus* existing at ICAR-Sugarcane Breeding Institute, Coimbatore, India, using SRAP markers. The germplasm consists of 208 accessions of *E. arundinaceus* representing major clones collected from India and clones from other countries such as Indonesia, Fiji, New Guinea, Pakistan, Burma and Philippines except Chinese collections. This is the first large-scale study on genetic diversity of world germplasm collection of *E. arundinaceus* maintained at ICAR-Sugarcane Breeding Institute, India, using functional marker system.

Material and Methods

Genetic Material

Germplasm accessions of 208 *E. arundinaceus* clones which are part of world germplasm collected from different geographical regions of India and other countries of the world (Indonesia, New Guinea, Fiji, Pakistan, Burma and Philippines) maintained at ICAR-Sugarcane Breeding Institute, Coimbatore, India, were used in the study (Table 1). The accessions were planted separately in 3 m rows and maintained for genetic characterization.

DNA Extraction and PCR Amplification

Young leaves from each accession were collected from well-established plants. The leaf samples were frozen immediately in ice and stored at -80°C until used for further DNA isolation. The leaf samples ground to a powder in liquid nitrogen and genomic DNA was extracted using CTAB method (Doyle 1991). Purity and quantity of extracted genomic DNA were estimated by Nanodrop and resolving in 1% (w/v) agarose gel.

Table 1 Details of 208 accessions of *E. arundinaceus* used in the study

| Sl. No | Country | Codes and clone names | No. of clones |
|--------|-------------|--|---------------|
| 1 | India | SES-3,7,17,27,75,79,89,133,136,149,153, 159,181,206,288,293,342, 347, E.a- Coimbatore, Pugalur, Cuttack, Gangtok, E.aB Munja, A and N islands, Lakshdweep, ENB, Rakki1, Rakki2, Rakki3, IND84-105, 85–547, 98–8, 99–871,873, 875, 877, 880, 884, 886, 888, 889, 890, 892, 895, 901,906, 907, 908, 956, 958, 965, 990, 991, 995, IND00-1013, 01–1091,1099,1101,1105,1107, 1110,1131,1134,1136, IND02-1201,1208, IND03-1223,1225,1228,1253,1255,1260,1262, IND04-1335, 1388,IND10-1575,1591,1594 | 80 |
| 2 | Indonesia | IK76-5,11,22,24,25,27,44,45,48,55,63,75,76,78,80,81, 88,90,91,92,93,99,101,103,105,109,111,IS76-120,126, 133,134,139,140,142,145,149,150,153,156,158,160,162,163, 169, 172,174,176,178,188,191,193,202,205,215,218,219,220 IM76-223,227,247,252,253,257, IMP-1536,1546,1547,2798 IJ76-327,332,333,334,340,341,342,345,346,347,357,358,359, 364,365,367, 370,374,379,381,383,384,388,389,390,394, 397,398,400,402,403,404,407,408,410, 411,476,502, 503, 508,511,513, IN84-17,83, Timor wild, Tongrang, US3-1,EaThornless | 117 |
| 3 | Pakistan | E.a Sarkender and E.a Lyalpur | 2 |
| 4 | Burma | Mythan A, Mythan B | 2 |
| 5 | Philippines | MINDANO, EC362-813, EC362-814, Eri 2384 | 4 |
| 6 | Fiji | Fiji 54 | 1 |
| 7 | New Guinea | NG77-188, NG7-28 | 2 |

Polymerase Chain Reaction (PCR) was carried out using 30 SRAP primer pair combinations which consisted of nine forward primers and eight reverse primers. The marker sequences are referred from Suman et al. (2008), and combinations of SRAP markers used for PCR amplification is given in Table 2. SRAP polymorphism analysis was conducted according to a previously established protocol (Li and Quiros 2001). PCR amplification reactions were performed in a Mastercycler (Eppendorf, Germany) using the following conditions: 5 min at 94 °C; 5 cycles of 94 °C, 1 min; 35 °C, 1 min; and 72 °C, 2 min; additional 35 cycles of 94 °C, 1 min; 50 °C, 1 min; and 72 °C, 1 min; extension of 5 min at 72 °C; and a final storage at 4 °C. PCR products were resolved using 3% (w/v) agarose gel. Clearly amplified PCR bands were visually scored for the presence (1) or absence (0) and then, were assembled into an excel matrix for genetic diversity analysis. Polymorphism Information Content (PIC) value was calculated using the formula, $PIC = 1 - \sum P_{ij}^2$ (Anderson et al. 1993), where P_{ij} is the frequency of the j th allele for i th locus summed across all alleles for the locus.

Genetic Distance and Phylogenetic Relationship Analysis

The number of polymorphic loci (N_p) and percentage of polymorphic bands (PPB) were calculated. Data analysis was carried out using DARwin to assess the pattern of grouping. The genetic distance and dissimilarity index were worked out based on simple matching. The phylogenetic tree based on dissimilarity index using WARD method was generated in DARwin 6.0. Bootstrap values were obtained by resampling replacements over loci in 1000 replicates.

