

PHENOTYPIC AND MOLECULAR EVALUATION OF SUGARCANE (*SACCHARUM* SPP. HYBRID) CLONES FOR HIGH YIELD AND QUALITY TRAITS USING AN AUGMENTED DESIGN AND EST-SSR MARKERS

Subhash Kumar Chaudhary¹, Ravi Kant^{2*}, Dharm Nath Kamat³, Thakur Prasad Mahto⁴, Rajan Kumar⁵, Sanjay Kumar⁴, Avinash Singh⁶, Suraj Kumar⁷

¹Department of Genetics and Plant Breeding, Post Graduate College of Agriculture, Dr. Rajendra Prasad Central Agricultural University, Pusa, Samastipur, Bihar - 848125, India

²Regional Research Station, Department of Plant Breeding and Genetics, Madhopur, West Champaran, Bihar-845454 (RPCAU Pusa), India

³Sugarcane Research Institute, Bihar-843125 (RPCAU Pusa), India

⁴PDUCH & F, Piprakothi, East Champaran, Bihar-845429 (RPCAU Pusa), India

⁵PDUCH & F, Piprakothi, East Champaran, Bihar-845429 (RPCAU Pusa), India

⁶PhD scholar, Department of Plant Breeding and Genetics, Punjab Agricultural University, Ludhiana-141004, India

⁷Regional Research Station, Madhopur, West Champaran, Bihar-845454 (RPCAU Pusa), India

*Corresponding author: Ravi Kant, Email: ravikant@rpcau.ac.in

ABSTRACT

The present investigation was undertaken at the Kalyanpur Farm of the Sugarcane Research Institute, Dr. Rajendra Prasad Central Agricultural University, Pusa, Samastipur, Bihar to assess genetic variability among sugarcane clones for yield and quality traits to examine the association among yield, quality and related attributes and to characterize the clones at the molecular level using EST-SSR markers. Forty-five CoXK-series clones derived from five bi-parental crosses were evaluated against two check varieties CoP16437 and CoP09437 in an augmented design with three blocks. Nineteen characters were recorded, juice quality being assessed independently at eight, ten and twelve months after planting. Twenty-two of the clones together with both checks were profiled using seven EST-SSR markers that resolved twenty-six allele bands. Treatment differences were significant ($P \leq 0.05$) for all nineteen characters. Substantial variability (GCV 11.16-17.28%) was evident for cane yield, CCS yield, cane diameter, plant height, single cane weight, germination and Brix and sucrose content at all three harvest ages, whereas juice purity showed the lowest variability of any character (GCV 1.37-1.57%). Broad-sense heritability was uniformly high, interpreted as an upper-bound estimate because the augmented design derives error variance from the replicated checks alone. Correlation analysis resolved the characters into two largely independent blocks: cane yield was most strongly associated with the number of millable canes per hectare ($r = 0.86$), single cane weight ($r = 0.79$), germination and plant height ($r = 0.77$ each) while quality characters were mutually correlated across harvest ages ($r = 0.82-0.87$) but showed no significant association with cane yield. Path coefficient analysis identified the number of millable canes per hectare as the principal direct determinant of cane yield (direct effect = 0.640; $R^2 = 0.788$). Ward's clustering grouped the forty-seven treatments into four clusters, one combining high biomass with low recovery and another the reverse. All twenty-six marker bands were polymorphic (mean PIC = 0.570; mean gene diversity = 0.624), with no two genotypes sharing an identical banding profile. UPGMA clustering of Jaccard distances resolved the panel into five groups (cophenetic correlation = 0.794). Four clones - CoXK25034, CoXK25036, CoXK25079 and CoXK25082 - combined a significant gain in CCS yield with recovery at par with the better check and fell into four different molecular clusters, indicating that superior performance arose independently in more than one lineage; these are recommended for advancement to replicated multi-location testing.

KEYWORDS: Sugarcane, augmented design, genetic variability, heritability, correlation, path coefficient analysis, cluster analysis, selection index, EST-SSR markers, commercial cane sugar, clonal evaluation

INTRODUCTION

Sugarcane (*Saccharum* spp. hybrid complex) is the principal source of sucrose and a major bioenergy feedstock, supplying about eighty per cent of the sugar produced worldwide and underpinning the livelihood of millions of cultivators across the tropics and sub-tropics (Cantarella et al., 2012; Godshall & Legendre, 2003). India is the second largest producer of sugarcane after Brazil, and the crop occupies a central place in the agrarian economy of Bihar and the wider Gangetic plain (Anonymous, 2023). Modern sugarcane cultivars are interspecific hybrids between *Saccharum officinarum* and *S. spontaneum*, carrying between eighty and one hundred and twenty chromosomes of mixed origin within a single, highly heterozygous genome (D'Hont et al., 1994; Gianotto et al., 2011; Zhang et al., 2018). This genomic complexity generates wide phenotypic segregation among full-sib progenies and constitutes both the principal challenge and the principal opportunity confronting the sugarcane breeder.

The improvement of cane yield and sugar recovery, the two attributes of overriding economic importance, depends upon the presence of adequate genetic variability in the base population and upon the correct identification of the characters through which that variability can most effectively be exploited (Falconer & Mackay, 1996; Dabholkar, 1992). Estimation of the genotypic and phenotypic coefficients of variation, together with heritability and genetic advance, indicates the extent to which a character is amenable to improvement through selection (Johnson et al., 1955; Burton & DeVane, 1953). Because a simple correlation coefficient confounds the direct effect of a character upon yield with the indirect effects transmitted through other, correlated characters, path coefficient analysis (Wright, 1921; Dewey & Lu, 1959) is additionally required to resolve the causal architecture of yield and to identify the components on which selection should be practiced. Multivariate clustering of the genotypes on the full complement of characters further assists the breeder in choosing genetically divergent parents for hybridization (Ward, 1963). Because newly generated sugarcane clones are available in a very limited quantity of planting material in the early stages of a breeding program, conventional replicated designs are frequently impracticable, and the augmented design of Federer (1956) is widely used to evaluate large numbers of unreplicated test entries against a small number of replicated check varieties, the check plots providing the basis for the estimation of error variance and for the adjustment of test entry means for block effects.

Molecular markers provide a complementary, environment-independent means of characterizing genetic diversity. Simple sequence repeat (SSR) markers derived from expressed sequence tags (EST-SSRs) combine the high polymorphism and co-dominant, multiallelic behavior of microsatellites with a location in transcribed and therefore relatively conserved regions of the genome, conferring high cross-transferability (Varshney et al., 2005; Gupta et al., 2003; Ellis & Burke, 2007). Large collections of EST-SSRs have been mined from the sugarcane expressed-sequence database (Pinto et al., 2004; Cardle et al., 2000), and candidate gene-based EST-SSR sets associated with sugar-related traits have recently been validated in *Saccharum* (Divakar et al., 2023), alongside genomic SSR panels such as those of Bilal et al. (2015). Because *Saccharum* spp. is an aneuploid polyploid in which allele dosage cannot be inferred from band intensity, SSR data in sugarcane are customarily analyzed as dominant, multiallelic marker phenotypes rather than under Hardy-Weinberg assumptions, and genetic similarity is estimated through coefficients such as that of Jaccard (1908) followed by hierarchical clustering (UPGMA). Combining a rigorous phenotypic evaluation with molecular characterization allows the breeder not only to identify superior clones but also to establish whether their superiority is associated with genetic distinctness, information of direct relevance to parental choice in subsequent crossing. The present investigation was accordingly undertaken with the following objectives: (i) to assess the extent of genetic variability, heritability and genetic advance for yield and juice quality traits among sugarcane clones derived from five bi-parental crosses; (ii) to examine the association among cane yield, juice quality and related component characters through correlation and path coefficient analysis, and to resolve the material into genetically divergent groups through cluster analysis; and (iii) to characterize a representative subset of the clones at the molecular level using EST-SSR markers, and to validate the phenotypically superior clones against this independent evidence of genetic distinctness.

MATERIAL AND METHODS

Experimental site and material

The field experiment was conducted during the crop season February 2025 to March 2026 at the Kalyanpur Farm, Sugarcane Research Institute, Dr. Rajendra Prasad Central Agricultural University (RPCAU), Pusa, Samastipur, Bihar (25.95° N, 85.67° E; altitude 52 m), which falls within the North-West Alluvial Plain Zone (Zone I) of Bihar and experiences a sub-tropical, sub-humid climate. The soil of the experimental field was a calcareous alluvium of sandy loam texture, slightly alkaline in reaction (pH 7.93), low in available nitrogen and phosphorus and medium in available potassium. Laboratory work relating to DNA extraction, PCR amplification and gel electrophoresis was carried out at the Department of Genetics and Plant Breeding, Post Graduate College of Agriculture, RPCAU, Pusa.

The experimental material comprised forty-five sugarcane clones of the CoXK series, derived from five bi-parental crosses (Table 1), together with two check varieties, CoP16437 (better check) and CoP09437, making a total of forty-seven treatments.

Experimental design and field layout

Because planting material of the newly generated clones was limited, precluding replication of every test entry, the experiment was laid out in an augmented design (Federer, 1956). The forty-five test clones were distributed unreplicated over three blocks, fifteen clones per block, while both check varieties were replicated once in every block, permitting block effects to be estimated from the checks and removed from test-entry performance to yield adjusted means on which all subsequent analyses were based. Three-budded setts were planted end to end in furrows spaced 90 cm apart, each clone occupying two rows of 3 m length (plot size 2 × 0.9 × 3 m). All recommended agronomic practices were followed uniformly.

Characters recorded

Nineteen characters comprising growth attributes, yield components and juice quality parameters were recorded: germination percentage at 45 days after planting (DAP); number of shoots at 120 and 240 DAP (°000/ha); number of millable canes per hectare (°000/ha); cane diameter (cm); plant height (cm); single cane weight (kg); cane yield

(t/ha); and, independently at 8, 10 and 12 months after planting, Brix (%), juice purity (%) and sucrose/pol (%), together with commercial cane sugar - CCS (%) and CCS yield (t/ha) computed at the twelve-month stage. Juice quality was assessed at three harvest ages so that the stability of clone ranking for quality over the maturity period could be examined directly. Derived characters were computed using the standard expressions of Meade & Chen (1977): Purity (%) = (Sucrose % / Brix %) × 100; CCS (%) = [Sucrose % - (Brix % - Sucrose %) × 0.4] × 0.73; Cane yield (t/ha) = millable canes/ha × single cane weight; CCS yield (t/ha) = (cane yield × CCS %)/100.

Statistical analysis

All phenotypic analyses were performed on the adjusted means obtained from the augmented-design analysis of variance (Federer, 1956) in the R statistical environment (R Core Team, 2022). Genotypic, phenotypic and environmental coefficients of variation (GCV, PCV, ECV) were computed following Burton & DeVane (1953), and classified as low (< 10%), moderate (10-20%) and high (> 20%) after Sivasubramanian & Menon (1973). Broad-sense heritability (h^2_{bs}) and genetic advance as per cent of mean (GAM) were estimated after Johnson et al. (1955), GAM being classified as low (< 10%), medium (10-20%) and high (> 20%). Pearson's phenotypic correlation coefficients were computed among all nineteen characters, and path coefficient analysis (Wright, 1921; Dewey & Lu, 1959) was carried out with cane yield as the dependent variable and seven growth and yield components as causal variables, using a reduced model to avoid multicollinearity. Standardized adjusted means were subjected to agglomerative hierarchical clustering by Ward's minimum-variance method on Euclidean distances. Two complementary selection indices were computed across all nineteen characters: the ranked summation index of Mulamba & Mock (1978) and a standardized (Z-score) selection index. Performance of each clone was additionally expressed as percentage deviation from the better check CoP16437 and tested against the critical difference ($P = 0.05$) derived from the augmented-design analysis.

Molecular characterization

Of the forty-five field-evaluated clones, twenty-two were selected on the basis of field performance for molecular characterization; together with both check varieties these constituted a panel of twenty-four genotypes. Genomic DNA was isolated from young leaf tissue by the CTAB method (Srivastava & Gupta, 2008) and its quality assessed by agarose gel electrophoresis. Fifteen SSR primer pairs of the SEM, CEMB, NKS and SMC series (Table 2) were screened; the SEM series was drawn from the candidate gene-based EST-SSRs of Divakar et al. (2023) and the CEMB series from the markers of Bilal et al. (2015). Seven primer pairs - SEM407, SEM435, SEM425, SEM2, SEM112, CEMB-14A and CEMB-5A - produced clear, reproducible, polymorphic amplification and were retained for analysis. PCR was carried out in a 12 µl reaction (2× PCR master mix, 0.5 µM each primer, template DNA and molecular-grade water) with an initial denaturation at 94 °C for 2 min, 35 cycles of 94 °C/30 s, 55 °C/30 s and 72 °C/30 s, and a final extension at 72 °C for 2 min. Amplicons were resolved on 2% agarose gel, stained with ethidium bromide and scored as a binary matrix (band present = 1, absent = 0); amplification failures were recorded as missing data rather than allele absence.

Because *Saccharum* spp. is an aneuploid polyploid in which allele dosage cannot be inferred from band intensity, the bands amplified by each primer pair were treated as alleles of a single multiallelic locus and all estimates were obtained without invoking Hardy-Weinberg equilibrium. Polymorphism information content ($PIC = 1 - \sum p_i^2$) was estimated after Botstein et al. (1980) and markers classified as highly informative ($PIC > 0.50$), reasonably informative (0.25-0.50) or slightly informative (< 0.25). Nei's gene diversity ($H = 1 - \sum p_i^2$), the effective number of alleles ($N_e = 1/\sum p_i^2$) and Shannon's information index ($I = -\sum p_i \ln p_i$) were also computed, together with resolving power ($R_p = \sum I^b$, where $I^b = 1 - 2|0.5 - p|$) and marker index ($MI = PIC \times N_a$). Pairwise genetic distance among genotypes was estimated as the complement of Jaccard's (1908) similarity coefficient [$J = a/(a+b+c)$], which disregards shared band absences, and the resulting matrix was subjected to UPGMA clustering with node support assessed by 100 bootstrap replicates; the cophenetic correlation between the original and dendrogram-implied distances was computed as a measure of fit. The molecular grouping was compared with the ranking obtained from field evaluation to establish whether phenotypically superior clones were also genetically distinct from the check varieties and from one another.

RESULT AND DISCUSSION

Analysis of variance

Treatment differences were significant ($P \leq 0.05$) for every one of the nineteen characters studied, establishing the presence of adequate variability in the experimental material (Table 3). Mean squares due to test entries exceeded those due to checks for all nineteen characters, indicating that the range of variation among the newly generated clones was wider than that between the two established checks. Block effects were significant for twelve of the nineteen characters and highly significant ($P \leq 0.01$) for the number of shoots at 240 DAP, confirming that adjustment of test-entry means for block position was necessary.

Mean performance and descriptive statistics

Germination at 45 DAP averaged 31.95% (range 25.58-41.48%). The number of millable canes per hectare averaged 99.58 thousand (88.41-118.06 thousand) and cane yield averaged 96.09 t/ha (73.97-127.69 t/ha); CoXK25036 recorded the highest values for shoots at 240 DAP, millable cane number, cane yield and CCS yield ($12.19 \pm$ range 7.80-16.34 t/ha), while CoXK25031 recorded the highest cane diameter and single cane weight. Juice quality characters increased progressively with maturity: Brix rose from 17.26% at 8 months to 20.60% at 12 months, sucrose from 14.81 to 18.33%, and purity from 85.50 to 88.69%, the range of purity being

conspicuously narrow at every harvest age (83.12-90.30%). Commercial cane sugar averaged 12.69% (9.25-15.34%). The clones occupying extreme positions for growth and yield attributes were distinct from those occupying extreme positions for juice quality, no clone appearing among the three highest for both classes of character (Table 4).

Genetic variability parameters

The phenotypic coefficient of variation exceeded the genotypic coefficient for every character, the margin being consistently narrow, and the environmental coefficient of variation was small throughout (0.20-2.38%). No character recorded a high GCV (> 20%). Moderate GCV was recorded for thirteen characters, the highest being CCS yield (17.28%), followed by cane yield (12.89%), cane diameter (12.65%), sucrose at 10 months (12.63%) and commercial cane sugar (12.35%); low GCV (< 2%) was confined to the three purity determinations and the three counts of shoot/cane number. Broad-sense heritability was high for all characters, a feature of the augmented design in which error variance is estimated from the replicated checks alone; these values are accordingly interpreted as upper-bound estimates. Genetic advance as per cent of mean ranged from 2.79% (purity at 12 months) to 35.31% (CCS yield). High heritability together with high GAM, indicating additive gene action and the likely effectiveness of simple selection, characterized CCS yield, cane yield, cane diameter, plant height, single cane weight, germination, commercial cane sugar and Brix/sucrose at all harvest ages, whereas the three purity determinations combined high heritability with low GAM, indicating that direct selection for juice purity would be unrewarding (Table 5).

Performance relative to the better check and selection indices

Performance of each clone was expressed as percentage deviation from the better check CoP16437 and tested against the critical difference ($P = 0.05$): $\pm 10.2\%$ for cane yield, $\pm 9.7\%$ for commercial cane sugar and $\pm 13.5\%$ for CCS yield. Seven clones were significantly superior to the check for CCS yield, five for cane yield and only four for commercial cane sugar; no clone was significantly superior for cane yield and commercial cane sugar simultaneously (Table 6). Of the seven clones superior for CCS yield, four - CoXK25034, CoXK25036, CoXK25079 and CoXK25082 - achieved this gain while remaining statistically at par with the check for recovery, whereas three (CoXK25017, CoXK25010, CoXK25046) gained CCS yield at a significant cost in recovery. A reciprocal pattern was observed among clones superior for recovery (CoXK25020, CoXK25035, CoXK25072, CoXK25059), each ranking among the lowest for CCS yield.

The ranked summation index (Mulamba & Mock, 1978) and a standardized (Z-score) selection index, computed across all nineteen characters, ranked CoXK25034 first on both criteria (RSI = 140, SI = 21.661), followed by CoXK25036 (SI = 18.438) and CoXK25082 (SI = 17.514). The better check CoP16437 occupied only the thirteenth position, indicating that twelve test clones surpassed it on aggregate performance (Table 7). The four clones combining a significant CCS-yield gain with recovery at par with the check occupied the first, second, third and eleventh positions on the selection index, providing independent concordance between the check-comparison and index-based approaches.

Correlation studies

Pearson's correlation coefficients among the nineteen characters (Table 8; Figure 1) resolved clearly into two blocks that were essentially independent of one another. Within the growth and yield block, cane yield was most strongly associated with the number of millable canes per hectare ($r = 0.86^{**}$), the number of shoots at 240 DAP ($r = 0.82^{**}$), single cane weight ($r = 0.79^{**}$), germination and plant height ($r = 0.77^{**}$ each), and cane diameter ($r = 0.74^{**}$). Within the quality block, Brix, purity and sucrose content were mutually and highly significantly correlated at every harvest age, and the correlations across ages were of comparable magnitude to those within a single age (e.g., Brix at 12 months with Brix at 8 and 10 months, $r = 0.87^{**}$ each), indicating that clone ranking for juice quality was established by eight months and maintained thereafter. Commercial cane sugar was most strongly determined by Brix at 12 months ($r = 0.91^{**}$). Between the two blocks, correlations were consistently negative and mostly non-significant: cane yield showed no significant association with any quality character ($r = -0.00$ to -0.27), while CCS yield was significantly but negatively correlated with commercial cane sugar ($r = -0.36^{*}$) and several other quality characters.

Path coefficient analysis

With cane yield as the dependent variable and seven growth and yield components as causal variables, the number of millable canes per hectare exerted by far the largest positive direct effect (0.640), more than twice that of the next character, the number of shoots at 240 DAP (0.292); germination (0.157) and shoots at 120 DAP (0.137) followed. Plant height (-0.156), cane diameter (-0.095) and single cane weight (-0.047) recorded negative direct effects despite strong positive total correlations with cane yield (0.775, 0.735 and 0.791 respectively), their positive association arising almost entirely from large positive indirect effects channelled through the number of millable canes (0.512-0.563). The seven characters together explained 78.80% of the variation in cane yield (residual effect = 0.4604) (Table 9).

Cluster analysis (phenotypic)

Ward's minimum-variance clustering of standardized adjusted means resolved the forty-seven treatments into four clusters (Table 10). Cluster II ($n = 8$) recorded the highest means for every growth and yield attribute (germination 36.85%, millable canes 109.52 '000/ha, cane yield 112.26 t/ha, CCS yield 15.18 t/ha) but the second-lowest commercial cane sugar (11.55%), assembling high-biomass, low-recovery material including CoXK25010, CoXK25017, CoXK25036 and CoXK25039. Cluster IV ($n = 18$), which contained the better check CoP16437,

recorded the highest means for every juice-quality character at every harvest age (commercial cane sugar 14.05%) while its yield attributes remained close to the overall mean, and contained the high-recovery clones CoXK25020, CoXK25035, CoXK25059 and CoXK25072 alongside CoXK25034 and CoXK25082. Cluster III ($n = 13$), containing check CoP09437, recorded the lowest means for every yield attribute; Cluster I ($n = 8$) occupied an intermediate yield position but the lowest quality means (Table 11). That the two checks fell into different clusters indicates their genetic distinctness and that the test material spans the range they define.

Molecular characterization

1. Amplification and polymorphism of EST-SSR markers

Of the fifteen SSR primer pairs screened, seven produced clear, reproducible and polymorphic amplification across the panel of twenty-two clones and two checks and were retained for analysis (Table 2). A total of twenty-six scorable allele bands were resolved, three to five bands per marker (mean 3.71); all twenty-six bands were polymorphic (100% polymorphism), and only one band occurred at a frequency below 5%. Bands per genotype ranged from five to nine (mean 6.8), and no two genotypes shared an identical banding profile. Amplification was complete for SEM407, SEM2, SEM112 and CEMB-5A; SEM435 and SEM425 each failed in a single genotype; CEMB-14A showed the poorest amplification, failing in seven of twenty-four genotypes. In total, 43 of 624 data points (6.9%) were missing and were treated as such rather than as allele absence.

2. Marker diversity parameters

Polymorphism information content (PIC) ranged from 0.432 (SEM112) to 0.725 (SEM435), mean 0.570; five of the seven markers exceeded the 0.50 threshold and were classed as highly informative (SEM435, CEMB-14A, SEM2, CEMB-5A, SEM425), the remaining two (SEM407, SEM112) being reasonably informative (Table 12; Figure 2). Nei's gene diversity paralleled PIC closely (range 0.483-0.762, mean 0.624). The effective number of alleles ranged from 1.935 to 4.195 (mean 2.884) against a mean observed allele number of 3.71 and Shannon's information index ranged from 0.839 to 1.520 (mean 1.132). SEM435 and CEMB-14A ranked highest on every parameter examined.

3. Genetic similarity and UPGMA cluster analysis

Pairwise genetic distance among the twenty-four genotypes, computed as the complement of Jaccard's (1908) similarity coefficient, was subjected to UPGMA clustering (cophenetic correlation $r = 0.794$; node support from 100 bootstrap replicates). Fusion distances ranged from 0.12 to 0.89. At a distance of 0.65 the panel resolved into five groups, two genotypes (CoXK25023, CoXK25029) remaining unattached at that level (Table 13; Figure 3). The closest pair, CoXK25039 and CoXK25044 (fusion distance 0.12, bootstrap 63%), derives from the same cross (Co 8371 $\times$ Co 88216). The two check varieties fused at 0.33 (bootstrap 81%) and were joined successively by CoXK25083 (0.47) and CoXK25082 (0.61); the node separating this group of four (Group I) from the remaining twenty genotypes, the deepest division in the dendrogram, was supported in all 100 bootstrap replicates.

4. Molecular validation of superior clones

The four clones combining a significant CCS-yield gain with recovery at par with the better check - CoXK25034, CoXK25036, CoXK25079 and CoXK25082 - were distributed among four of the five UPGMA groups (Groups V, III, IV and I respectively) rather than confined to one, indicating that superior performance arose independently in more than one lineage. CoXK25034, which ranked first on both selection indices, fell within Group V on the opposite side of the deepest dendrogram division from both checks, representing the most molecularly divergent of the four superior clones; CoXK25082 joined the checks within Group I, representing the same broad genetic type with improved performance. The molecular grouping did not correspond closely with the phenotypic classification of section 4.4/4.7 - clones of markedly different phenotypic behavior were frequently placed within the same UPGMA group - which is expected given that the two data sets sample different components of a genome of great size and complexity. No two genotypes in the panel shared an identical banding profile, confirming that the selected clones are not narrowly related duplicates of one another or of the check varieties.

DISCUSSION

Genetic variability, heritability and genetic advance

Significant treatment differences for all nineteen characters confirm that the material assembled from five biparental crosses carried substantial genetic variation, in agreement with the very highly significant variation reported by Tolera et al. (2023) for eighteen agro-morphological and biochemical characters among 196 Ethiopian genotypes, and with the findings of Tena et al. (2018), Tadesse et al. (2014), Alam et al. (2017) and Anna Durai et al. (2015) in other sugarcane germplasm. Cane yield, cane diameter, plant height, single cane weight and germination recorded moderate GCV (11.83-12.89%), comparable with the moderate-to-high coefficients reported by Rakesh et al. (2023), Gowda et al. (2016), Anbanandan & Eswaran (2018) and Tesfaye et al. (2020), whereas the counts of shoots and millable canes recorded the lowest GCV among the yield attributes, consistent with their constitution as area-based counts constrained at both extremes rather than a genuine absence of genetic difference.

Juice purity recorded the narrowest variability of any character at every harvest age (GCV 1.37-1.57%), a finding that recurs with striking consistency across widely separated environments: Tolera et al. (2023) obtained the lowest coefficients of their entire study (1.31 and 2.66%) for purity, Rakesh et al. (2023) recorded closely

comparable values (1.43-2.26%) at the twelve-month stage, and Karthik et al. (2025), working at the same Kalyanpur farm, reported distinctly lower variability for Brix, purity and sucrose than for the growth attributes; Mali & Patel (2013), Behou & Pene (2019), Japheth et al. (2019), Belwal & Ahmad (2020) and Kumar et al. (2018) reported similarly low coefficients. Because purity is a ratio of sucrose to total soluble solids, it tends to remain stable across genotypes that differ widely in absolute sugar content, and because successive cycles of selection for sugar accumulation have already narrowed the range of non-sucrose solids surviving in advanced breeding material, purity offers little further scope for improvement by direct selection; the breeder seeking to raise recovery must instead operate on Brix and sucrose content, both of which retained moderate variability in the present study.

Broad-sense heritability was uniformly high (96.94-98.51%), substantially exceeding values obtained in replicated trials such as those of Tolera et al. (2023, 2024) and Rakesh et al. (2023). This inflation is a direct consequence of the augmented design, in which the error variance is estimated only from the small, genetically uniform block of replicated checks; heritability computed on this basis approaches unity irrespective of the true transmissibility of the character and is accordingly interpreted here as an upper-bound estimate rather than a direct measure. Genetic advance as per cent of mean, being less sensitive to this inflation, provided the more informative criterion and ranged over a twelve-fold spread (2.79-35.31%). Following Johnson et al. (1955), high heritability accompanied by high GAM - obtained for CCS yield, cane yield, cane diameter, plant height, single cane weight, germination, commercial cane sugar and Brix/sucrose at all harvest ages - indicates the operation of predominantly additive gene action and predicts an effective response to simple phenotypic selection; the reverse combination in the three purity determinations, also reported by Tolera et al. (2023), Belwal & Ahmad (2020), Japheth et al. (2019) and Kumar et al. (2018), confirms that direct selection for purity would be unrewarding in this material.

Association among yield and quality characters

The correlation analysis resolved the nineteen characters into two blocks, cohesive within themselves and essentially independent of one another - a structure that governs the interpretation of the remaining analyses. Within the yield block, the strong association of cane yield with millable cane number ($r = 0.86$), single cane weight (0.79), germination and plant height (0.77 each) accords closely with Tena et al. (2016), Tolera et al. (2024), Vennela et al. (2024), Ahmed et al. (2019), Alam et al. (2014), Ebid et al. (2022) and Babu et al. (2009), and indicates that the yield components behave as a coherent system in which selection for one member tends to carry the others along. Within the quality block, the near-equal magnitude of correlations across harvest ages (e.g., Brix at 12 months with Brix at 8 and 10 months, $r = 0.87$ each) demonstrates that clone ranking for juice quality is established as early as eight months after planting and maintained thereafter, a pattern consistent with the multi-age approach of Coutinho et al. (2022).

The consistently negative and largely non-significant correlation between the two blocks - cane yield showing no significant association with any quality character and CCS yield significantly but negatively correlated with commercial cane sugar ($r = -0.36$) - closely parallels the findings of Divakar et al. (2023) on the same research station, who reported a strong negative association between commercial cane sugar and cane yield and comparatively weak correlations of sugar yield with the individual quality components. A plausible physiological basis lies in the competition between sustained vegetative growth and sucrose accumulation for available photosynthate: a clone that sustains vigorous growth over a longer period reaches harvest with a lower juice sucrose concentration than one that ceases elongation earlier and diverts assimilate to storage, consistent with the significant negative correlations of commercial cane sugar with germination and with shoot number at 240 DAP. Chandrakant et al. (2007) likewise cautioned that no single composite measure is adequate for early clonal screening in *Saccharum*; the present results, in which the correlation between cane yield and commercial cane sugar was not merely low but slightly negative, indicate that a composite index combining the two would in this material be positively misleading and that cane yield and commercial cane sugar should instead be treated as independent selection criteria.

Path coefficient analysis and selection criteria

The pre-eminence of millable cane number as the principal direct determinant of cane yield (direct effect = 0.640) is the most consistently reported finding in the sugarcane path-analysis literature, corroborated by Tena et al. (2016) (direct effect 0.812), Ahmed et al. (2019), Alam et al. (2014), Ebid et al. (2022), Barbosa et al. (2017), Esposito et al. (2012), Ferreira et al. (2007), Masri (2015) and Masri et al. (2022). The negative direct effects recorded by plant height, cane diameter and single cane weight, despite their strong positive total correlation with cane yield, resolve through their large positive indirect effects channelled via millable cane number (0.512-0.563), a pattern also reported by Tena et al. (2016) for stalk diameter and Brix. This is best interpreted as reflecting the compensation that operates between stalk number and stalk size within a fixed area: an increase in the number of millable canes is achieved at some cost to the dimensions of the individual cane, so that within any given cane population an increase in cane size can only have been achieved by a reduction in some other yield component. The practical conclusion is that the number of millable canes per hectare and the number of shoots at 240 DAP as a secondary criterion available earlier in the crop cycle, are the most reliable bases for selecting high-yielding clones, whereas the characters measured on the individual cane should be regarded as descriptive of cane type rather than employed as independent selection criteria. The residual effect of 0.4604 ($R^2 = 0.788$) indicates that unmeasured factors such as tillering uniformity, lodging incidence and localized soil variation contribute

appreciably to cane yield, consistent with the magnitude of residuals commonly reported in sugarcane, a crop that occupies the ground for twelve months and is thereby exposed to a wide range of unrecorded influences.

Genetic divergence and clustering

Ward's clustering reproduced at the level of whole genotypes, the same division that the correlation analysis had established at the level of characters: Cluster II assembled high-biomass, low-recovery material and Cluster IV the reciprocal, high-recovery material of near-average biomass, the two clusters occupying opposite extremes of the two largely independent axes. That the convergence of two independently conducted analyses - one at the level of characters, the other at the level of genotypes - upon the same underlying structure lends weight to the interpretation placed on each. The distribution of clones derived from a common cross across two or more clusters, and the converse, is consistent with expectation in a crop whose interspecific hybrid genome carries eighty to one hundred and twenty chromosomes of mixed origin (D'Hont et al., 1994; Zhang et al., 2018): the segregation of so large a number of homoeologous chromosomes generates wide phenotypic variation among full sibs, so that progenies of the same cross need not cluster together, a discordance also reported by Mebrahtom et al. (2016), Al-Sayed et al. (2012) and in an unrelated crop by Pecetti et al. (1992). Crosses between Clusters II and IV, the most divergent groups with respect to the characters on which each excelled, offer the greatest scope for recombining high biomass with high recovery, in line with the recommendation of Borém et al. (2016) that grouping on several characters simultaneously is preferable to sequential single-character selection.

Molecular characterization

Marker polymorphism, genetic relationships and breeding implications

The mean PIC of 0.570 and mean gene diversity of 0.624 obtained from seven EST-SSR markers are consistent with the highly heterozygous, interspecific hybrid constitution of *Saccharum* described by D'Hont et al. (1994) and Zhang et al. (2018), in which alleles derived from *S. officinarum* and *S. spontaneum* coexist within a single genome. The values obtained are somewhat lower than the 0.60-0.92 range reported by Chaudhary et al. (2024) using a mixed panel of genomic and EST-SSR markers, a difference attributable to the transcribed, relatively conserved location of EST-SSRs, which typically reveal less polymorphism than genomic SSRs while offering greater cross-transferability (Ellis & Burke, 2007; Varshney et al., 2005; Gupta et al., 2003) and to the selection of the SEM markers by Divakar et al. (2023) for functional association with sugar content rather than for maximum polymorphism. That no two genotypes shared an identical banding profile across all twenty-six bands confirms, in agreement with Glynn et al. (2009), that a small set of well-chosen EST-SSR markers suffices to fingerprint a clone set of this size.

UPGMA clustering, with a cophenetic correlation of 0.794 and 100% bootstrap support for the deepest division, correctly recovered the two check varieties as a pair (fusion distance 0.33, bootstrap 81%) and grouped full sibs of two of the five crosses together, confirming that the marker set retains genuine resolving power; the partial correspondence with pedigree, in which full sibs of other crosses were dispersed across separate groups, mirrors the incomplete correspondence between molecular grouping and phenotypic classification, and is expected given that twenty-six bands sample only a small, arbitrary portion of a genome of great size and complexity - a discordance also reported by Perera et al. (2012), Al-Sayed et al. (2012) and Mebrahtom et al. (2016). Consistent with the polygenic control of yield-related stalk traits emphasized by Aitken et al. (2008) and the caution of Wei et al. (2006) regarding spurious marker-trait associations arising from population substructure, the present marker panel is best understood as a tool for characterizing genetic distinctness rather than for inferring the genetic basis of performance. On this more limited but well-supported basis, the distribution of the four superior clones (CoXK25034, CoXK25036, CoXK25079, CoXK25082) across four separate molecular groups demonstrates that their superior field performance arose independently in more than one lineage, a result of direct consequence for the crossing program: CoXK25034, the most molecularly divergent of the four, offers the widest scope for generating novel recombinants as a parent, while CoXK25082 which clustered with the established check varieties, represents the more conventional route of incremental improvement within a known genetic background. Significant and substantial genetic variability, coupled with high heritability and genetic advance, was demonstrated for cane yield, CCS yield and the majority of growth, yield and juice-quality characters, indicating good prospects for improvement through direct phenotypic selection; juice purity by contrast offers little scope for direct selection. Cane yield and sugar recovery behaved as independent, weakly antagonistic attributes throughout the correlation, path-coefficient and cluster analyses and should accordingly be handled as separate selection criteria rather than combined into a single composite index. The number of millable canes per hectare was identified as the single most reliable criterion for selecting high-yielding clones. Molecular characterization with seven EST-SSR markers confirmed that the material carries substantial, informative polymorphism and that the four clones recommended for advancement - CoXK25034, CoXK25036, CoXK25079 and CoXK25082 - are genetically distinct from one another and from the check varieties, their superior performance having arisen independently in more than one lineage. These four clones, together with the divergent Clusters II and IV identified by phenotypic clustering, constitute the material of principal value for advancement to replicated multi-location testing and for use as parents in the next cycle of hybridization.

ACKNOWLEDGMENT

The authors are thankful to the Head, Department of Genetics and Plant Breeding and the Dean, Post Graduate College of Agriculture, Dr. Rajendra Prasad Central Agricultural University, Pusa, Samastipur, for providing the

facilities necessary to carry out this research and to the technical staff of the Sugarcane Research Institute and the Molecular Biology Laboratory for their assistance during the field and laboratory work.

Author Contribution

SKC and RK conceived and designed the study. SKC conducted the field experiment, the laboratory analysis and the statistical analysis, and prepared the original draft. RK and DNK supervised the investigation and provided laboratory facilities. RK and DNK also reviewed the manuscript and provided critical suggestions. All authors read and approved the final manuscript.

Statements & Declarations:

Funding

The authors did not receive support from any organization for the submitted work.

Data Availability

The datasets generated and analyzed during the present study are available from the corresponding author on reasonable request.

Ethical Approval and Consent to Participate

Not applicable.

Consent for Publication

Not applicable.

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

1. Ahmed KI, Patil SB, Moger NB, Hanumaratti NG, et al. (2019). Correlation and path analysis in sugarcane hybrid clones of proven cross. *J. Pharmacogn. Phytochem.*, 8(2):781-783.
2. Aitken KS, Hermann S, Karno K, Bonnett GD, et al. (2008). Genetic control of yield related stalk traits in sugarcane. *Theor. Appl. Genet.*, 117:1191-1203.
3. Alam MN, Nath UK, Karim KMR, Ahmed MM, et al. (2017). Genetic variability of exotic sugarcane genotypes. *Scientifica*, 2017:5202913.
4. Alam R, Rahman MA, Karim KMR, Tarique HM, et al. (2014). Genetic variability, correlation and path analysis of some yield components of sugarcane genotypes. *Bangladesh J. Sugarcane*, 35:28-36.
5. Al-Sayed HM, Fateh HS, Fares WM, Attaya AS (2012). Multivariate analysis of sugar yield factors in sugar cane. *Am.-Eurasian J. Sustain. Agric.*, 6(1):44-50.
6. Anbanandan V, Eswaran R (2018). Genetic variability, heritability and genetic advance in sugarcane. *Int. J. Recent Sci. Res.*, 9(2):24217-24219.
7. Anna Durai A, Premachandran MN, Govindaraj P, Malathi P, et al. (2015). Variability in breeding pool of sugarcane (*Saccharum* spp.) for yield, quality and resistance to different biotic and abiotic stress factors. *Sugar Tech*, 17:107-115.
8. Anonymous (2023). Sugarcane in India – a statistical profile: All India. ICAR–Sugarcane Breeding Institute, Coimbatore, India.
9. Babu C, Koodalingam K, Natarajan US, Shanthi RM, et al. (2009). Interrelationships of sugarcane yield and quality components and their utility in family selection. *Madras Agric. J.*, 96(7-12):309-313.
10. Barbosa RPB, Neto FA, Gravina LM, Gravina GA, et al. (2017). Early selection of sugarcane using path analysis. *Genet. Mol. Res.*, 16(1):gmr16019038.
11. Behou MY, Pene CB (2019). Genetic variability and heritability among sugarcane genotypes in plant crop for some agronomic traits under tropical dry climate of Ferke, Ivory Coast. *J. Exp. Agric. Int.*, 38(1):1-14.
12. Belwal V, Ahmad S (2020). Estimation of variability parameters in sugarcane under water logging conditions. *Int. J. Agric. Sci.*, 12:9451-9453.
13. Bilal M, Saeed M, Nasir IA, Tabassum B, et al. (2015). Association mapping of cane weight and tillers per plant in sugarcane. *Biotechnol. Biotechnol. Equip.*, 29(4):617-623.
14. Borém A, Guimarães MR, Daros E, Souza VQ (2016). Multi-trait selection for sugarcane breeding: a review of recent strategies. *Crop Sci.*, 56(4):1670-1684.
15. Botstein D, White RL, Skolnick M, Davis RW (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.*, 32(3):314-331.
16. Burton GW, DeVane EH (1953). Estimating heritability in tall fescue (*Festuca arundinacea*) from replicated clonal material. *Agron. J.*, 45(10):478-481.
17. Cantarella H, Buckeridge M, Van Sluys MA, Souza A, et al. (2012). Sugarcane. In: Vitousek P, Meinzer FC & Zhu YG (eds), *Comparative Ecosystem Physiology*. Elsevier.
18. Cardle L, Ramsay L, Milbourne D, Macaulay M, et al. (2000). Computational and experimental characterization of physically clustered simple sequence repeats in plants. *Genetics*, 156(2):847-854.
19. Chandrakant K, Kant R, Singh P (2007). Screening criteria for selection of superior clones in early clonal generation of *Saccharum* complex hybrids. *Crop Improv.*, 34(1):63-71.
20. Chaudhary D, Jeena AS, Pal N, Rohit R, et al. (2024). DNA fingerprinting and genetic diversity analysis in advanced clones of sugarcane (*Saccharum* sp. complex) using STMS markers. *Braz. Arch. Biol. Technol.*, 67:e24231022.

21. Coutinho AE, Silva MF, Perecin D, Carvalheiro R, et al. (2022). Association mapping for sugarcane quality traits at three harvest times. *Sugar Tech*, 24(2):448-462.
22. D'Hont A, Lu YH, Léon DGD, Grivet L, et al. (1994). A molecular approach to unraveling the genetics of sugarcane, a complex polyploid of the Andropogoneae tribe. *Genome*, 37(2):222-230.
23. Dabholkar AR (1992). Elements of Biometrical Genetics. Concept Publishing Company, New Delhi.
24. Dewey DR, Lu KH (1959). A correlation and path-coefficient analysis of components of crested wheatgrass seed production. *Agron. J.*, 51(9):515-518.
25. Divakar S, Jha RK, Kamat DN, Singh A (2023). Validation of candidate gene-based EST-SSR markers for sugar yield in sugarcane. *Front. Plant Sci.*, 14:1273740.
26. Ebid MHM, Khalil SRA, El-Gabry YA (2022). Evaluation the performance and estimating the association of the important traits for some sugarcane genotypes. *Egypt. J. Plant Breed.*, 26(1):105-125.
27. Ellis JR, Burke JM (2007). EST-SSRs as a resource for population genetic analyses. *Heredity*, 99(2):125-132.
28. Esposito DP, Peternelli LA, Paula TOM, Barbosa MHP (2012). Path analysis using phenotypic and genotypic values for yield components in the selection of sugarcane families. *Cienc. Rural*, 42(1):38-44.
29. Falconer DS, Mackay TFC (1996). Introduction to Quantitative Genetics, 4th edn. Benjamin Cummings, Harlow.
30. Federer WT (1956). Augmented designs. *Hawaiian Plant. Rec.*, 55:191-208.
31. Ferreira FM, Barros WS, Silva FL, Barbosa MHP, et al. (2007). Phenotypic and genotypic relationships among yield components in sugarcane. *Bragantia*, 66(4):605-610.
32. Gianotto AC, Abreu HMC, Arruda P, Filho JCB, et al. (2011). Sugarcane (*Saccharum* × *officinarum*): a reference study for the regulation of genetically modified cultivars in Brazil. *Trop. Plant Biol.*, 4(1):62-89.
33. Glynn NC, McCorkle K, Comstock JC (2009). Diversity among mainland USA sugarcane cultivars examined by SSR genotyping. *J. Am. Soc. Sugar Cane Technol.*, 29:36-52.
34. Godshall MA, Legendre BL (2003). Sugar/Sugarcane. In: Encyclopedia of Food Sciences and Nutrition, 2nd edn. Elsevier.
35. Gowda SNS, Saravanan K, