GMR
The Original

DIVERSITY PROFILING OF INDIAN-ORIGIN *SACCHARUM OFFICINARUM* CLONES THROUGH MOLECULAR CHARACTERIZATION

S Swetha[1], M Nisha[2*], Safeeda Parakkandi[3], Chandana Sudheer[4], R Sathya[5], Prathima Perumal Thirugnanasambandam[6]

[1]Student, SRM College of Agricultural Sciences, SRMIST, Chengalpattu, India. E. mail: s.swethashs@gmail.com
[2*]Senior Scientist & Officer In-Charge, ICAR-Sugarcane Breeding Institute Research Centre Kannur Civil station post, Kannur. Pin: 670002 E-mail: Nisha.M@icar.org.in
[3]Student, ICAR-Sugarcane Breeding Institute Research Centre Kannur Civil station post, Kannur. E. mail: safeedaparakkandi@gmail.com
[4]Student, ICAR-Sugarcane Breeding Institute Research Centre Kannur Civil station post, Kannur. E. mail: Chandanasudhi2002@gmail.com
[5]Assistant Professor Senior Grade, Department of Genetics and Plant Breeding SRM College of Agricultural Sciences, SRMIST, Chengalpattu, India. E. mail: rsathyapbg@gmail.com
[6]Principal Scientist (Plant Biotechnology), Division of Crop Improvement, ICAR-Sugarcane Breeding Institute, Coimbatore, Tamil Nadu, India. Pin: 641 007. E. mail: prathimasambandam@gmail.com, prathima.PT@icar.gov.in
***Corresponding Author:** M Nisha[2*], E-mail: Nisha.M@icar.org.in

ABSTRACT:
Saccharum officinarum, commonly known as noble cane, contributes approximately 80–90% of the genome of modern nobilized sugarcane cultivars and is an important source of high sucrose content and superior juice quality. However, the genetic diversity present in Indian-origin *S. officinarum* germplasm has received limited molecular characterization. In the present study, 24 elite Indian-origin *S. officinarum* accessions conserved at the ICAR-Sugarcane Breeding Institute Research Centre, Kannur, were genotyped using twenty simple sequence repeat (SSR) markers to assess genetic diversity, clustering patterns, and population structure. A total of 241 alleles were amplified, with an average of 12.05 alleles per marker and an overall polymorphism of 98.76%. The mean polymorphic information content values ranged from 0.209 to 0.467. Analysis of Molecular Variance showed that 83.5% of the total molecular variation was distributed within clusters, while significant differentiation was observed among clusters (PhiPT = 0.165, P = 0.0001). Jaccard's dissimilarity coefficients ranged from 0.3273 to 0.8162, with a mean of 0.6156 ± 0.0866, indicating that all 24 accessions retained a distinct molecular identity. UPGMA clustering grouped the accessions into four clusters, with a single cluster comprising 79.17% of the germplasm. This clustering pattern was broadly supported by Principal Coordinate Analysis. Overall, the findings suggest that although the collection is genetically cohesive, possibly reflecting a historically narrow breeding base, each accession remains molecularly distinguishable. The SSR panel used in this study therefore provides a useful tool for germplasm curation and for identifying genetically distinct clones for future sugarcane breeding programmes.

KEYWORDS: *Saccharum officinarum*; SSR markers; Genetic diversity; Germplasm characterization; UPGMA clustering

INTRODUCTION

Sugarcane (*Saccharum* spp.) is one of the world's most economically important agro-industrial crops, serving as the primary global source of table sugar as well as an important feedstock for bioethanol and biofuel production (Waclawovsky et al. 2010). India is among the world's largest sugarcane producers, and the crop plays a major role in supporting millions of farming households and sustaining the country's rural sugar economy (Mahadevaiah et al. 2021). Modern commercial sugarcane cultivars are highly polyploid and aneuploid interspecific hybrids ("nobilized" canes), derived mainly from crosses between *S. officinarum* L. (2n = 8x = 80, x = 10) and the wild species *S. spontaneum* L. (2n = 40-128, x = 8) (D'Hont et al. 1998). As 80-90% of the genome of hybrid cultivars is contributed by *S. officinarum*, this "noble cane" species continues to be the major source of high sucrose content, thick internodes, and superior juice quality traits, and therefore remains highly valuable in sugarcane breeding programmes (Mahadevaiah et al. 2021).

S. officinarum is believed to have originated in New Guinea and subsequently diversified across the Melanesian and South-East Asian islands before being introduced and naturalized on the Indian subcontinent. In India, it has traditionally been maintained as thick, sweet "chewing cane" types across diverse agro-climatic zones. Successive germplasm exploration and collection missions conducted across India, including the north-eastern hill states, the Cauvery River basin, and coastal Tamil Nadu, have resulted in the assembly of a considerable pool of indigenous officinarum accessions. These accessions are now conserved ex situ in national repositories such as the ICAR-Sugarcane Breeding Institute, Coimbatore (Nair et al. 2006; Nair and Sekharan 2009). However, more than a century of intensive breeding based on a limited number of founder parental clones has progressively narrowed the genetic base of cultivated sugarcane (Mahadevaiah et al. 2021). This has increased the need for systematic characterization and conservation of the remaining diversity within indigenous officinarum germplasm before it is lost through genetic erosion or clonal duplication in field gene banks.

As sugarcane is vegetatively propagated, highly heterozygous, and strongly influenced by environmental conditions and crop age, morphological descriptors alone are often insufficient to clearly distinguish genetic identity and relatedness among closely related clones. Morphological characterization can also be complicated by synonymous accessions carrying different local names or by true duplicates being maintained as separate types (Siraree et al. 2018). DNA-based molecular markers help overcome these limitations by directly detecting variation at the genome level, independent of environmental conditions, developmental stage, and scoring subjectivity. Consequently, they have become important tools for accurate germplasm characterization, fingerprinting, and diversity assessment in *Saccharum* (Singh et al. 2020).

A wide range of molecular marker systems has been used to investigate genetic diversity in sugarcane over the past two decades. Maize-derived microsatellite (SSR) markers were among the earliest cross-transferable marker systems evaluated for genetic fingerprinting in sugarcane. These markers showed high transferability and proved useful for diversity analysis despite the complex polyploid genome of the crop (Selvi et al. 2003). Subsequently, AFLP markers were applied to 30 clones representing *S. officinarum*, *S. robustum*, *S. spontaneum*, *S. barberi*, *S. sinense*, and the related genus *Erianthus*. The resulting phenetic tree was consistent with known taxonomic relationships and provided substantially higher resolution than earlier RAPD- and RFLP-based approaches (Selvi et al. 2006). A related AFLP-based study of tropical and subtropical Indian sugarcane cultivars further contributed to understanding the genomic constitution and relationships within cultivated Indian germplasm (Selvi et al. 2005). Khan et al. (2009) used RAPD markers to assess diversity among 20 accessions of *S. officinarum* in Pakistan and reported 86.8% polymorphism across 21 primers. Similarly, a RAPD-based investigation of chewing sugarcane varieties demonstrated considerable intraspecific variability within *S. officinarum* germplasm (Ullah et al. 2013). Combined RAPD and sequence-tagged microsatellite site (STMS) markers have also been employed to assess diversity among high-biomass-producing sugarcane hybrids in India (Saravanakumar et al. 2014). Furthermore, AFLP profiling of *S. officinarum* accessions in comparison with commercial hybrid cultivars showed that this species retains markedly greater allelic richness than modern nobilized hybrids, further emphasizing its importance as a source of novel breeding variability (Aitken et al. 2006).

Microsatellite-based studies of clonal germplasm representing five *Saccharum* species *S. barberi*, *S. robustum*, *S. officinarum*, *S. sinense*, and *S. spontaneum* have provided further insights into inter- and intra-specific relationships within the genus and confirmed the distinctiveness of *S. officinarum* as a genetically coherent group (Brown et al. 2007). Chloroplast microsatellite (coSSR) markers have also been used to investigate maternal lineage diversity across the wider *Saccharum* complex (Raj et al. 2016). In India, SSR marker-based studies have characterized the genetic structure of subtropical sugarcane varieties (Siraree et al. 2018), as well as cultivars, wild progenitor species, and related genera collectively (Singh et al. 2020). In addition, a molecular investigation of the wild relative *S. spontaneum* across four geographical regions of India revealed pronounced region-wise genetic differentiation, with accessions from Arunachal Pradesh showing the greatest diversity (Mary et al. 2006). A subsequent worldwide assessment of *S. spontaneum* diversity showed that only a fraction of the wild diversity available globally has been incorporated into active sugarcane breeding programmes, highlighting the broader issue of underutilized genetic variability within the *Saccharum* complex (Aitken et al. 2018).

At the global level, large-panel SSR genotyping of the world collection of *Saccharum* germplasm has facilitated the development of representative core collections that capture maximum genetic diversity using a relatively small number of accessions. Such collections improve the efficiency with which breeders can access and utilize conserved genetic variability (Nayak et al. 2014). More recently, SSR-based population structure analysis involving sugarcane hybrids and six related *Saccharum* species demonstrated that *S. officinarum*, together with *S. robustum* and *S. barberi*, possesses among the highest numbers of private alleles and the highest Shannon diversity index values within the genus. These findings further support the exceptional value of *S. officinarum* as a reservoir of untapped genetic variation (Xiong et al. 2022).

Collectively, these studies demonstrate that molecular markers can not only quantify the extent of genetic diversity present within *Saccharum* germplasm but also provide valuable information on population structure, clonal redundancy, and species-specific allelic patterns that cannot be reliably identified through field observation alone. Despite the substantial and growing body of molecular research on sugarcane hybrids, commercial cultivars, and the wild species *S. spontaneum*, dedicated molecular diversity profiling of Indian-origin *S. officinarum* clones remains comparatively limited. Most existing marker-based studies have focused predominantly on interspecific hybrid panels, elite parental lines, or globally pooled germplasm collections rather than on indigenous noble-cane accessions considered separately (Xiong et al. 2022). Since *S. officinarum* continues to serve as the principal source of sucrose-related genes in nobilization breeding, focused molecular characterization of the Indian *officinarum* gene pool is therefore important to (i) quantify the extent and pattern of genetic variability retained among indigenous clones, (ii) resolve fine-scale genetic relationships, redundancy, and population structure within the collection, and (iii) identify genetically distinct and previously unexploited clones that could be utilized as parents in genetic base-broadening and pre-breeding programmes. Developing such a molecular diversity profile would provide breeders with an objective and reproducible basis for prioritizing germplasm conservation and selecting parents, thereby supporting long-term genetic gain and the sustainability of sugarcane improvement in India (Mahadevaiah et al. 2021). In this context, the present study was undertaken to characterize a set of Indian-origin *S. officinarum* clones using molecular marker-based analysis. The specific objectives were to assess the extent of DNA-level genetic polymorphism, clustering patterns and population structure among the accessions, and identify genetically divergent, breeding-worthy clones within this valuable indigenous germplasm resource.

MATERIALS AND METHODS

Plant materials

A total of 24 elite Indian-origin *S. officinarum* accessions were selected from the field-maintained germplasm collection at the ICAR-Sugarcane Breeding Institute Research Centre (ICAR-SBIRC), Kannur, Kerala, India. (Table S1).

Genomic DNA isolation and quantification

Genomic DNA was extracted from young leaf tissue using a rapid sodium dodecyl sulfate (SDS) -based DNA isolation method. About 100 mg of finely cut fresh inner leaf tissue was transferred into a sterile 2 mL microcentrifuge tube containing 400 µL of SDS extraction buffer along with two medium-sized Hi Media ceramic beads. The tissue was homogenized using a TissueLyser III Bead Mill Homogenizer at 30 Hz for 30 s. The homogenized samples were centrifuged at 12,000 rpm for 10 min at 4°C in a pre-cooled refrigerated centrifuge (Sorvall ST 8R, Thermo Fisher Scientific, USA). The clear supernatant was transferred to a fresh sterile 1.5 mL microcentrifuge tube, and an equal volume of chilled isopropanol was added to precipitate the genomic DNA. The samples were incubated at -20°C for 10 min and then centrifuged at 12,000 rpm for 10 min to collect the DNA pellet. The pellet was air-dried at room temperature and dissolved in 50 µL of TE buffer. The extracted DNA was stored at -20°C until further use. DNA concentration and purity were checked using a NanoDrop Lite Plus Spectrophotometer (Thermo Fisher Scientific, USA). Samples showing an A260/A280 ratio between 1.8 and 2.0 were considered suitable for further molecular analysis. The DNA samples were then diluted to a working concentration of 50 ng/µL for subsequent downstream analysis.

SSR-PCR amplification

Twenty SSR primers were used for PCR amplification of the genomic sugarcane DNA (Table 1). PCR amplification of the SSR markers was carried out in a total reaction volume of 10 µL. Each reaction consisted of 2 µL of template DNA, 1 µL of Taq buffer A (10X buffer containing 15 mM $MgCl_2$; GeNei™), 0.3 µL of dNTP mix (10 mM; 2.5 mM each; GeNei™), 0.05 µL of Taq DNA polymerase (3 U/µL; GeNei™), 0.5 µL each of forward and reverse primers, and 5.6 µL of molecular-grade water. Amplification was performed using an Eppendorf Mastercycler nexus GX2. The PCR programme consisted of an initial denaturation at 95°C for 2 min, followed by 35 cycles of denaturation at 94°C, primer annealing for 40 s at 50-65°C depending on the primer, and extension at 72°C for 40 s. A final extension was performed at 72°C for 7 min. The amplified SSR products were separated by electrophoresis on an 8% non-denaturing polyacrylamide gel prepared using a 29:1 ratio of acrylamide to bis-acrylamide in 1× TBE buffer. Electrophoresis was carried out at 160 V for 2-3 h, with a 100 bp DNA ladder included as a molecular size reference. After electrophoresis, the gels were silver-stained and the amplified banding profiles were documented using a ProteinSimple AlphaImager HP. Among the 20 SSR markers used in this study, twelve markers, namely mSSCIR9, mSSCIR24, mSSCIR68, SMC278CS, SMC361BS, SMC334BS, mSSCIR1, SMC336BS, mSSCIR53, SMC597CS, SMC545MS and SMC851MS were obtained based on a sugarcane germplasm evaluation study (Pan, 2006), one primer, UGSM417, was obtained from the microsatellite marker study of Singh et al. (2008), and seven primers, namely NKS25, NKS32, NKS23, NKS28, NKS38, NKS24 and NKS49 were obtained from the sugarcane-specific STMS markers developed by Govindaraj et al. (2005).

Table. 1 Forward and reverse sequences, base pair length and annealing temperatures (Ta) of the SSR primers

S. No	Primer	Forward (5′-3′) / Reverse (3′-5′)	Ta
1	mSSCIR9	TCTCTATGCACCCTATCGT	54
		TAACTTGACCCCCTCTTGA	
2	mSSCIR24	AGATGAACCCAAAAACTTA	50
		TTACTCCGCCTCTTTACT	
3	mSSCIR68	CGTCTCTATGCACCCTATC	52
		GCCTTCTTTTGTTTTCCTC	
4	SMC278CS	TTCTAGTGCCAATCCATCTCAGA	53
		CATGCCAACTTCCAAACAGACT	
5	SMC361BS	GCAGCCAGCCACAGTGTTTA	54
		CACCGCTAAGGACATCAACTAGG	
6	SMC334BS	CAATTCTGACCGTGCAAAGAT	50
		CGATGAGCTTGATTGCGAATG	
7	UGSM417	GCTAGCAACAGATCGGAGTGTC	55
		GTGTACCGTGTGTATGTCTGTC	
8	NKS25	TCCATGCATGCGTGTAGTTT	50
		AGT GCA CAA CGT TCT TGC TG	
9	NKS32	CCAACTCACTCACCCCAGTT	52

		ATGAGAGTGCAGATGCATGG	
10	mSSCIR1	CTTGTGGATTGGATTGGAT	54
		AGGAAATGGATTGCTCAGG	
11	SMC336BS	ATTCTAGTGCCAATCCATCTCA	52
		CATGCCAACTTCCAAACAGAC	
12	mSSCIR53	TGGTCTACTGAAGTTCGTG	53
		TGCTTCTAAGTCAACCAAA	
13	NKS23	TAAACCCCCGAAAAAGAACC	52
		TCCGGAGGTAGATCCATTTG	
14	NKS28	GTGCTGGGATTCTGAGCTTC	54
		GCAAGTTCTTGGCCTTTGTT	
15	SMC597CS	GCACACCACTCGAATAACGGAT	55
		AGTATATCGTCCCTGGCATTCA	
16	NKS38	TGAACTCGGCAACAGTTTTT	52
		CCCACCAAGTCGTTCTGAAT	
17	NKS24	TATATGGCGAGGACAGATGC	53
		GGGTTCAGAATTAGAGCAATCG	
18	NKS49	CTCACGTCCTGTTGGTGCTA	54
		TACATGGGACACATGCTTGC	
19	SMC545MS	AGGCTACATGCTTACAGCCAT	53
		TGGTCTATCACTTAATCAGCCAC	
20	SMC851MS	ACTAAAATGGCAAGGGTGGT	54
		CGTGAGCCCACATATCATGC	

SSR data scoring and statistical analysis

The SSR alleles were scored manually, with the presence and absence of individual alleles recorded as 1 and 0, respectively, to generate a binary data matrix. The polymorphic information content (PIC) of each scored allele was calculated using the formula PIC = $2f$ (1 - f), where f represents the frequency of the allele across the evaluated accessions. For each SSR marker, the mean PIC was obtained by calculating the arithmetic mean of the PIC values of all alleles amplified by that marker. The binary SSR data were further analysed to assess the genetic diversity among the accessions. Analysis of Molecular Variance (AMOVA) was calculated using the GenAlEx 6.5 software. Shannon's diversity index (H′), Simpson's diversity index (1-D), and Simpson's dominance index (D) were calculated using PAST version 4.03. Pairwise Jaccard's dissimilarity coefficients were calculated from the binary matrix using DARwin version 6.0.15. The resulting dissimilarity matrix was used to generate a UPGMA dendrogram for assessing the genetic relationships among the accessions. Principal Coordinate Analysis (PCoA) was performed using the Jaccard's dissimilarity matrix to visualize the genetic relationships among the accessions and to support the clustering pattern obtained from the UPGMA analysis.

RESULTS

SSR amplification and marker characterization

The amplified fragments ranged from 100-500 bp across the markers (Table 2). Among the twenty SSR markers, six markers (mSSCIR53, NKS28, NKS38, NKS24, NKS49, and SMC851MS) amplified the widest fragment size range (100-500 bp), whereas mSSCIR68 produced the narrowest fragment size range (160-453 bp). The largest amplified fragment (500 bp) was detected with nine markers (mSSCIR53, NKS23, NKS28, SMC597CS, NKS38, NKS24, NKS49, SMC545MS, and SMC851MS), while the smallest fragment (100 bp) was generated by seven markers (mSSCIR9, mSSCIR53, NKS28, NKS38, NKS24, NKS49, and SMC851MS).

Table. 2 The molecular weight range, number of alleles, number of polymorphic and monomorphic alleles, and percent polymorphism of the 20 SSR markers

S. No	Markers	Mol. Wt range	No. of alleles	No. of Polymorphic alleles	No. of Monomorphic alleles	% polymorphism
1	mSSCIR9	100 - 460	6	6	0	100
2	mSSCIR24	125 - 470	8	8	0	100
3	mSSCIR68	160 - 453	7	7	0	100

4	SMC278CS	115 - 467	8	8	0	100
5	SMC361BS	102 - 488	7	7	0	100
6	SMC334BS	113 - 465	8	8	0	100
7	UGSM417	106 - 488	8	8	0	100
8	NKS25	113 - 464	8	8	0	100
9	NKS32	105 - 424	6	6	0	100
10	mSSCIR1	103 - 490	6	6	0	100
11	SMC336BS	120 - 499	12	11	1	91.67
12	mSSCIR53	100 - 500	15	15	0	100
13	NKS23	105 - 500	12	12	0	100
14	NKS28	100 - 500	17	17	0	100
15	SMC597CS	107 - 500	19	18	1	94.74
16	NKS38	100 - 500	17	17	0	100
17	NKS24	100 - 500	22	22	0	100
18	NKS49	100 - 500	19	19	0	100
19	SMC545MS	132 - 500	19	19	0	100
20	SMC851MS	100 - 500	17	16	1	94.12
Total			241	238	3	98.76

SSR marker allelic diversity

A total of 241 alleles were generated. The number of alleles amplified per marker ranged from a minimum of 6 (mSSCIR9, NKS32, and mSSCIR1) to a maximum of 22 (NKS24), with an average of 12.05 alleles per marker. Of the 241 alleles, 238 alleles were polymorphic and 3 alleles were monomorphic, giving an overall percent polymorphism of 98.76% across the marker panel. Seventeen of the 20 markers (mSSCIR9, mSSCIR24, mSSCIR68, SMC278CS, SMC361BS, SMC334BS, UGSM417, NKS25, NKS32, mSSCIR1, mSSCIR53, NKS23, NKS28, NKS38, NKS24, NKS49, and SMC545MS) were completely polymorphic, exhibiting 100% polymorphism with no monomorphic alleles detected. Monomorphic alleles were detected at three of the 20 markers, namely SMC336BS, SMC597CS and SMC851MS. These three markers represented the lowest percent polymorphism values recorded among the 20 SSR markers used in the study (Table 2).

SSR Fingerprint Profiles of Indian origin *S. officinarum* Accessions

SSR fingerprint profiles were generated for all 24 accessions using the binary matrix derived from the presence (1) and absence (0) of amplified SSR alleles. The columns represent individual accessions and rows represent distinct alleles generated by the corresponding SSR primer, where black cells indicate allele presence and white cells indicate allele absence. The fingerprint profile generated using the 20 SSR markers is presented in Fig. S1.

Allele frequency distribution and Polymorphic information content (PIC)

Individual allele frequencies across the 20 SSR markers ranged from 0.0417 to 1.0000. The widest allele frequency range was observed for SMC597CS and SMC851MS (0.0417-1.0000), whereas mSSCIR1 exhibited the narrowest range (0.0417-0.2500). The highest mean allele frequency was recorded for mSSCIR24 (0.677), while mSSCIR1 exhibited the lowest mean allele frequency (0.125). The mean PIC values ranged from 0.209 to 0.467, indicating differences in the discriminatory power of the markers. Among the 20 SSR markers, SMC361BS exhibited the highest mean PIC value (0.467), and mSSCIR1 recorded the lowest mean PIC value (0.209). Individual PIC values across all markers ranged from 0.000 to 0.500, a PIC value of 0.000 was recorded at SMC336BS, SMC597CS, and SMC851MS, while a mean PIC 0.500 was recorded at SMC361BS, SMC334BS, UGSM417, SMC597CS, NKS38, and NKS49 (Table 3). The allele frequency and PIC values for each allele across all SSR markers are provided in Table S2.

Table. 3 Allele frequency distribution and polymorphic information content of the SSR markers

Markers	Allele frequency range	Mean allele frequency	PIC range	Mean PIC
mSSCIR9	0.2917 - 0.8333	0.542	0.278 - 0.497	0.431
mSSCIR24	0.2083 - 0.9583	0.677	0.080 - 0.469	0.286
mSSCIR68	0.2083 - 0.7917	0.381	0.330 - 0.486	0.389
SMC278CS	0.2083 - 0.6667	0.464	0.330 - 0.497	0.447
SMC361BS	0.3333 - 0.7083	0.452	0.413 - 0.500	0.467
SMC334BS	0.2500 - 0.8750	0.589	0.219 - 0.500	0.411
UGSM417	0.3333 - 0.8333	0.563	0.278 - 0.500	0.445
NKS25	0.0833 - 0.8333	0.464	0.153 - 0.497	0.387

NKS32	0.1250 - 0.4167	0.236	0.219 - 0.486	0.337
mSSCIR1	0.0417 - 0.2500	0.125	0.080 - 0.375	0.209
SMC336BS	0.1250 - 1.0000	0.545	0.00 - 0.497	0.397
mSSCIR53	0.0417 - 0.7500	0.389	0.080 - 0.497	0.397
NKS23	0.0417 - 0.8333	0.517	0.080 - 0.497	0.411
NKS28	0.0833 - 0.8750	0.409	0.153 - 0.497	0.399
SMC597CS	0.0417 - 1.0000	0.456	0.000 - 0.500	0.365
NKS38	0.1667 - 0.7500	0.449	0.278 - 0.500	0.442
NKS24	0.1250 - 0.7500	0.426	0.219 - 0.497	0.426
NKS49	0.1250 - 0.9583	0.542	0.080 - 0.500	0.414
SMC545MS	0.0417 - 0.7917	0.351	0.080 - 0.497	0.333
SMC851MS	0.0417 - 1.0000	0.507	0.000 - 0.486	0.301

Analysis of molecular variance (AMOVA)

AMOVA revealed that the majority of the molecular variation among the 24 *S. officinarum* accessions was distributed within clusters (Table 4). Of the total molecular variance (σ^2 = 54.318), 16.50% was attributed to variation among clusters (σ^2 = 8.965), whereas 83.50% was attributed to variation within clusters (σ^2 = 45.353). The among-cluster source of variation had a sum of squares of 166.162 (df = 2; MS = 83.081), while the within-clusters exhibited a sum of squares of 952.421 (df = 21; MS = 45.353). The estimated PhiPT value of 0.165 was statistically significant (P = 0.0001) based on 9999 permutations, indicating significant genetic differentiation among the three clusters.

Table 4. Analysis of Molecular Variance (AMOVA) showing the distribution of molecular variation among and within clusters

Source of Variation	df	Sum of Squares (SS)	Mean Square (MS)	Variance Component (σ^2)	Percentage of Variation (%)	PhiPT
Among clusters	2	166.162	83.081	8.965	16.5	0.165
Within clusters	21	952.421	45.353	45.353	83.5	-
Total	23	1118.583	-	54.318	100	-

Genetic diversity indices

Across the 241 SSR alleles amplified by the twenty SSR markers, the Shannon diversity index (H′) ranged from 0.000 to 3.178, with an overall mean of 2.206 ± 0.704. Simpson's Diversity index (1-D) ranged from 0.000 to 0.958, with an overall mean of 0.847 ± 0.178, whereas Simpson's Dominance index (D) ranged from 0.042 to 1.000, with an overall mean of 0.153 ± 0.178 (Table 5). The complete allele-wise diversity indices for all 241 SSR alleles are presented in Fig.1 and Table S3.

Among the SSR markers, mSSCIR24 exhibited the highest mean Shannon diversity index (H′ = 2.659 ± 0.607), followed by SMC334BS (2.584 ± 0.403) and UGSM417 (2.563 ± 0.306). In contrast, mSSCIR1 recorded the lowest mean Shannon diversity (H′ = 0.944 ± 0.625), followed by NKS32 (1.626 ± 0.511) and SMC545MS (1.829 ± 0.857).

A similar pattern was observed for Simpson's Diversity index (1−D), with the highest mean values recorded for UGSM417 (0.920 ± 0.026), SMC334BS (0.918 ± 0.038), and mSSCIR24 (0.916 ± 0.062). The marker mSSCIR9 also exhibited a high mean diversity value of 0.913 ± 0.033. Conversely, mSSCIR1 showed the lowest mean Simpson's Diversity index (0.542 ± 0.297), followed by SMC545MS (0.768 ± 0.233) and SMC597CS (0.777 ± 0.305).

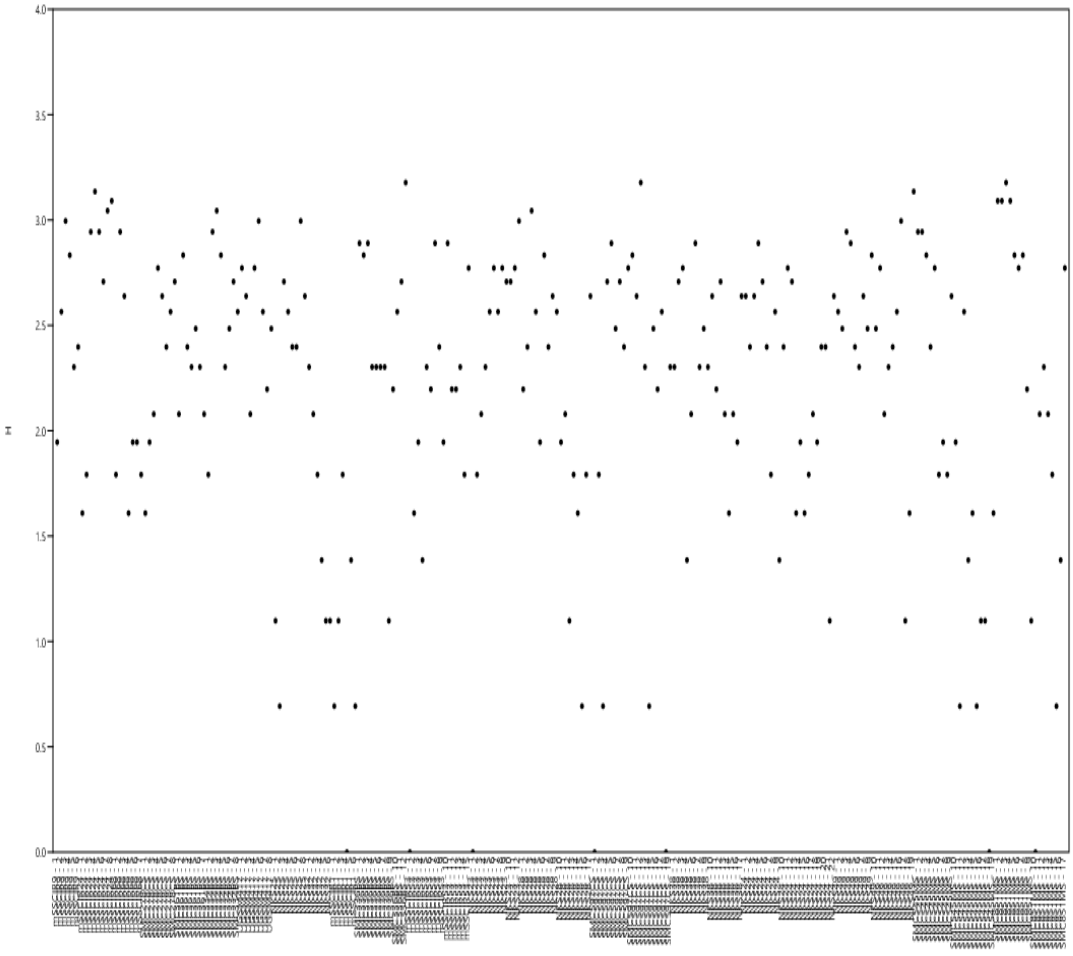
The Dominance index showed the reverse pattern, with mSSCIR1 recording the highest mean value (D = 0.458 ± 0.297), followed by SMC545MS (0.232 ± 0.233), SMC597CS (0.223 ± 0.305), and NKS32 (0.218 ± 0.103). In contrast, UGSM417 exhibited the lowest mean Dominance index (D = 0.080 ± 0.026), followed by SMC334BS (0.082 ± 0.038) and mSSCIR24 (0.084 ± 0.062). The minimum values of H′ and 1-D (0.000) and the maximum Dominance value (D = 1.000) occurred among alleles of mSSCIR1, mSSCIR53, NKS23, SMC597CS, SMC545MS, and SMC851MS, indicating variation in allele-wise diversity across these marker loci (Table S3).

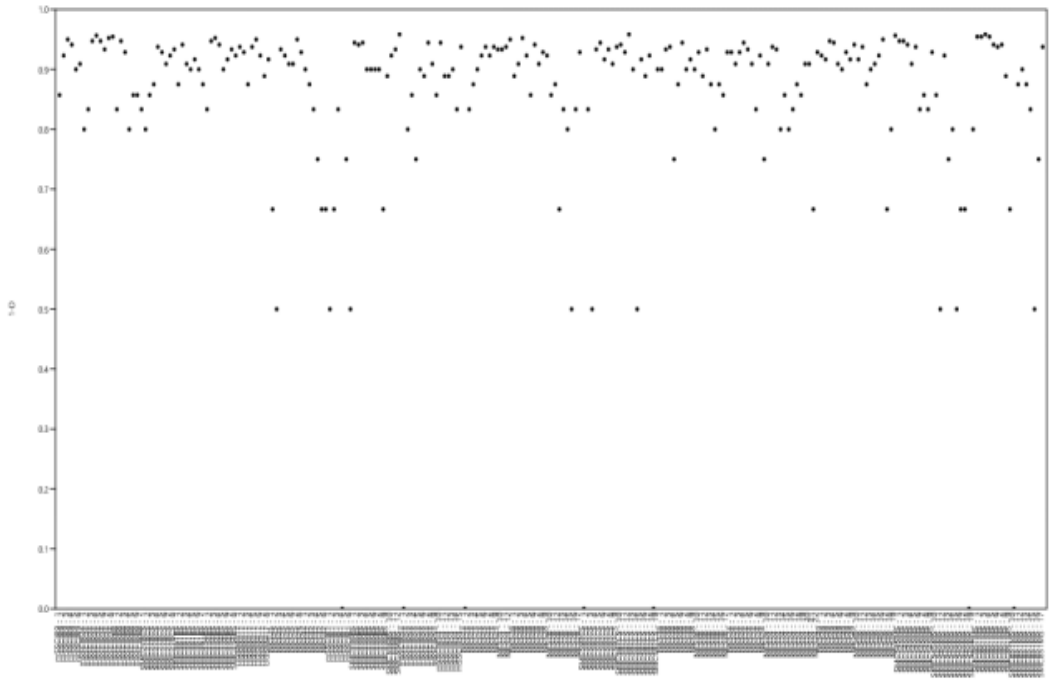
Table. 5 Summary of Shannon's (H'), Simpson's (1-D) and Dominance (D) diversity indices across groups M1 - M20

Marker No.	Marker	No. of alleles	Mean H′ ± SD	Range (Min–Max)	Mean 1−D ± SD	Range (Min–Max)	Mean D ± SD	Range (Min–Max)
M1	mSSCIR9	6	2.507 ± 0.379	1.946–2.996	0.913 ± 0.033	0.857–0.950	0.087 ± 0.033	0.050–0.143
M2	mSSCIR24	8	2.659 ± 0.607	1.609–3.135	0.916 ± 0.062	0.800–0.957	0.084 ± 0.062	0.043–0.200

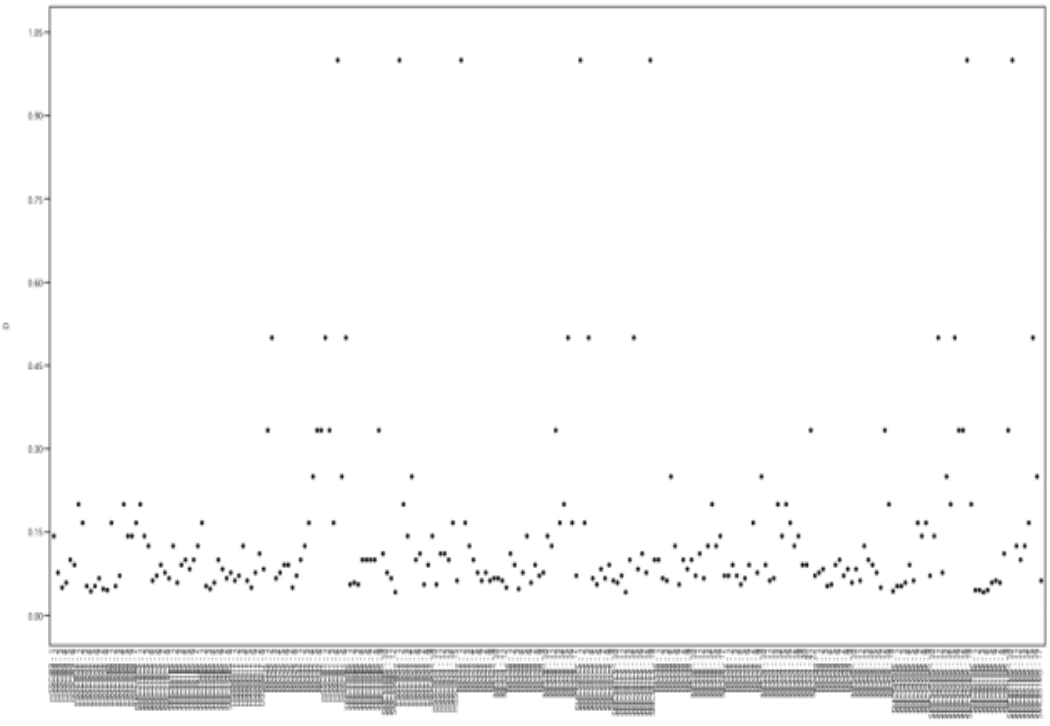
M3	mSSCIR68	7	2.095 ± 0.497	1.609–2.944	0.865 ± 0.054	0.800–0.947	0.135 ± 0.054	0.053–0.200
M4	SMC278CS	8	2.340 ± 0.418	1.609–2.773	0.895 ± 0.048	0.800–0.938	0.105 ± 0.048	0.062–0.200
M5	SMC361BS	7	2.354 ± 0.260	2.079–2.833	0.902 ± 0.023	0.875–0.941	0.098 ± 0.023	0.059–0.125
M6	SMC334BS	8	2.584 ± 0.403	1.792–3.045	0.918 ± 0.038	0.833–0.952	0.082 ± 0.038	0.048–0.167
M7	UGSM417	8	2.563 ± 0.306	2.079–2.996	0.920 ± 0.026	0.875–0.950	0.080 ± 0.026	0.050–0.125
M8	NKS25	8	2.187 ± 0.826	0.693–2.996	0.840 ± 0.165	0.500–0.950	0.160 ± 0.165	0.050–0.500
M9	NKS32	6	1.626 ± 0.511	1.099–2.303	0.782 ± 0.103	0.667–0.900	0.218 ± 0.103	0.100–0.333
M10	mSSCIR1	6	0.944 ± 0.625	0.000–1.792	0.542 ± 0.297	0.000–0.833	0.458 ± 0.297	0.167–1.000
M11	SMC336BS	12	2.464 ± 0.532	1.099–3.178	0.900 ± 0.077	0.667–0.958	0.100 ± 0.077	0.042–0.333
M12	mSSCIR53	15	2.055 ± 0.716	0.000–2.890	0.820 ± 0.233	0.000–0.944	0.180 ± 0.233	0.056–1.000
M13	NKS23	12	2.336 ± 0.810	0.000–2.996	0.840 ± 0.267	0.000–0.950	0.160 ± 0.267	0.050–1.000
M14	NKS28	17	2.131 ± 0.616	0.693–3.045	0.855 ± 0.115	0.500–0.952	0.146 ± 0.115	0.048–0.500
M15	SMC597CS	19	2.086 ± 0.982	0.000–3.178	0.777 ± 0.305	0.000–0.958	0.223 ± 0.305	0.042–1.000
M16	NKS38	17	2.302 ± 0.416	1.386–2.890	0.891 ± 0.051	0.750–0.944	0.109 ± 0.051	0.056–0.250
M17	NKS24	22	2.219 ± 0.501	1.099–2.890	0.876 ± 0.071	0.667–0.944	0.124 ± 0.071	0.056–0.333
M18	NKS49	19	2.473 ± 0.485	1.099–3.135	0.903 ± 0.067	0.667–0.957	0.097 ± 0.067	0.043–0.333
M19	SMC545MS	19	1.829 ± 0.857	0.000–2.944	0.768 ± 0.233	0.000–0.947	0.232 ± 0.233	0.053–1.000
M20	SMC851MS	17	2.194 ± 0.935	0.000–3.178	0.816 ± 0.244	0.000–0.958	0.184 ± 0.244	0.042–1.000
Total	-	241	2.206 ± 0.704	0.000–3.178	0.847 ± 0.178	0.000–0.958	0.153 ± 0.178	0.042–1.000

(a)





(b)



(c)

Fig .1 Marker-wise distribution of (a) Shannon's (H'), (b) Simpson's (1-D) and (c) Dominance (D) diversity indices

Jaccard's dissimilarity coefficient

Genetic dissimilarity among the 24 accessions, estimated using Jaccard's coefficient, ranged from 0.3273 to 0.8162 across all 276 pairwise comparisons, with a mean pairwise dissimilarity of 0.6156 ± 0.0866 (Table 6). The majority of comparisons (40.22%) fell within the 0.60-0.70 dissimilarity class, followed by 33.70% within the 0.50-0.60 class. Overall, 16.30% of the pairwise comparisons exhibited dissimilarity values greater than 0.70. No accession pair exhibited a dissimilarity coefficient below 0.30, and no pair showed identical SSR profiles (dissimilarity coefficient = 0.000); the closest pair, Kea 21 and Khajuria, recorded the lowest dissimilarity coefficient of 0.3273, while the most divergent pair, Poona and Vellai, recorded the highest dissimilarity coefficient of 0.8162. The complete pairwise Jaccard's dissimilarity matrix for the 24 germplasm accessions is provided in Fig. S2.

Table 6. Frequency distribution of Jaccard's pairwise dissimilarity coefficients

Dissimilarity class	Number of pairs	Percentage (%)
0.00 - 0.10	0	0
0.10 - 0.20	0	0
0.20 - 0.30	0	0
0.30 - 0.40	1	0.36
0.40 - 0.50	26	9.42
0.50 - 0.60	93	33.7
0.60 - 0.70	111	40.22
0.70 - 0.80	42	15.22
0.80 - 0.90	3	1.09
0.90 - 1.00	0	0
Total	276	100

UPGMA dendrogram

UPGMA clustering based on the Jaccard's dissimilarity matrix resolved all 24 germplasm accessions into a single tree, which, when cut at the mean dissimilarity coefficient (0.6156), yielded four clusters (I-IV) of variable size, namely Cluster I (19 accessions), Cluster II (3 accessions), and Clusters III and IV comprising single accessions each (Thellacherukku and Vellai, respectively) (Table 7, Fig. 2). Cluster I accounted for 79.17% of the germplasm evaluated, Cluster II accounted for 12.50%, and Clusters III and IV each accounted for 4.17% of the total accessions.

Table 7. Distribution of 24 germplasm accessions into clusters based on the UPGMA dendrogram, cut at the mean dissimilarity coefficient (0.6156)

Cluster	No. of Accessions	Coalescence Height	Accessions
I	19	0.6145	Khajuria, Kea 21, Kaludai Boothan, Manjri Red, Lakhapur, Kajla, Javari Kabbu, Pakaweli-2, Namam, Desi Pounda, Chittan, Assam Red, Malabar, Pattapatti, Pattacherukku 86, Pakaweli-2 Striped, Poovan, Poona, Penang
II	3	0.5779	Saharanpur Black, Rasdali, Ramgarh
III	1	-	Thellacherukku
IV	1	-	Vellai

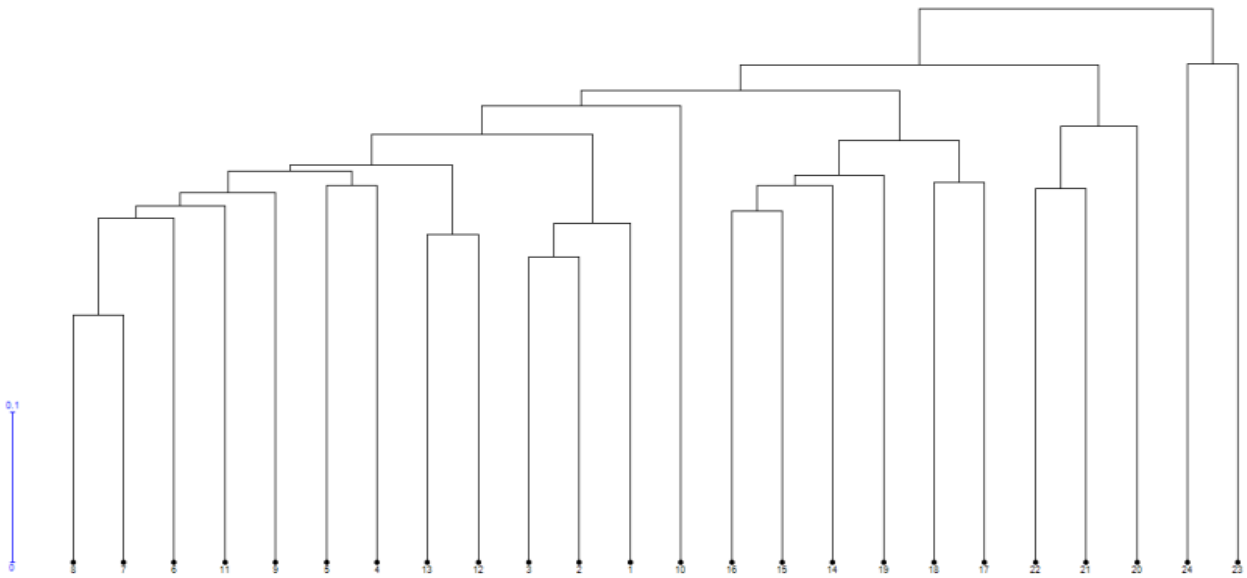


Fig.2 UPGMA dendrogram of 24 Indian-origin *S. officinarum* accessions

Principal Coordinate Analysis (PCoA)

A total of 24 principal coordinate axes were generated from the Jaccard's dissimilarity matrix of the 24 germplasm accessions. All 24 axes had non-negative eigenvalues, with the eigenvalue of the last axis (Axis 24) being negligible (≈0.000), consistent with the expected number of informative dimensions for a matrix of 24 accessions, no axis exhibited a negative eigenvalue. The first five axes together explained 51.857% of the total variation, while the first two axes (PC1 and PC2) together accounted for 27.485% of the total variation (Table 8).

The ordination plot (Fig. 3) showed a wide distribution of the 24 accessions across the first two principal coordinates. Along PC1, scores ranged from -0.3468 (Vellai) to 0.2549 (Kajla), while along PC2, scores ranged from -0.2899 (Pattapatti) to 0.5577 (Vellai). No accession pair recorded a zero Jaccard's dissimilarity coefficient in the pairwise analysis, and accordingly, no two accessions occupied an identical position in the ordination. A distinct group of six accessions, namely codes Pattacherukku 86, Pattapatti, Penang, Poona, Poovan and Rasdali, recorded negative scores on both PC1 and PC2 and were positioned separately from the remaining 18 accessions in the PCoA plot.

Table 8. Eigenvalues and percentage of variation explained by the first ten principal coordinate axes

Axis	Eigenvalue	Variation explained (%)	Cumulative variation (%)
PC1	0.6822	15.351	15.351
PC2	0.5392	12.134	27.485
PC3	0.444	9.99	37.475
PC4	0.3455	7.774	45.249
PC5	0.2936	6.608	51.857
PC6	0.2741	6.169	58.026
PC7	0.2403	5.408	63.433
PC8	0.1812	4.077	67.51
PC9	0.1718	3.866	71.376
PC10	0.1582	3.56	74.936

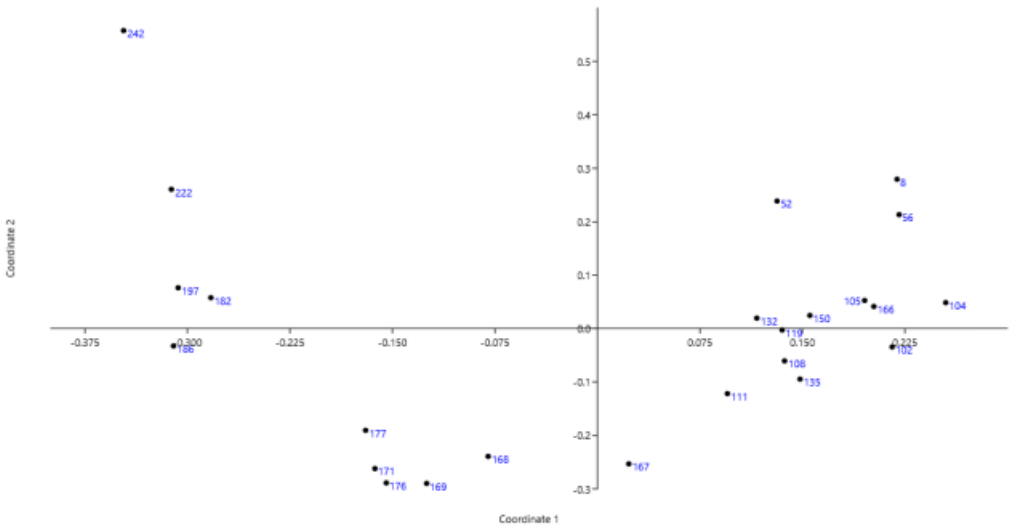


Fig.3 Principal Coordinate Analysis (PCoA) plot showing the genetic relationships among the accessions

DISCUSSION

The clear and reproducible amplification obtained across all twenty SSR markers, spanning a fragment size range of 100-500 bp, is consistent with the amplicon size ranges reported in recent comparative evaluations of genomic and EST-derived microsatellite panels in sugarcane. These studies similarly reported fragment sizes in the 90-500 bp range for markers belonging to the mSSCIR, SMC, NKS, and related series when applied to Indian *Saccharum* germplasm (Parthiban et al., 2018; Singh et al., 2020). A total of 241 alleles were generated across the twenty markers, with an average of 12.05 alleles per locus and 98.76 per cent overall polymorphism. The complex, highly heterozygous, aneuploid interspecific-hybrid genome of sugarcane allows multiple allelic size classes to occur at individual loci, a pattern that has been repeatedly reported in recent SSR-based population studies involving Indian breeding material, introgressed hybrids, and wild *Saccharum* species (Singh et al., 2020; Elumalai and Srinivasan, 2023). Only three markers, namely SMC336BS, SMC597CS, and SMC851MS, exhibited a single monomorphic allele each. This

represents a minor, locus-specific departure and most plausibly reflects the presence of a near-fixed allele within this particular, geographically restricted set of 24 accessions.

The mean PIC values recorded across the twenty markers ranged from 0.209 for mSSCIR1 to 0.467 for SMC361BS. Based on the informativeness classification proposed by Botstein et al. (1980), in which loci with PIC > 0.5 are considered highly informative, those with PIC values of 0.25-0.5 are regarded as reasonably informative, and values < 0.25 indicate slightly informative loci, the majority of the markers evaluated in the present study fall within the reasonably informative category. Among them, SMC361BS, SMC278CS, and UGSM417, which recorded mean PIC values > 0.44, approached the highly informative threshold. Notably, the greater discriminatory power observed for genomic SSR markers (SMC-series) compared with EST-derived SSR markers (mSSCIR-, UGSM-series) is consistent with the findings of a recent comparison of these two marker classes in Indian sugarcane germplasm. That study reported a substantially higher mean PIC for genomic SSR markers (0.72) than for EST-SSR markers (0.62), and attributed the lower informativeness of EST-SSRs to their occurrence in more conserved, transcribed genomic regions compared with the comparatively less constrained, non-coding regions targeted by genomic SSR primers (Parthiban et al., 2018). This provides a direct explanation for the lower informativeness observed for mSSCIR1, an EST-derived marker, which recorded both the lowest mean PIC (0.209) and the narrowest and most skewed allele frequency range (0.0417-0.2500; mean frequency 0.125) among all twenty markers evaluated in the present study.

The Analysis of Molecular Variance revealed that 83.5 per cent of the total molecular variance was distributed within, rather than among, the identified clusters (PhiPT = 0.165, P = 0.0001). Although the among-cluster component was statistically significant, indicating genuine genetic differentiation, the predominance of within-cluster variation is consistent with recent SSR-based diversity assessments of Indian sugarcane germplasm collections. These studies have similarly reported that a greater proportion of the total molecular variance is retained within rather than between clusters or sub-populations. This pattern has been attributed to the highly heterozygous and multi-allelic constitution of individual sugarcane genotypes, together with the historically narrow founder base underlying commercial and conserved Indian *Saccharum* material (Singh et al., 2020; Elumalai and Srinivasan, 2023).

The overall mean Shannon diversity index ($H' = 2.206 \pm 0.704$), Simpson's diversity index ($1-D = 0.847 \pm 0.178$), and Dominance index ($D = 0.153 \pm 0.178$) indicate a moderate-to-high level of allelic diversity accompanied by relatively low dominance across the marker panel. Markers such as mSSCIR24, SMC334BS, and UGSM417 consistently recorded the highest diversity and lowest dominance across all three indices, whereas mSSCIR1 consistently showed the lowest diversity and highest dominance. Several markers, including mSSCIR1, mSSCIR53, NKS23, SMC597CS, SMC545MS, and SMC851MS, also showed individual alleles fixed at the theoretical extremes ($H' = 0$; $D = 1.000$). Such locus-specific fixation within an otherwise diverse marker panel has been documented in recent diversity studies of introgressed and cultivated sugarcane germplasm and is generally attributed to the disproportionately high frequency of a single predominant allele within a genetically related subset of accessions. This pattern is particularly likely to occur in collections representing a defined regional or institutional breeding context rather than a globally representative sample (Elumalai and Srinivasan, 2023).

Jaccard's dissimilarity coefficient among the 24 accessions ranged from 0.3273 to 0.8162, with a mean of 0.6156 ± 0.0866. Notably, no pairwise comparison recorded a dissimilarity value of 0.000. This is an important finding when compared with recent SSR-based surveys of larger and more heterogeneous sugarcane collections, in which completely identical marker profiles have occasionally been detected between accession pairs and interpreted as evidence of clonal duplication or synonymous naming (Singh et al., 2020). The complete absence of identical pairs among the present 24 Indian-origin accessions indicates that each accession retained a molecularly distinguishable identity at the twenty SSR loci examined. This is an encouraging outcome for germplasm curation. The UPGMA dendrogram, cut at the mean dissimilarity coefficient, resolved the accessions into four clusters, with Cluster I alone comprising 79.17 per cent of the germplasm. This markedly skewed distribution is consistent with the AMOVA-based inference of a genetically cohesive collection. Principal Coordinate Analysis, with the first five axes explaining 51.857 per cent of the total variation, broadly corroborated this substructuring. The concordance between UPGMA and PCoA-based groupings is also consistent with patterns reported in recent sugarcane population-structure studies, where agreement between these approaches has been used to support the reliability of inferred genetic relationships (Xiong et al., 2022; Perera et al., 2023).

CONCLUSION

The present findings demonstrate that the twenty-marker SSR panel was effective in resolving genetic diversity among Indian-origin *S. officinarum* accessions. The consistently superior performance of genomic SSR markers (SMC-series) over EST-derived markers (mSSCIR-, UGSM-series) in the present study further supports recent methodological findings on marker-class efficiency in sugarcane. At the same time, the predominance of within-cluster molecular variance and the relatively narrow range of dissimilarity values indicate that the collection represents a genetically cohesive group, reflecting the historically limited breeding base of Indian sugarcane germplasm. This observation is consistent with recent regional diversity assessments.

Acknowledgements: The authors sincerely thank Dr. P. Govindaraj, Director, ICAR-Sugarcane Breeding Institute, Coimbatore, for his support in successfully carrying out the research work. The authors also gratefully acknowledge ICAR-Sugarcane Breeding Institute, Coimbatore, for providing the financial support required for this study. Research was supported by the Indian Council of Agricultural Research, Department of Agricultural Research an Education, Government of India.

Author Contributions: S.S. and N.M. contributed equally to this work and share first authorship. S.S. analysed the data, and prepared the manuscript; N.M. conceptualized the study, designed the research methodology, carried out the

research work and reviewed the manuscript; S.P. carried out the research work; C.S. assisted in research work; P.P.T assisted in data analysis and reviewed the manuscript and S.R. student's internal guide and reviewed the manuscript.

Data Availability: All relevant data are included within the manuscript.

Statements and Declarations

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

Ethics approval: This article does not contain any studies with human participants.

REFERENCES

1. Aitken KS, Li JC, Jackson P, Piperidis G, McIntyre CL (2006) AFLP analysis of genetic diversity within Saccharum officinarum and comparison with sugarcane cultivars. Australian Journal of Agricultural Research 57(11):1167–1184
2. Aitken K, Li J, Piperidis G, Qing C, Yuanhong F, Jackson P (2018) Worldwide genetic diversity of the wild species Saccharum spontaneum and level of diversity captured within sugarcane breeding programs. Crop Science 58(1):218–229
3. Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. American Journal of Human Genetics 32(3):314–331
4. Brown JS, Schnell RJ, Power EJ, Douglas SL, Kuhn DN (2007) Analysis of clonal germplasm from five Saccharum species: S. barberi, S. robustum, S. officinarum, S. sinense and S. spontaneum. A study of inter- and intra-species relationships using microsatellite markers. Genetic Resources and Crop Evolution 54(3):627–648
5. D'Hont A, Ison D, Alix K, Roux C, Glaszmann JC (1998) Determination of basic chromosome numbers in the genus Saccharum by physical mapping of ribosomal RNA genes. Genome 41(2):221–225
6. Elumalai K, Srinivasan A (2023) Genetic diversity and population structure of sugarcane introgressed hybrids by SSR markers. 3 Biotech 13:399
7. Govindaraj P, Balasundaram N, Sharma TR, Bansal KC, Koundal KR, Singh NK (2009) Development of new STMS markers for sugarcane. Sugarcane Breeding Institute, Coimbatore
8. Khan FA, Khan A, Azhar FM, Rauf S (2009) Genetic diversity of Saccharum officinarum accessions in Pakistan as revealed by random amplified polymorphic DNA. Genetics and Molecular Research 8(4):1376–1382
9. Mahadevaiah C, Appunu C, Aitken K, Suresha GS, Vignesh P, Mahadeva Swamy HK, Valarmathi R, Hemaprabha G, Alagarasan G, Ram B (2021) Genomic selection in sugarcane: Current status and future prospects. Frontiers in Plant Science 12:708233
10. Mary S, Nair NV, Chaturvedi PK, Selvi A (2006) Analysis of genetic diversity among Saccharum spontaneum L. from four geographical regions of India, using molecular markers. Genetic Resources and Crop Evolution 53(6):1221–1231
11. Nair NV, Nagarajan R, Amalraj VA (2006) Saccharum germplasm collection from the Cauvery river basin and coastal Tamil Nadu, India. IPGRI Plant Genetic Resources Newsletter 146:56–59
12. Nair NV, Sekharan S (2009) Saccharum germplasm collection in Mizoram, India. Sugar Tech 11(3):288–291
13. Nayak SN, Song J, Villa A, Pathak B, Ayala-Silva T, Yang X, Todd J, Glynn NC, Kuhn DN, Glaz B, Gilbert RA, Comstock JC (2014) Promoting utilization of Saccharum spp. genetic resources through genetic diversity analysis and core collection construction. PLoS ONE 9(10):e110856
14. Pan YB (2006) Highly polymorphic microsatellite DNA markers for sugarcane germplasm evaluation and variety identity testing. Sugar Tech 8:246–256
15. Parthiban S, Govindaraj P, Senthilkumar S (2018) Comparison of relative efficiency of genomic SSR and EST-SSR markers in estimating genetic diversity in sugarcane. 3 Biotech 8(3):144
16. Perera MF, Ostengo S, Malavera ANP, Balsalobre TWA, Onorato GD, Noguera AS, Hoffmann HP, Carneiro MS (2023) Genetic diversity and population structure of Saccharum hybrids. PLoS One 18(8):e0289504
17. Raj P, Selvi A, Prathima PT, Nair NV (2016) Analysis of genetic diversity of Saccharum complex using chloroplast microsatellite markers. Sugar Tech 18(2):141–148
18. Saravanakumar K, Govindaraj P, Appunu C, Senthilkumar S, Kumar R (2014) Analysis of genetic diversity in high biomass producing sugarcane hybrids (Saccharum spp. complex) using RAPD and STMS markers. Indian Journal of Biotechnology 13:214–220
19. Selvi A, Nair NV, Balasundaram N, Mohapatra T (2003) Evaluation of maize microsatellite markers for genetic diversity analysis and fingerprinting in sugarcane. Genome 46(3):394–403
20. Selvi A, Nair NV, Noyer JL, Singh NK, Balasundaram N, Bansal KC, Koundal KR, Mohapatra T (2005) Genomic constitution and genetic relationship among the tropical and subtropical Indian sugarcane cultivars revealed by AFLP. Crop Science 45(5):1750–1757
21. Selvi A, Nair NV, Noyer JL, Singh NK, Balasundaram N, Bansal KC, Koundal KR, Mohapatra T (2006) AFLP analysis of the phenetic organization and genetic diversity in the sugarcane complex, Saccharum and Erianthus. Genetic Resources and Crop Evolution 53(4):831–842
22. Singh RB, Mahenderakar MD, Jugran AK, Singh RK, Srivastava RK (2020) Assessing genetic diversity and population structure of sugarcane cultivars, progenitor species and genera using microsatellite (SSR) markers. Gene 753:144800
23. Singh RK, Srivastava S, Singh SP, Sharma ML, Mohopatra T, Singh NK, Singh SB (2008) Identification of new microsatellite DNA markers for sugar and related traits in sugarcane. Sugar Tech 10(4):327–333
24. Siraree A, Banerjee N, Kumar S, Khan MS, Singh PK, Kumar S, Sharma S, Singh RK, Singh J (2018) Agro-morphological description, genetic diversity and population structure of sugarcane varieties from sub-tropical India. 3 Biotech 8(11):469

25. Ullah SMS, Hossain MA, Hossain MM, Barman S et al (2013) Genetic diversity analysis of chewing sugarcane (Saccharum officinarum L.) varieties by using RAPD markers. Journal of Bioscience and Biotechnology 2:145–150
26. Waclawovsky AJ, Sato PM, Lembke CG, Moore PH, Souza GM (2010) Sugarcane for bioenergy production: an assessment of yield and regulation of sucrose content. Plant Biotechnology Journal 8(3):263–276
27. Xiong H, Chen Y, Gao SJ, Pan YB, Shi A (2022) Population structure and genetic diversity analysis in sugarcane (Saccharum spp. hybrids) and six related Saccharum species. Agronomy 12(2):412

SUPPLEMENTARY FILES

Supplementary tables

Table. S1 Details of the 24 selected Indian-origin *S. officinarum* accessions

S. No	Accession names
1	Assam Red
2	Chittan
3	Desi Pounda
4	Javari Kabbu
5	Kajla
6	Kaludai Boothan
7	Kea 21
8	Khajuria
9	Lakhapur
10	Malabar
11	Manjri Red
12	Namam
13	Pakaweli-2
14	Pakaweli-2 Striped
15	Pattacherukku 86
16	Pattapatti
17	Penang
18	Poona
19	Poovan
20	Ramgarh
21	Rasdali
22	Saharanpur Black
23	Thellacherukku
24	Vellai

Note: (S. No = Sample ID): 1=8, 2=52, 3=56, 4=102, 5=104, 6=105, 7=108, 8=111, 9=119, 10=132, 11=135, 12=150, 13=166, 14=167, 15=168, 16=169, 17=171, 18=176, 19=177, 20=182, 21=186, 22=197, 23=222, 24=242

Table. S2 Allele frequencies and polymorphic information content (PIC) values of individual alleles detected by 20 SSR markers in 24 accessions

SSR marker	Allele size (bp)	Allele frequency	PIC value
mSSCIR9	100	0.2917	0.413
	170	0.5417	0.497
	206	0.8333	0.278
	250	0.7083	0.413
	320	0.4167	0.486
	460	0.4583	0.497
mSSCIR24	125	0.2083	0.33
	170	0.25	0.375

	226	0.7917	0.33
	250	0.9583	0.08
	306	0.7917	0.33
	350	0.625	0.469
	415	0.875	0.219
	470	0.9167	0.153
mSSCIR68	160	0.25	0.375
	214	0.7917	0.33
	240	0.5833	0.486
	270	0.2083	0.33
	352	0.2917	0.413
	386	0.2917	0.413
	453	0.25	0.375
SMC278CS	115	0.2083	0.33
	158	0.2917	0.413
	185	0.3333	0.444
	214	0.6667	0.444
	251	0.5833	0.486
	340	0.4583	0.497
	384	0.5417	0.497
	467	0.625	0.469
SMC361BS	102	0.3333	0.444
	197	0.7083	0.413
	226	0.4583	0.497
	284	0.4167	0.486
	336	0.5	0.5
	386	0.4167	0.486
	488	0.3333	0.444
SMC334BS	113	0.25	0.375
	160	0.7917	0.33
	187	0.875	0.219
	203	0.7083	0.413
	240	0.4167	0.486
	333	0.5	0.5
	360	0.625	0.469
	465	0.5417	0.497
UGSM417	106	0.6667	0.444
	170	0.5833	0.486
	230	0.3333	0.444
	273	0.6667	0.444
	300	0.8333	0.278
	332	0.5417	0.497
	355	0.375	0.469
	488	0.5	0.5
NKS25	113	0.125	0.219
	140	0.0833	0.153
	186	0.625	0.469

	203	0.5417	0.497
	243	0.4583	0.497
	270	0.4583	0.497
	365	0.8333	0.278
	464	0.5833	0.486
NKS32	105	0.4167	0.486
	181	0.3333	0.444
	203	0.25	0.375
	250	0.1667	0.278
	313	0.125	0.219
	424	0.125	0.219
mSSCIR1	103	0.0833	0.153
	135	0.125	0.219
	170	0.25	0.375
	225	0.0417	0.08
	260	0.1667	0.278
	490	0.0833	0.153
SMC336BS	120	0.75	0.375
	179	0.7083	0.413
	209	0.75	0.375
	222	0.4167	0.486
	243	0.4167	0.486
	278	0.4167	0.486
	315	0.4167	0.486
	350	0.125	0.219
	376	0.375	0.469
	403	0.5417	0.497
	468	0.625	0.469
	499	1	0
mSSCIR53	100	0.0417	0.08
	170	0.2083	0.33
	195	0.2917	0.413
	219	0.1667	0.278
	230	0.4167	0.486
	263	0.375	0.469
	285	0.75	0.375
	314	0.4583	0.497
	334	0.2917	0.413
	350	0.75	0.375
	392	0.375	0.469
	418	0.375	0.469
	453	0.4167	0.486
	480	0.25	0.375
	500	0.6667	0.444
NKS23	105	0.0417	0.08
	130	0.25	0.375
	166	0.3333	0.444

	199	0.4167	0.486
	244	0.5417	0.497
	260	0.6667	0.444
	294	0.5417	0.497
	307	0.6667	0.444
	367	0.625	0.469
	400	0.625	0.469
	469	0.6667	0.444
	500	0.8333	0.278
NKS28	100	0.375	0.469
	113	0.4583	0.497
	146	0.875	0.219
	169	0.5417	0.497
	192	0.2917	0.413
	211	0.7083	0.413
	228	0.4583	0.497
	258	0.5833	0.486
	279	0.5417	0.497
	312	0.2917	0.413
	349	0.3333	0.444
	377	0.125	0.219
	385	0.25	0.375
	417	0.2083	0.33
	437	0.0833	0.153
	458	0.25	0.375
	500	0.5833	0.486
SMC597CS	107	0.0417	0.08
	120	0.25	0.375
	140	0.0833	0.153
	160	0.625	0.469
	177	0.75	0.375
	197	0.5	0.5
	217	0.625	0.469
	235	0.4583	0.497
	246	0.6667	0.444
	260	0.7083	0.413
	288	0.5833	0.486
	323	1	0
	352	0.4167	0.486
	378	0.0833	0.153
	403	0.5	0.5
	432	0.375	0.469
	447	0.5417	0.497
	482	0.0417	0.08
	500	0.4167	0.486
NKS38	100	0.4167	0.486
	141	0.625	0.469

	189	0.6667	0.444
	200	0.1667	0.278
	223	0.3333	0.444
	239	0.75	0.375
	278	0.4167	0.486
	292	0.5	0.5
	321	0.4167	0.486
	341	0.5833	0.486
	389	0.375	0.469
	409	0.625	0.469
	436	0.3333	0.444
	459	0.2083	0.33
	472	0.3333	0.444
	488	0.2917	0.413
	500	0.5833	0.486
NKS24	100	0.5833	0.486
	145	0.4583	0.497
	155	0.5833	0.486
	163	0.75	0.375
	189	0.625	0.469
	204	0.4583	0.497
	214	0.25	0.375
	227	0.5417	0.497
	251	0.1667	0.278
	262	0.4583	0.497
	271	0.6667	0.444
	300	0.625	0.469
	313	0.2083	0.33
	334	0.2917	0.413
	347	0.2083	0.33
	361	0.25	0.375
	376	0.3333	0.444
	394	0.2917	0.413
	429	0.4583	0.497
	440	0.4583	0.497
	458	0.125	0.219
	500	0.5833	0.486
NKS49	100	0.5417	0.497
	119	0.5	0.5
	141	0.7917	0.33
	152	0.75	0.375
	165	0.4583	0.497
	183	0.4167	0.486
	195	0.5833	0.486
	217	0.5	0.5
	230	0.7083	0.413
	254	0.5	0.5

	271	0.6667	0.444
	292	0.3333	0.444
	314	0.4167	0.486
	320	0.4583	0.497
	357	0.5417	0.497
	401	0.8333	0.278
	442	0.125	0.219
	484	0.2083	0.33
	500	0.9583	0.08
SMC545MS	132	0.7917	0.33
	148	0.7917	0.33
	169	0.7083	0.413
	185	0.4583	0.497
	193	0.6667	0.444
	217	0.25	0.375
	226	0.2917	0.413
	250	0.25	0.375
	266	0.5833	0.486
	296	0.2917	0.413
	317	0.0833	0.153
	330	0.5417	0.497
	347	0.1667	0.278
	359	0.2083	0.33
	392	0.0833	0.153
	427	0.125	0.219
	441	0.125	0.219
	464	0.0417	0.08
	500	0.2083	0.33
SMC851MS	100	0.9167	0.153
	120	0.9167	0.153
	138	1	0
	169	0.9167	0.153
	194	0.7083	0.413
	221	0.6667	0.444
	246	0.7083	0.413
	279	0.375	0.469
	294	0.125	0.219
	342	0.0417	0.08
	362	0.3333	0.444
	386	0.4167	0.486
	419	0.3333	0.444
	433	0.25	0.375
	450	0.0833	0.153
	478	0.1667	0.278
	500	0.6667	0.444

Table. S3 Allele-wise Shannon diversity index (H′), Simpson's diversity index (1-D), and Dominance index (D) values of 241 SSR alleles

Marker/alleles	Shannon (H)	Simpson (1-D)	Dominance (D)	Marker/alleles	Shannon (H)	Simpson (1-D)	Dominance (D)
mSSCIR9 - 1	1.946	0.857	0.143	NKS28 - 11	2.079	0.875	0.125
mSSCIR9 - 2	2.565	0.923	0.077	NKS28 - 12	1.099	0.667	0.333
mSSCIR9 - 3	2.996	0.95	0.05	NKS28 - 13	1.792	0.833	0.167
mSSCIR9 - 4	2.833	0.941	0.059	NKS28 - 14	1.609	0.8	0.2
mSSCIR9 - 5	2.303	0.9	0.1	NKS28 - 15	0.6931	0.5	0.5
mSSCIR9 - 6	2.398	0.909	0.091	NKS28 - 16	1.792	0.833	0.167
mSSCIR24 - 1	1.609	0.8	0.2	NKS28 - 17	2.639	0.929	0.071
mSSCIR24 - 2	1.792	0.833	0.167	SMC597CS - 1	0	0	1
mSSCIR24 - 3	2.944	0.947	0.053	SMC597CS - 2	1.792	0.833	0.167
mSSCIR24 - 4	3.135	0.957	0.043	SMC597CS - 3	0.6931	0.5	0.5
mSSCIR24 - 5	2.944	0.947	0.053	SMC597CS - 4	2.708	0.933	0.067
mSSCIR24 - 6	2.708	0.933	0.067	SMC597CS - 5	2.89	0.944	0.056
mSSCIR24 - 7	3.045	0.952	0.048	SMC597CS - 6	2.485	0.917	0.083
mSSCIR24 - 8	3.091	0.955	0.045	SMC597CS - 7	2.708	0.933	0.067
mSSCIR68 - 1	1.792	0.833	0.167	SMC597CS - 8	2.398	0.909	0.091
mSSCIR68 - 2	2.944	0.947	0.053	SMC597CS - 9	2.773	0.938	0.063
mSSCIR68 - 3	2.639	0.929	0.071	SMC597CS - 10	2.833	0.941	0.059
mSSCIR68 - 4	1.609	0.8	0.2	SMC597CS - 11	2.639	0.929	0.071
mSSCIR68 - 5	1.946	0.857	0.143	SMC597CS - 12	3.178	0.958	0.042
mSSCIR68 - 6	1.946	0.857	0.143	SMC597CS - 13	2.303	0.9	0.1
mSSCIR68 - 7	1.792	0.833	0.167	SMC597CS - 14	0.6931	0.5	0.5
SMC278CS - 1	1.609	0.8	0.2	SMC597CS - 15	2.485	0.917	0.083
SMC278CS - 2	1.946	0.857	0.143	SMC597CS - 16	2.197	0.889	0.111
SMC278CS - 3	2.079	0.875	0.125	SMC597CS - 17	2.565	0.923	0.077
SMC278CS - 4	2.773	0.938	0.063	SMC597CS - 18	0	0	1
SMC278CS - 5	2.639	0.929	0.071	SMC597CS - 19	2.303	0.9	0.1
SMC278CS - 6	2.398	0.909	0.091	NKS38 - 1	2.303	0.9	0.1
SMC278CS - 7	2.565	0.923	0.077	NKS38 - 2	2.708	0.933	0.067
SMC278CS - 8	2.708	0.933	0.067	NKS38 - 3	2.773	0.938	0.063
SMC361BS - 1	2.079	0.875	0.125	NKS38 - 4	1.386	0.75	0.25
SMC361BS - 2	2.833	0.941	0.059	NKS38 - 5	2.079	0.875	0.125
SMC361BS - 3	2.398	0.909	0.091	NKS38 - 6	2.89	0.944	0.056
SMC361BS - 4	2.303	0.9	0.1	NKS38 - 7	2.303	0.9	0.1
SMC361BS - 5	2.485	0.917	0.083	NKS38 - 8	2.485	0.917	0.083
SMC361BS - 6	2.303	0.9	0.1	NKS38 - 9	2.303	0.9	0.1
SMC361BS - 7	2.079	0.875	0.125	NKS38 - 10	2.639	0.929	0.071
SMC334BS - 1	1.792	0.833	0.167	NKS38 - 11	2.197	0.889	0.111
SMC334BS - 2	2.944	0.947	0.053	NKS38 - 12	2.708	0.933	0.067
SMC334BS - 3	3.045	0.952	0.048	NKS38 - 13	2.079	0.875	0.125
SMC334BS - 4	2.833	0.941	0.059	NKS38 - 14	1.609	0.8	0.2
SMC334BS - 5	2.303	0.9	0.1	NKS38 - 15	2.079	0.875	0.125
SMC334BS - 6	2.485	0.917	0.083	NKS38 - 16	1.946	0.857	0.143
SMC334BS - 7	2.708	0.933	0.067	NKS38 - 17	2.639	0.929	0.071

SMC334BS - 8	2.565	0.923	0.077	NKS24 - 1	2.639	0.929	0.071
UGSM417 - 1	2.773	0.938	0.063	NKS24 - 2	2.398	0.909	0.091
UGSM417 - 2	2.639	0.929	0.071	NKS24 - 3	2.639	0.929	0.071
UGSM417 - 3	2.079	0.875	0.125	NKS24 - 4	2.89	0.944	0.056
UGSM417 - 4	2.773	0.938	0.063	NKS24 - 5	2.708	0.933	0.067
UGSM417 - 5	2.996	0.95	0.05	NKS24 - 6	2.398	0.909	0.091
UGSM417 - 6	2.565	0.923	0.077	NKS24 - 7	1.792	0.833	0.167
UGSM417 - 7	2.197	0.889	0.111	NKS24 - 8	2.565	0.923	0.077
UGSM417 - 8	2.485	0.917	0.083	NKS24 - 9	1.386	0.75	0.25
NKS25 - 1	1.099	0.667	0.333	NKS24 - 10	2.398	0.909	0.091
NKS25 - 2	0.693	0.5	0.5	NKS24 - 11	2.773	0.938	0.063
NKS25 - 3	2.708	0.933	0.067	NKS24 - 12	2.708	0.933	0.067
NKS25 - 4	2.565	0.923	0.077	NKS24 - 13	1.609	0.8	0.2
NKS25 - 5	2.398	0.909	0.091	NKS24 - 14	1.946	0.857	0.143
NKS25 - 6	2.398	0.909	0.091	NKS24 - 15	1.609	0.8	0.2
NKS25 - 7	2.996	0.95	0.05	NKS24 - 16	1.792	0.833	0.167
NKS25 - 8	2.639	0.929	0.071	NKS24 - 17	2.079	0.875	0.125
NKS32 - 1	2.303	0.9	0.1	NKS24 - 18	1.946	0.857	0.143
NKS32 - 2	2.079	0.875	0.125	NKS24 - 19	2.398	0.909	0.091
NKS32 - 3	1.792	0.833	0.167	NKS24 - 20	2.398	0.909	0.091
NKS32 - 4	1.386	0.75	0.25	NKS24 - 21	1.099	0.667	0.333
NKS32 - 5	1.099	0.667	0.333	NKS24 - 22	2.639	0.929	0.071
NKS32 - 6	1.099	0.667	0.333	NKS49 - 1	2.565	0.923	0.077
mSSCIR1 - 1	0.693	0.5	0.5	NKS49 - 2	2.485	0.917	0.083
mSSCIR1 - 2	1.099	0.667	0.333	NKS49 - 3	2.944	0.947	0.053
mSSCIR1 - 3	1.792	0.833	0.167	NKS49 - 4	2.89	0.944	0.056
mSSCIR1 - 4	0	0	1	NKS49 - 5	2.398	0.909	0.091
mSSCIR1 - 5	1.386	0.75	0.25	NKS49 - 6	2.303	0.9	0.1
mSSCIR1 - 6	0.693	0.5	0.5	NKS49 - 7	2.639	0.929	0.071
SMC336BS - 1	2.89	0.944	0.056	NKS49 - 8	2.485	0.917	0.083
SMC336BS - 2	2.833	0.941	0.059	NKS49 - 9	2.833	0.941	0.059
SMC336BS - 3	2.89	0.944	0.056	NKS49 - 10	2.485	0.917	0.083
SMC336BS - 4	2.303	0.9	0.1	NKS49 - 11	2.773	0.938	0.063
SMC336BS - 5	2.303	0.9	0.1	NKS49 - 12	2.079	0.875	0.125
SMC336BS - 6	2.303	0.9	0.1	NKS49 - 13	2.303	0.9	0.1
SMC336BS - 7	2.303	0.9	0.1	NKS49 - 14	2.398	0.909	0.091
SMC336BS - 8	1.099	0.667	0.333	NKS49 - 15	2.565	0.923	0.077
SMC336BS - 9	2.197	0.889	0.111	NKS49 - 16	2.996	0.95	0.05
SMC336BS - 10	2.565	0.923	0.077	NKS49 - 17	1.099	0.667	0.333
SMC336BS - 11	2.708	0.933	0.067	NKS49 - 18	1.609	0.8	0.2
SMC336BS - 12	3.178	0.958	0.042	NKS49 - 19	3.135	0.957	0.043
mSSCIR53 - 1	0	0	1	SMC545MS - 1	2.944	0.947	0.053
mSSCIR53 - 2	1.609	0.8	0.2	SMC545MS - 2	2.944	0.947	0.053
mSSCIR53 - 3	1.946	0.857	0.143	SMC545MS - 3	2.833	0.941	0.059
mSSCIR53 - 4	1.386	0.75	0.25	SMC545MS - 4	2.398	0.909	0.091
mSSCIR53 - 5	2.303	0.9	0.1	SMC545MS - 5	2.773	0.938	0.063
mSSCIR53 - 6	2.197	0.889	0.111	SMC545MS - 6	1.792	0.833	0.167

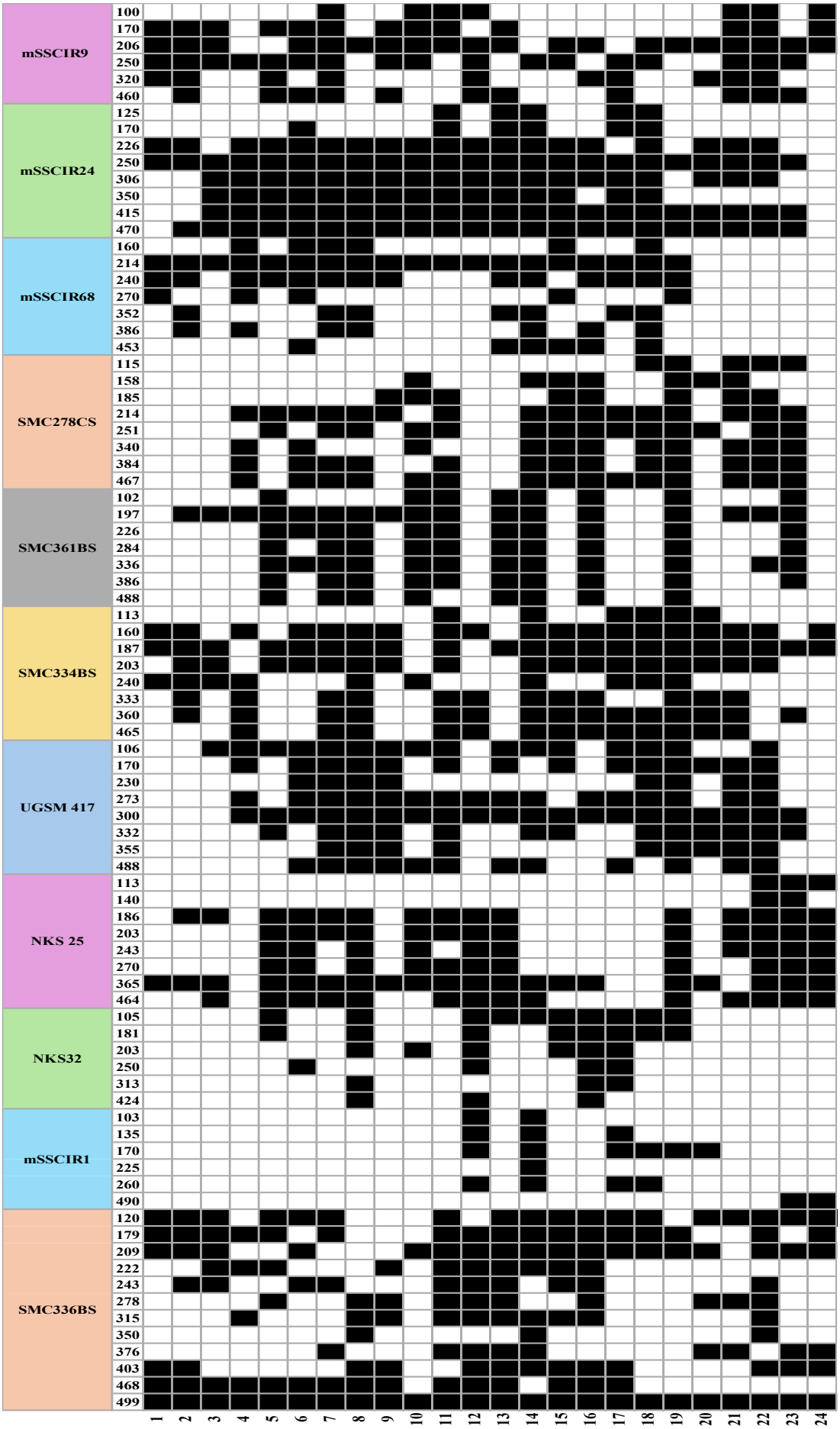
mSSCIR53 - 7	2.89	0.944	0.056	SMC545MS - 7	1.946	0.857	0.143
mSSCIR53 - 8	2.398	0.909	0.091	SMC545MS - 8	1.792	0.833	0.167
mSSCIR53 - 9	1.946	0.857	0.143	SMC545MS - 9	2.639	0.929	0.071
mSSCIR53 - 10	2.89	0.944	0.056	SMC545MS - 10	1.946	0.857	0.143
mSSCIR53 - 11	2.197	0.889	0.111	SMC545MS - 11	0.693	0.5	0.5
mSSCIR53 - 12	2.197	0.889	0.111	SMC545MS - 12	2.565	0.923	0.077
mSSCIR53 - 13	2.303	0.9	0.1	SMC545MS - 13	1.386	0.75	0.25
mSSCIR53 - 14	1.792	0.833	0.167	SMC545MS - 14	1.609	0.8	0.2
mSSCIR53 - 15	2.773	0.938	0.063	SMC545MS - 15	0.693	0.5	0.5
NKS23 - 1	0	0	1	SMC545MS - 16	1.099	0.667	0.333
NKS23 - 2	1.792	0.833	0.167	SMC545MS - 17	1.099	0.667	0.333
NKS23 - 3	2.079	0.875	0.125	SMC545MS - 18	0	0	1
NKS23 - 4	2.303	0.9	0.1	SMC545MS - 19	1.609	0.8	0.2
NKS23 - 5	2.565	0.923	0.077	SMC851MS - 1	3.091	0.955	0.045
NKS23 - 6	2.773	0.938	0.063	SMC851MS - 2	3.091	0.955	0.045
NKS23 - 7	2.565	0.923	0.077	SMC851MS - 3	3.178	0.958	0.042
NKS23 - 8	2.773	0.938	0.063	SMC851MS - 4	3.091	0.955	0.045
NKS23 - 9	2.708	0.933	0.067	SMC851MS - 5	2.833	0.941	0.059
NKS23 - 10	2.708	0.933	0.067	SMC851MS - 6	2.773	0.938	0.063
NKS23 - 11	2.773	0.938	0.063	SMC851MS - 7	2.833	0.941	0.059
NKS23 - 12	2.996	0.95	0.05	SMC851MS - 8	2.197	0.889	0.111
NKS28 - 1	2.197	0.889	0.111	SMC851MS - 9	1.099	0.667	0.333
NKS28 - 2	2.398	0.909	0.091	SMC851MS - 10	0	0	1
NKS28 - 3	3.045	0.952	0.048	SMC851MS - 11	2.079	0.875	0.125
NKS28 - 4	2.565	0.923	0.077	SMC851MS - 12	2.303	0.9	0.1
NKS28 - 5	1.946	0.857	0.143	SMC851MS - 13	2.079	0.875	0.125
NKS28 - 6	2.833	0.941	0.059	SMC851MS - 14	1.792	0.833	0.167
NKS28 - 7	2.398	0.909	0.091	SMC851MS - 15	0.693	0.5	0.5
NKS28 - 8	2.639	0.929	0.071	SMC851MS - 16	1.386	0.75	0.25
NKS28 - 9	2.565	0.923	0.077	SMC851MS - 17	2.773	0.938	0.063
NKS28 - 10	1.946	0.857	0.143				

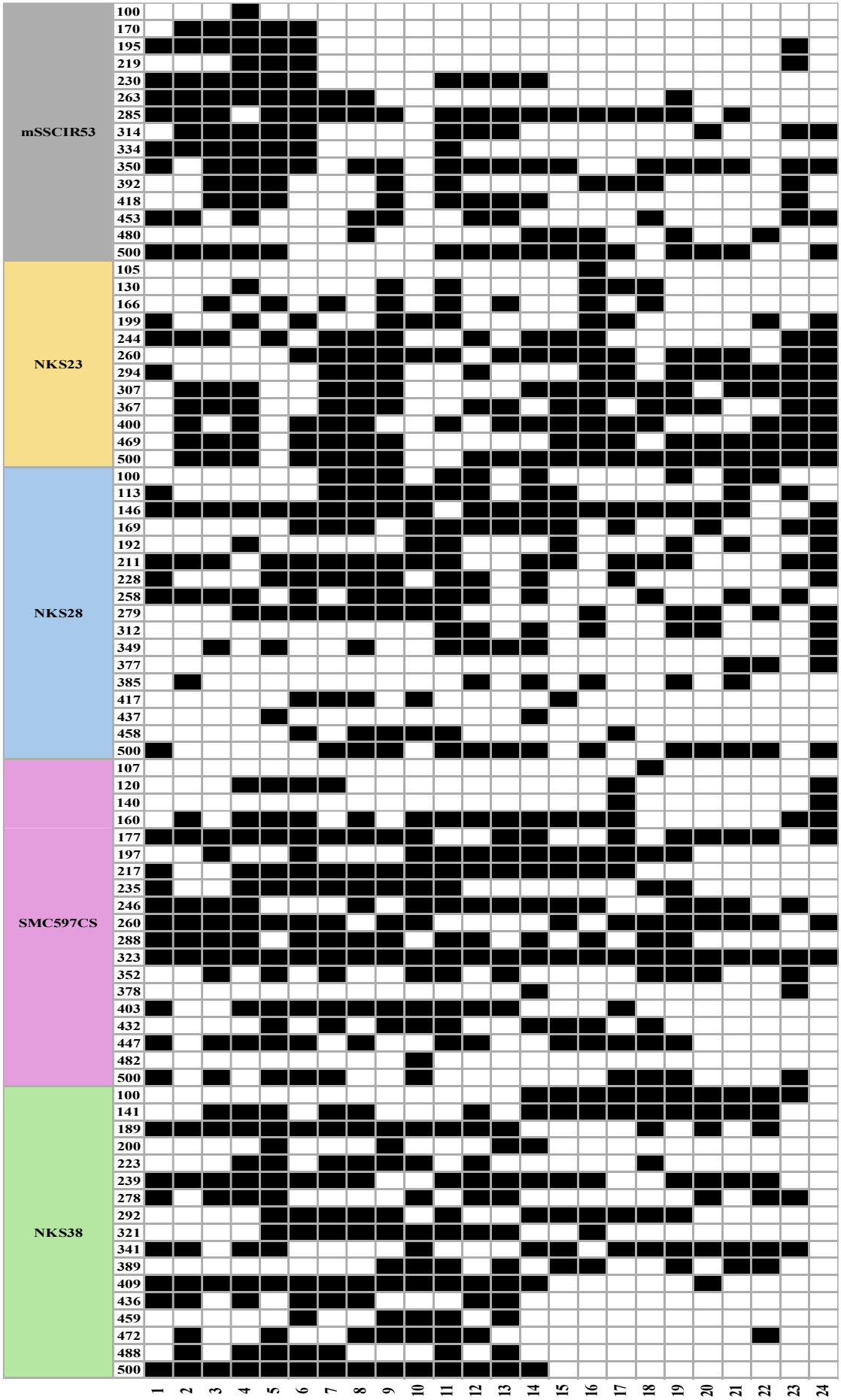
Supplementary figures

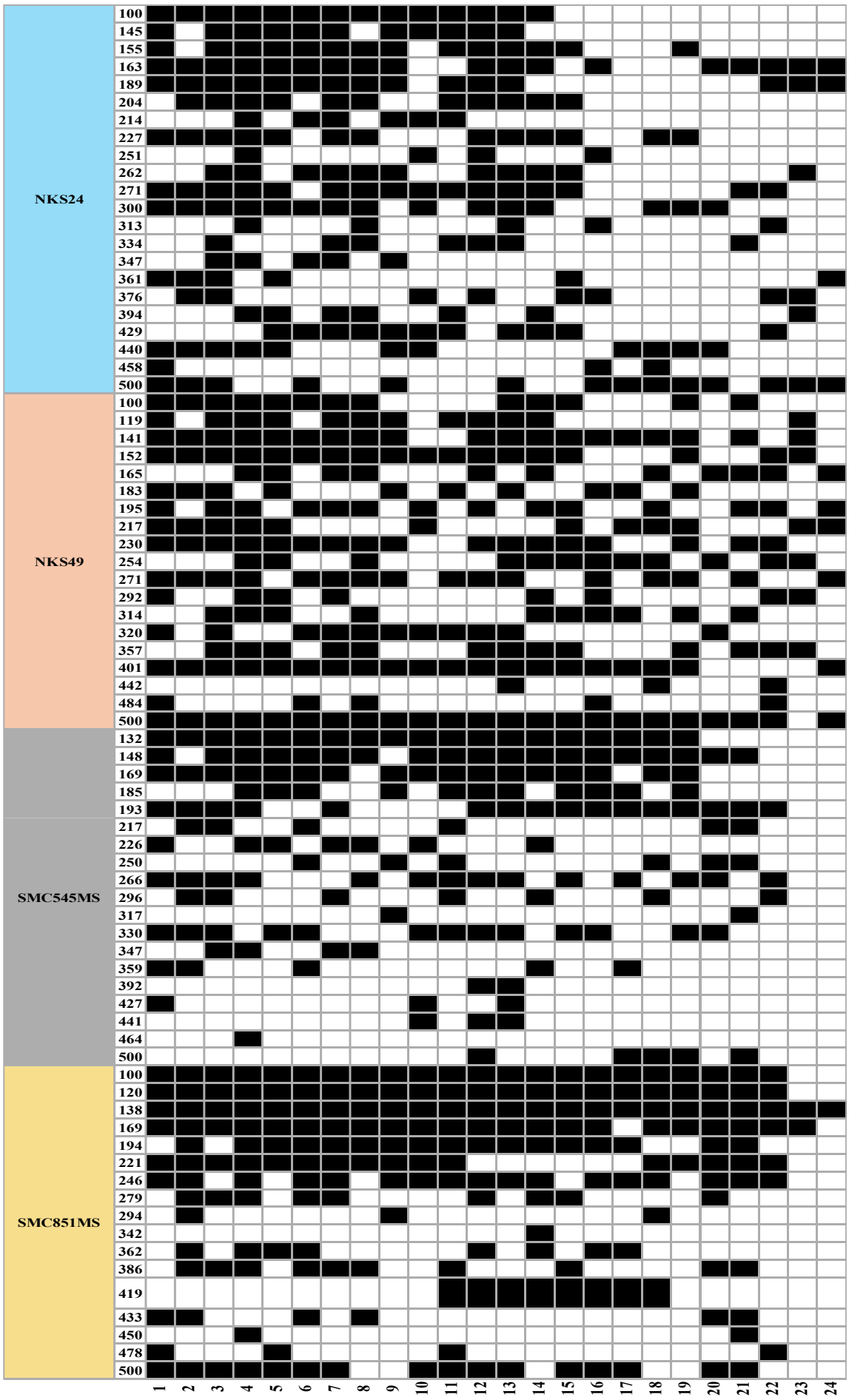
Fig. S2 Complete pairwise Jaccard's dissimilarity matrix for the 24 germplasm accessions

Units	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
2	0.4331																						
3	0.4662	0.4046																					
4	0.5484	0.5256	0.4839																				
5	0.5455	0.5776	0.4805	0.5000																			
6	0.5484	0.5350	0.5220	0.4941	0.5088																		
7	0.5904	0.5602	0.5212	0.4770	0.4651	0.4226																	
8	0.6199	0.6149	0.5943	0.5000	0.5140	0.4915	0.3273																
9	0.5890	0.6209	0.5882	0.5723	0.5610	0.5000	0.4720	0.5060															
10	0.6345	0.6987	0.6581	0.6450	0.5750	0.5506	0.5749	0.5965	0.6000														
11	0.6512	0.6457	0.5765	0.5860	0.5029	0.4709	0.4545	0.4945	0.4845	0.5250													
12	0.5796	0.5750	0.5346	0.5402	0.5625	0.5810	0.5307	0.4943	0.5928	0.6325	0.5254												
13	0.5975	0.5839	0.5159	0.5810	0.4759	0.4970	0.4971	0.5028	0.5494	0.5723	0.4371	0.4348											
14	0.6552	0.6185	0.6136	0.5519	0.5333	0.5979	0.4693	0.4581	0.6257	0.6554	0.5217	0.5225	0.5141										
15	0.6601	0.6000	0.5570	0.5432	0.6095	0.5868	0.5503	0.5298	0.6352	0.6429	0.5789	0.5644	0.5904	0.4783									
16	0.6933	0.6375	0.6805	0.6215	0.6271	0.6292	0.5944	0.5593	0.6196	0.6848	0.5657	0.5765	0.5765	0.5205	0.4658								
17	0.6892	0.6644	0.7006	0.6446	0.6587	0.6196	0.6552	0.6514	0.6536	0.7059	0.6590	0.6566	0.6726	0.5818	0.5461	0.5411							
18	0.6821	0.6579	0.6429	0.6061	0.6763	0.6391	0.5858	0.6229	0.6209	0.7070	0.6457	0.6897	0.6897	0.5689	0.5616	0.6026	0.5038						
19	0.6474	0.6415	0.5924	0.6171	0.5824	0.5930	0.5407	0.5030	0.6319	0.6129	0.5690	0.6441	0.6047	0.5322	0.5000	0.5065	0.5839	0.5205					
20	0.6818	0.6336	0.6569	0.6859	0.7170	0.6859	0.6456	0.6748	0.6993	0.7101	0.6582	0.6556	0.6818	0.6790	0.6296	0.6503	0.6615	0.6439	0.6058				
21	0.7181	0.6575	0.6776	0.6545	0.7151	0.6380	0.5563	0.5964	0.6284	0.7086	0.6369	0.6420	0.6905	0.6250	0.5874	0.6364	0.6828	0.6483	0.5248	0.5259			
22	0.7152	0.6821	0.7006	0.6919	0.6746	0.6446	0.5818	0.5952	0.6267	0.6711	0.6433	0.6882	0.6485	0.6552	0.6333	0.6258	0.7067	0.6644	0.5933	0.6299	0.4961		
23	0.7778	0.7343	0.7083	0.7378	0.6795	0.7143	0.6909	0.6788	0.7467	0.7234	0.7186	0.7346	0.6859	0.6667	0.6923	0.7086	0.7606	0.7343	0.6597	0.7317	0.6992	0.6250	
24	0.7317	0.7222	0.7424	0.8089	0.7843	0.7450	0.7516	0.7547	0.7647	0.8060	0.7722	0.7415	0.7500	0.7975	0.7721	0.7847	0.7619	0.8162	0.7464	0.7290	0.7355	0.7000	0.6602

Fig. S1. SSR fingerprint profile of 24 Indian-origin *S. officinarum* accessions




mSSCIR53
NKS23
NKS28
SMC597CS
NKS38
100 170 195 219 230 263 285 314 334 350 392 418 453 480 500
105 130 166 199 244 260 294 307 367 400 469 500
100 113 146 169 192 211 228 258 279 312 349 377 385 417 437 458 500
107 120 140 160 177 197 217 235 246 260 288 323 352 378 403 432 447 482 500
100 141 189 200 223 239 278 292 321 341 389 409 436 459 472 488 500
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24


NKS24
NKS49
SMC545MS
SMC851MS
100 145 155 163 189 204 214 227 251 262 271 300 313 334 347 361 376 394 429 440 458 500
100 119 141 152 165 183 195 217 230 254 271 292 314 320 357 401 442 484 500
132 148 169 185 193
217 226 250 266 296 317 330 347 359 392 427 441 464 500
100 120 138 169 194 221 246 279 294 342 362 386 419 433 450 478 500
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24