Population Structure

Population structure among the *E. arundinaceus* population using the SRAP marker dataset was analyzed using the model-based program STRUCTURE (Pritchard et al. 2000). To infer population structure by a Bayesian approach, STRUCTURE 2.3.3 was run with the admixture model, and a burn-in period of 5000 and 50,000 Markov chain Monte Carlo was given. Ten iterations were performed for each K value, ranging from 1 to 10. The optimum K value was detected using Structure Harvester (an ad hoc summary statistic ΔK), which is based on the rate of

Table 2 List of SRAP marker combination and polymorphism obtained among the 208 accessions of *E. arundinaceus*

| Primer combinations | No. of bands obtained | No. of polymorphic bands | Percent polymorphic bands | PIC value |
|---------------------|-----------------------|--------------------------|---------------------------|-----------|
| E1M1 | 11 | 8 | 72.7 | 0.50 |
| E1M13 | 8 | 5 | 62.5 | 0.40 |
| E1M5 | 12 | 10 | 83.3 | 0.22 |
| E1M6 | 14 | 12 | 85.7 | 0.21 |
| E1M11 | 8 | 5 | 62.5 | 0.34 |
| E2M2 | 9 | 7 | 77.7 | 0.50 |
| E3M3 | 7 | 4 | 57.1 | 0.32 |
| E2M5 | 10 | 8 | 80 | 0.29 |
| E3M4 | 11 | 9 | 81.8 | 0.38 |
| E2M1 | 15 | 13 | 86.7 | 0.43 |
| E3M6 | 12 | 10 | 83.3 | 0.47 |
| E2M6 | 10 | 7 | 70 | 0.27 |
| E2M13 | 11 | 10 | 90.9 | 0.41 |
| E3M5 | 8 | 6 | 75 | 0.18 |
| E3M6 | 9 | 7 | 77.7 | 0.36 |
| E4M5 | 6 | 4 | 66.6 | 0.16 |
| E4M6 | 15 | 10 | 66.7 | 0.32 |
| E5M1 | 12 | 11 | 91.7 | 0.50 |
| E5M3 | 11 | 8 | 72.7 | 0.28 |
| E5M4 | 8 | 4 | 50 | 0.39 |
| E5M6 | 10 | 10 | 100 | 0.48 |
| E6M2 | 7 | 5 | 71.4 | 0.44 |
| E6M3 | 9 | 8 | 88.9 | 0.31 |
| E6M5 | 16 | 12 | 75 | 0.27 |
| E6M6 | 12 | 10 | 83.3 | 0.26 |
| E11M2 | 10 | 8 | 80 | 0.21 |
| E11M3 | 7 | 4 | 57.1 | 0.37 |
| E11M14 | 9 | 7 | 77.8 | 0.22 |
| E11M5 | 13 | 11 | 84.6 | 0.34 |
| E14M2 | 6 | 4 | 66.7 | 0.28 |
| Total | 306 | 237 | 75.9 | |

change in the log probability of data between successive K values (Evanno et al. 2005). The gene flow (Nm) value was calculated by the formula of $Nm = 0.25 \times (1-F_{st})/F_{st}$ (Zhou et al. 2015).

Sequencing of Selected Polymorphic Bands

Ten polymorphic bands of size 500–1000 bp were selected for sequence characterization. Amplified PCR products were resolved by 1% agarose gel electrophoresis, and the bands corresponding to the expected size were cut-out from the gel and eluted using gel extraction kit (M/s. Favorgen gel/PCR purification kit). The eluted products were checked again through agarose gel electrophoresis and

used for ligation. The eluted fragment was ligated into pTZ57R/T vector, and the ligated product was transformed into competent cells of *E. coli* (XL₁ Blue) using heat-shock method. The XL₁ Blue cells transformed with the plasmid DNA were selected on Luria Broth agar plates containing ampicillin (100 mg/l) and confirmed through PCR using vector specific universal M13 primers. Plasmid DNA was isolated from the positive transformants and subjected to Sanger sequencing. The sequencing of the positive plasmids was outsourced at Bioserve Biotechnologies India Pvt Ltd.

Results

To understand the genetic diversity among the world germplasm collection *E. arundinaceus* clones, a polymorphic survey using SRAP markers were carried out. An initial screening of 60 SRAP primer pairs (PP) was carried out on 10 *E. arundinaceus* clones to check the amplification. Out of 60 combinations, screened 30 PP combinations resulted in clear amplification and were used in the final survey. SRAP marker survey with 30 combinations of SRAP marker generated distinct DNA profiles with fragments ranging in size from 120 to 1000 bp in all the germplasm lines. Thirty selected PPs yielded a total of 306 scorable bands in the range of 6–16 bands. Out of 306 scorable bands, 237 (77.4%) bands were found to be polymorphic in the range of 4–13 bands in each PP. The mean percentage of polymorphic bands (PPB) among all the population was 75.9% with minimum percentage (50%) recorded in combination E5M4 and maximum (91.7%) recorded in E5M1 (Table 2). The PIC value varied for all the combinations tried ranging from 0.16 to 0.50 with an average value of 0.33. A PIC value of 0.50 was observed for the marker combinations E1M1, E2M2 and E5M1.

Genetic Distance and Phylogenetic Clustering of *E. arundinaceus* Accessions

The genetic distance among the *Erianthus* population was measured using simple matching method in DARwin (Supplementary Table 1). The distance showed clear segregation among the populations such as SES, IK76, IJ76, IS, IM and IND. The genetic distance among the population ranged from 0.01 to 0.60. The maximum distance was measured between IK76 (Indonesian) and SES (Indian) collections which ranged from 0.4 to 0.6, and the minimum distance of 0.21 to 0.40 was measured between the Indian population IND and SES clones. Among the other population, the genetic distance was in the range of 0.21–0.49. The maximum genetic distance (0.49) was measured between the Indonesian and Indian collections, and the minimum distance of 0.040 was measured between the population of Fiji and Philippines, and within the population of Philippines (Supplementary Table 2). The Indian population showed a higher genetic distance with the populations of all other countries.

The phylogenetic tree of 208 *E. arundinaceus* accessions generated using the dissimilarity index formed ten clusters (Fig. 1). The tree showed three major clusters (1, 2 and 3) and the major clusters 2 and 3 formed four sub-clusters (2a, 2b, 3a and 3b), the sub cluster 3b was found to be further divided into three sub-sub-clusters (3b1, 3b2 and 3b3). The major population of Indonesian collection

showed cluster based on the geographical origin. The IJ, IS, IM and IK collections reported to be collected from different geographical places of Indonesia clustered separately in the sub-clusters 2a, 2b and 3a. The first major cluster (1) was formed by the Indian accessions SES clones, while IND and other collections from India clustered separately in the sub-cluster 3b. The populations of New Guinea, Fiji, Philippines, Pakistan and Burma were clustered along with Indian collections in the sub-cluster 3b. Within the sub-cluster 3b, the accessions from New Guinea, Fiji, Philippines, Pakistan and Burma showed country-specific minor grouping in three sub-sub-clusters (3b1, 3b2 and 3b3). The germplasm accessions from Burma (Mythan A and Mythan B), New Guinea (NG77-188 and NG7-28) and Pakistan (E.a Sarkender and E.a Lyalpur) clustered along with other Indian accessions in the sub-sub-cluster 3b1. The germplasm accessions from Philippines (MINDANO, EC362-813, 814, Eri 2384) and Fiji (Fiji 54) formed a small sub-group in the sub-sub-cluster 3b3, while the accessions from India (IND) formed a major group in the same sub-sub-cluster 3b3. Very few intermixing population was observed in the clustering such as two IND clones IND 98–8 and IND 85–547 were found to cluster in between the Indonesian IJ collections. Three Indonesian clones Timor wild, Tongrang and US3-1 were grouped along with other accessions in the sub-sub-cluster 3b1.

Erianthus Population Structure Inferred from SRAP Marker Data

From the structure output using the SRAP data set, the best K value or the uppermost level of structure obtained was nine (Table 3). The K value indicated nine optimal clusters existing among the 208 germplasm clones. Each of the nine optimal clusters had a considerable proportion of memberships sharing among the clusters with very few admixed clones. Major proportion of a particular accession in a cluster is labelled in Fig. 2. Cluster 1 contained intermixed population from New Guinea, Pakistan and Burma, and cluster 2 contained few accessions of IS and IK. The Indonesian collection IK76 formed two clusters (fifth and ninth cluster), while all other clusters showed domination of a particular accession from a country. The mean alpha value obtained from the analysis was 0.0299. The Mean Fixation Index (Fst) was 0.79 with a range from 0.88 to 0.70 (Table. 3). Cluster 3 and cluster 9 dominated by Indonesian accessions showed highest Fst of 0.88, and cluster 1 dominated by intermixed population from New Guinea, Pakistan and Burma showed the lowest Fst of 0.70, respectively. High Fst values correlated with low gene flow among the clusters. Low gene flow of 0.03 was observed in cluster 9 and cluster 1 showed a maximum gene flow of

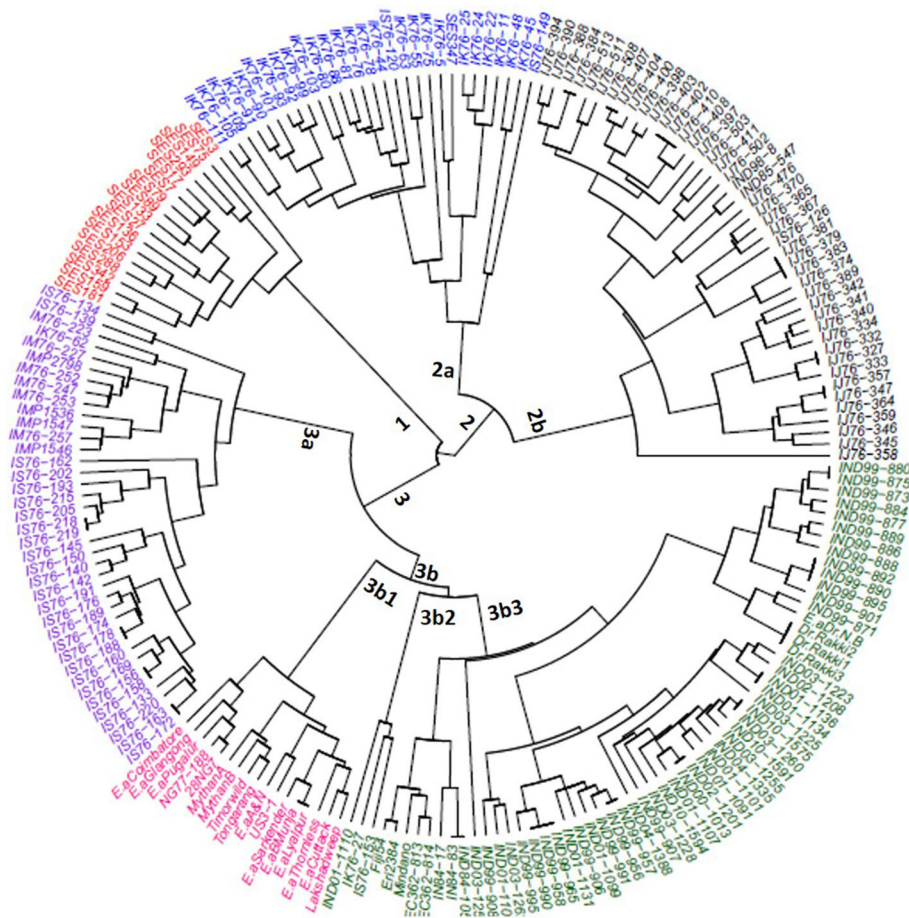


Fig. 1 The phylogenetic tree derived based on the dissimilarity index and calculated based on simple matching using DARwin 6.0. The clones from different geographical origin are given in different colours, and the numbers in the figure indicate the number of major clusters (1, 2 and 3), sub-clusters (2a, 2b, 3a and 3b) and sub-sub-clusters (3b1, 3b2 and 3b3) formed by the population

Table 3 Fixation Index and gene flow observed among the 208 accessions of *E. arundinaceus*

| Sub-population | Fixation index (Fst) | Gene flow (Nm) | Proportion of individuals which are having > 60% membership in the cluster |
|----------------|----------------------|----------------|--|
| 1 | 0.70 | 0.11 | 93.3 |
| 2 | 0.73 | 0.09 | 50.0 |
| 3 | 0.80 | 0.06 | 83.3 |
| 4 | 0.84 | 0.05 | 89.2 |
| 5 | 0.73 | 0.09 | 61.5 |
| 6 | 0.80 | 0.06 | 83.3 |
| 7 | 0.80 | 0.06 | 94.4 |
| 8 | 0.86 | 0.04 | 95.7 |
| 9 | 0.88 | 0.03 | 87.5 |

0.11. Proportion of individuals sharing > 60% membership within cluster ranged from 50 (cluster 2) to 95.7 (cluster 8). In Fig. 2 the bar colours indicate the membership proportion of each individuals.

Sequencing of Amplified Polymorphic Bands

Ten polymorphic bands (≥ 500 bp) obtained from different SRAP marker combinations were selected for further

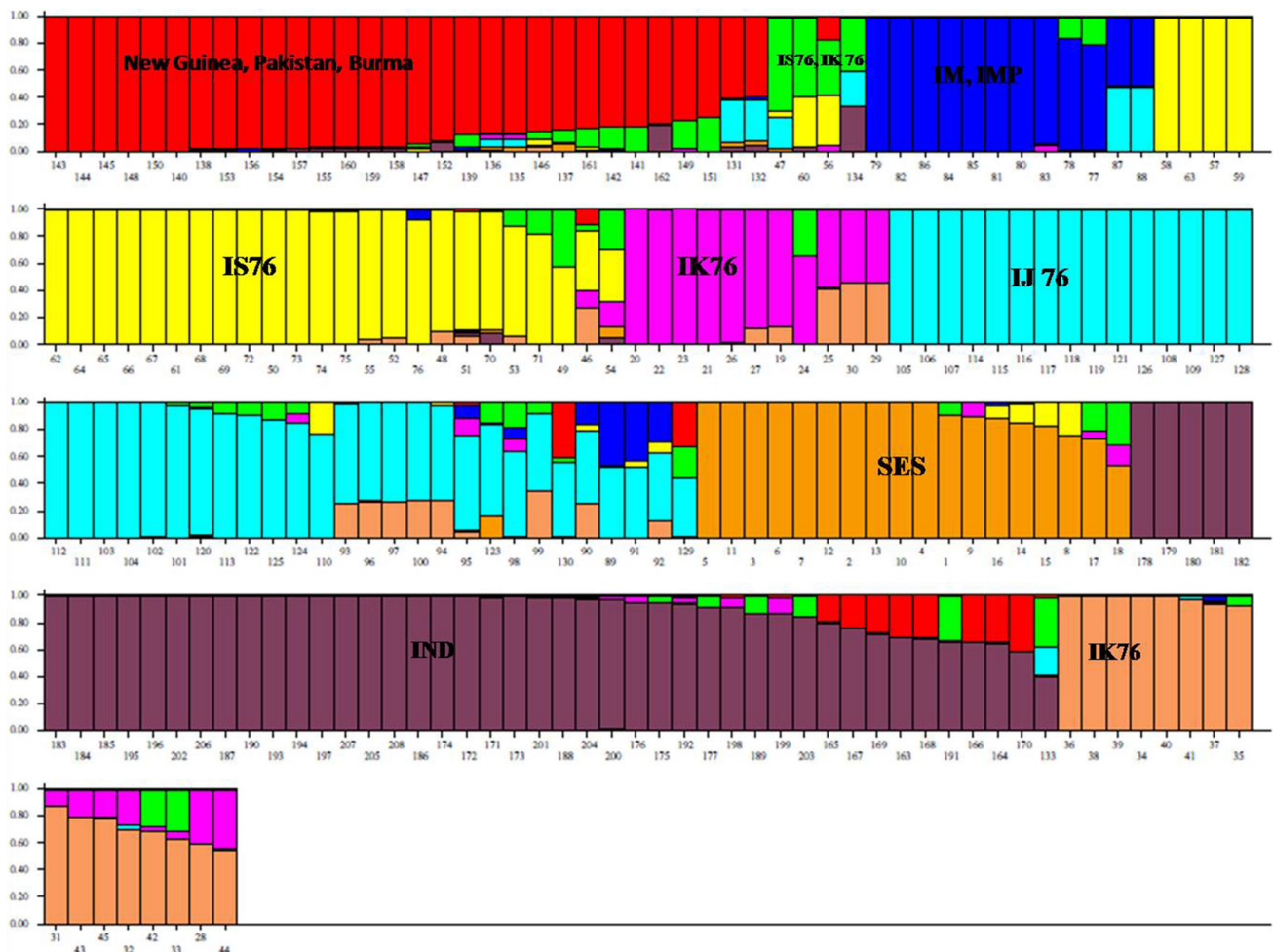


Fig. 2 Population structure analysis of 208 *E. arundinaceus* clones derived using Structure version 2.3.4. Each vertical bar represents one individual genotype. Vertical bars with multiple colours contains

admixed genotypes from multiple clusters. Individual colours with maximum number of dominated clones in one cluster is labelled

Table 4 Sequence homology details of polymorphic fragments amplified using SRAP markers

| Sl. No | Primer pair combination | Amplicon size (bp) | Homology hit |
|--------|-------------------------|--------------------|---|
| 1 | E5M1 | 500 | <i>Sorghum bicolor</i> MADS-box protein SOC1 (LOC8085162) |
| 2 | E2M6 | 1000 | <i>Sorghum bicolor</i> uncharacterized protein (LOC110437092) |
| 3 | E6M2 | 500 | <i>Sorghum bicolor</i> jmjC domain-containing protein 4 (LOC8075464) |
| 4 | E3M13 | 1000 | <i>Sorghum bicolor</i> pseudo histidine-containing phosphotransfer protein 2 (LOC8055697) |
| 5 | E1M4 | 500 | <i>Arachishypogaea</i> probable pectinesterase/pectinesterase inhibitor 51 (LOC112791131) |
| 6 | E1M11 | 700 | <i>Sorghum bicolor</i> uncharacterized protein (LOC8064011) |
| 7 | E3M4 | 600 | <i>Saccharum officinarum</i> Ribosomal protein L22-like mRNA |
| 8 | E1M14 | 800 | <i>Saccharum spontaneum</i> xyloglucan endotransglucosylase/hydrolase protein 6 |
| 9 | E1M1 | 600 | <i>Zea mays</i> uncharacterized protein (LOC103647121) |
| 10 | E2M5 | 500 | <i>Zea mays</i> protein kinase-like gene (LOC100285244) |

sequence characterization in order to verify whether the SRAP marker binding sites correspond to the gene-rich

regions. The amplicons were gel eluted, cloned and subjected to sequencing. After removing the vector sequences

blast analysis was carried out using NCBI BLASTn. Blast analysis of all the ten fragment sequences showed homology to functional genes (Table 4). Among the ten fragments, four showed homology to genes of *Sorghum bicolor* such as MADs-box protein, pseudohistidine-containing phosphotransfer protein 2 and uncharacterized protein. A fragment amplified using combination E3M4 showed homology to *Saccharum officinarum* ribosomal protein L22-like mRNA and one fragment from combination E1M14 matched with *Saccharum spontaneum* xyloglucan endotransglucosylase/hydrolase protein 6. Two fragments amplified using E2M5 and E1M1 showed homology to *Zea mays* uncharacterized protein (LOC103647121) and protein kinase-like gene (LOC100285244).

Discussion

Due to its high adaptability to various biotic and abiotic stress as well as extensive high biomass producing ability, *E. arundinaceus* clones have been utilized in developing intergeneric hybrids. As a prerequisite to utilize these germplasm clones in the crop improvement programs, several studies have been carried out on understanding the genetic diversity among different populations of *E. arundinaceus*. Studies on genetic diversity among Chinese *E. arundinaceus* using DNA markers such as AFLP, RAPD reported a high molecular diversity (Zhang et al. 2004, 2008). Zhang et al. (2013) analyzed genetic variation using SRAP primer pairs and reported a high polymorphism among and within 164 *E. arundinaceus* clones of 18 populations collected from six Chinese provinces. Similarly, Cai et al. (2012) used AFLP markers to study 207 *E. arundinaceus* accessions which consisted majority of population from China and Indonesia, with few representative from India, New Guinea, Philippines and Vietnam. The study indicated a high molecular diversity among Chinese population and low diversity among *E. arundinaceus* from Indonesia and intermediate diversity in the collections of India (Besse et al. 1996, 1997; Cai et al. 2012). However, to conclude intermediate diversity among the Indian population the sample size used in the study was very less.

So far, no large-scale genetic diversity studies on the world collection of *E. arundinaceus* maintained at ICAR-SBI have been carried out. In a study by Nair and Mary (2006) using RAPD markers analyzed the genetic diversity among twelve representative clones from India (SES, IND), Andaman and Nicobar Islands and Indonesia (IJ and IK). The current study is the first report on the genetic diversity of all the world germplasm of *E. arundinaceus* maintained at ICAR-Sugarcane Breeding Institute, Coimbatore.

In this study, the SRAP marker survey of 208 *E. arundinaceus* showed 75.9% PPB which was lesser than the PPB value of 99% obtained using AFLP in Chinese population, but higher than the PPB reported in Chinese population using ISSR (64%) and RAPD (70%) (Zhang et al. 2004, 2008). As reported in earlier studies, the polymorphism among the *E. arundinaceus* clones of Indonesia was lesser (65%) than Indian germplasm accessions (Cai et al. 2012).

The genetic distance calculated among the population (0.01–0.60) indicates the genetic variation. The maximum genetic distance between the IK76 and SES collections (0.4–0.6) indicates high genetic variation among these population. The maximum genetic distance was observed between Indian and Indonesian population which could be attributed to the maximum variation among these population. The phylogenetic tree derived using WARD method clustered all the clones based on geographical origin. However, the Indian clones SES and IND collected from different states within India formed separate cluster. Most of the SES clones are reported to be collected before 1942 through *Spontaneum* Expedition Scheme and could be genetically diverse from the recent IND collections. Clustering of the germplasm accessions from Philippines (MINDANO, EC362-813, 814, Eri 2384) and Fiji (Fiji 54) along with the IND collections and the clustering of accessions from New Guinea, Pakistan (E.a Sarkender and E.a Lyalpur) and Burma (Mythan A, Mythan B) with other Indian accessions indicates very less variation among these population.

The Bayesian approach for population structure analysis is considered a quantitative clustering method as it computes the proportion of the genome of an individual originating from each inferred population. All the 208 *Erianthus* accessions investigated showed 80–90% proportion of individuals with more than 60% memberships sharing among the seven clusters, while only two clusters had 50–60% individuals indicating very few admixtures. The mean alpha value obtained from the analysis was 0.0299 which indicates most individuals in the cluster are essentially from the same population (Evanno et al. 2005).

The clustering in structure was more or less similar to the phylogenetic tree drawn using the dissimilarity indices. The population structure showed clusters based on the origin, and there were sub-groups within a population originated from the same country. Mean F_{st} values confirmed the existence of differences among clusters (Zhao et al. 2010). Population with high Mean Fixation Index (F_{st}) showed low gene flow. High F_{st} and lowest gene flow among the Indonesian accessions (cluster 3 and cluster 9) indicate very low genetic differentiation. Low F_{st} with highest gene flow among intermixed population from New Guinea, Pakistan and Burma (cluster 1) indicates

significant genetic differentiation among these population. The Indian population IND and SES (seventh and eighth cluster) showed intermediate fixation index and gene flow in the analysis. Out-crossing species are reported to have higher levels of gene flow, but the maximum Nm value of *E. arundinaceus* (an outcrossing species) populations in this study was only 0.11. The homology of SRAP polymorphic markers with gene rich regions could be used as markers in the future studies.

Conclusion

This is the first large-scale genetic diversity study on world germplasm collection of *E. arundinaceus* maintained at ICAR-Sugarcane Breeding Institute, Coimbatore, India. The study supports the utility of SRAP markers for assessing the genetic diversity of *Erianthus* germplasm. Results from the sequencing and BLASTn analyses of the SRAP amplicons also support the fact that SRAP markers target gene regions. The phylogenetic clusters of germplasm clones used in the study correlated with geographical origin of the clones, indicating a strong genetic differentiation existing among the world germplasm collection. The population structure obtained in the study also indicated a lower gene flow and significant genetic variation among the population collected from India and from other countries such as Indonesia, Fiji, New Guinea, Pakistan, Burma and Philippines. *Erianthus* clones are high biomass producing clones which grows well under marginal inputs and is utilized in developing high fibre energy canes. They are also utilized in developing biotic and abiotic stress tolerant intergeneric hybrids between *Saccharum* and *Erianthus*. The diversity data provided in this study have potential practical applications to choose the most diverse parents with desirable traits to be utilized in sugarcane improvement as well as energy cane breeding programs. The breeding programs with more divergent parent from unexploited basic germplasm will also help in broadening the sugarcane genetic pool.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12355-021-01037-8>.

Acknowledgements DST-SERB Early Career Research Award for carrying out the study to the first author is kindly acknowledged.

References

- Aitken, K., and M. McNeill. 2010. Diversity analysis. In *Genetics, genomics and breeding of sugarcane*, 1st ed., ed. R. Henry and C. Kole, 19–42. Enfield: Science Publishers.
- Anderson, J.A., G.A. Churchill, J.E. Autrique, S.D. Tanksley, and M.E. Sorrell. 1993. Optimizing parental selection for genetic linkage maps. *Genome* 36: 181–186.
- Berding, N., and H. Koike. 1980. Germplasm conservation of the *Saccharum* complex: A collection from the Indonesian Archipelago. *Hawaii Planters Record* 59: 176–187.
- Besse, P., C.L. McIntyre, and N. Berding. 1996. Ribosomal DNA variations in *Erianthus*, a wild sugarcane relative (Andropogoneae-Saccharinae). *Theoretical and Applied Genetics* 92: 733–743.
- Besse, P., C.L. McIntyre, and N. Berding. 1997. Characterization of *Erianthus* sect. *Ripidium* and *Saccharum* germplasm (Andropogoneae-Saccharum) using RFLP markers. *Euphytica* 93: 283–292.
- Cai, Q., J.C. Wen, Y.H. Fan, L.P. Wang, and L. Ma. 2002. Chromosome analysis of *Saccharum* L. and related plants. *Southwest China Journal of Agricultural Sciences* 15: 16–19.
- Cai, Q., K.S. Aitken, H. Deng, X. Chen, C. Fu, P.A. Jackson, and C. McIntyre. 2008. Verification of the introgression of *Erianthus arundinaceus* germplasm into sugarcane using molecular markers. *Plant Breeding* 124: 322–328.
- Cai, Q., K.S. Aitken, Y.H. Fan, G. Piperidis, X.L. Liu, C.L. McIntyre, X.Q. Huang, and P. Jackson. 2012. Assessment of the genetic diversity in a collection of *Erianthus arundinaceus*. *Genetic Resources and Crop Evolution* 59: 1483–1491.
- Chang, D., F.Y. Yang, J.J. Yan, Y.Q. Wu, S.Q. Bai, X.Z. Liang, Y.W. Zhang, and Y.M. Gan. 2012. SRAP analysis of genetic diversity of nine native populations of wild sugarcane, *Saccharum spontaneum*, from Sichuan, China. *Genetics and Molecular Research* 11: 1245–1253.
- Dao, Z., G. Yan, J. Zhang, D. Chang, S. Bai, Z.H. Chen, D.X. Li, M.H. You, Y. Zhang, C.B. Zhang, J. Zhang, X. Yan, and C. Hu. 2013. Investigation and collection of wild *Erianthus arundinaceus* germplasm resources. *Journal of Plant Genetic Resources* 14: 816–820.
- Deng, H.H., Z.Z. Liao, Q.W. Li, F.Y. Lao, C. Fu, and X. Chen. 2002. Breeding and isozyme marker assisted selection of F2 hybrids from *Saccharum* spp. x *Erianthus arundinaceus*. *Sugarcane and Cane Sugar* 1: 1–5.
- Doyle, J.J. 1991. DNA Protocols for plants-CTAB total DNA isolation. In *Molecular techniques in taxonomy*, ed. G.M. Hewitt and A. Johnston, 283–293. Berlin: Springer.
- Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology* 14: 2611–2620.
- ISMA, 2020. Indian Sugar Mills Association (ISMA), Sugar recovery data as on 31st January, 2020
- Jackson, P., and R.J. Henry. 2011. *Erianthus*. In *Wild crop relatives: Genomic and breeding resources: Industrial crops*, ed. C. Kole, 97–109. Berlin: Springer.
- Li, G., and C.F. Quiros. 2001. Sequence-related amplified polymorphism (SRAP), a new marker system based on a simple PCR reaction: Its application to mapping and gene tagging in *Brassica*. *Theoretical and Applied Genetics* 103: 455–461.
- Manechini, J.R.V., J.B. da Costa, B.T. Pereira, L.A. Carlini-Garcia, M.A. Xavier, M.G.A. de Landell, and L.R. Pinto. 2018. Unraveling the genetic structure of Brazilian commercial sugarcane cultivars through microsatellite markers. *PLoS ONE* 13 (4): e0195623.
- Ming, R., P.H. Moore, K.K. Wu, and A. D’Hont. 2006. Sugarcane improvement through breeding and biotechnology. *Plant Breeding Reviews* 27: 115–118.
- Mugehu, E., L.T. Mpfu, and P. Manjeru. 2020. Genetic diversity evaluation among fifty sugarcane accessions from the Zimbabwe Sugar Industry using DNA-SSR markers and qualitative traits. *BioRxiv preprint*.

- Mukherjee, S.K. 1957. Origin and distribution of *Saccharum*. *Botanical Gazette* 119: 55–61.
- Nair, N.V., and S.J. Mary. 2006. RAPD analysis reveals the presence of mainland Indian and Indonesian forms of *Erianthus arundinaceus* (Retz.) Jeswiet in the Andaman-Nicobar Islands. *Current Science* 90: 1118–1122.
- Nair, N.V., S. Nair, T. Sreenivasan, and M. Mohan. 1999. Analysis of genetic diversity and phylogeny in *Saccharum* and related genera using RAPD markers. *Genetic Resources and Crop Evolution* 46: 73–79.
- Nair, N.V., K. Mohanraj, K. Sunadaravelpandian, A. Suganya, A. Selvi, and C. Appunu. 2017. Characterization of an intergeneric hybrid of *Erianthus procerus* x *Saccharum officinarum* and its backcross progenies. *Euphytica* 213: 267.
- Pritchard, J., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.
- Suman, A., A.K. Collins, J.E. Serge, and V. John. 2008. Sequence-related amplified polymorphism (SRAP) markers for assessing genetic relationships and diversity in sugarcane germplasm collections. *Plant Genetic Resources: Characterization and Utilization* 6 (3): 222–231.
- Tew, T.L., and R.M. Cobill. 2008. Genetic improvement of sugarcane (*Saccharum* spp.) as an energy crop. In *Genetic improvement of bioenergy crops*, ed. W. Vermerris, 273–294. Berlin: Springer.
- Tsuruta, S., M. Ebina, M. Kobayashi, T. Hattori, and T. Terauchi. 2012. Analysis of genetic diversity in the bioenergy plant *Erianthus arundinaceus* (Poaceae: Andropogoneae) using amplified fragment length polymorphism markers. *Grassland Science* 58: 174–177.
- Wang, W., R. Li, H. Wang, B. Qi, X. Jiang, Q. Zhu, D. Cai, X. Tang, and Q. Zhao. 2019. Sweetcane (*Erianthus arundinaceus*) as a native bioenergy crop with environmental remediation potential in southern China: A review. *Bioenergy* 11: 1–14.
- Yang, Q.H., F.S. Li, F.H. Xiao, and S.C. He. 1997. Studies on the chromosomes and botanical characters of *Erianthus Arundinaceus* (Retz.) Jeswiet. *Journal of Yunnan Agricultural University* 12: 253–256.
- Zhang, M.Q., Y.X. Hong, Q.W. Li, S.M. Liu, C.M. Zhang, and R.Z. Yang. 2004. Molecular polymorphic analyses for the germplasm of *Erianthus arundinaceus* collected in China. *Journal of Plant Research and Environment* 13: 1–6.
- Zhang, H.Y., F.S. Li, X.Z. Liu, L.L. He, Q.H. Yang, and S.C. He. 2008. Analysis of genetic variation in *Erianthus arundinaceus* by random amplified polymorphic DNA markers. *African Journal of Biotechnology* 7: 3414–3418.
- Zhang, J., J. Yan, Y. Zhang, X. Ma, S. Bai, Y. Wu, Z. Dhao, D. Li, C. Zhang, Y. Zhang, M. You, F. Yang, and J. Zhang. 2013. Molecular insights of genetic variation in *Erianthus arundinaceus* populations native to china. *PLoS ONE* 8 (11): e80388.
- Zhang, J., J. Yan, X. Shen, S. Bai, D. Li, Y. Zhang, M.H. You, and Y. Zhang. 2017. Development of polymorphic microsatellite markers and their use in collections of *Erianthus arundinaceus* in China. *Grassland Science* 63 (1): 54–60.
- Zhao, Z.H., J.Z. Xi, Q. Jia, S.F. Li, and H.Y. Huang. 2010. Analysis of genetic structure and diversity of Chai chicken breed using microsatellite markers. *Journal of Animal and Veterinary Advances* 9: 1197–1200.
- Zhou, L.X., Y. Xiao, W. Xia, and Y.D. Yang. 2015. Analysis of genetic diversity and population structure of oil palm (*Elaeis guineensis*) from China and Malaysia based on species-specific simple sequence repeat markers. *Genetics and Molecular Research* 14: 16247–16254.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.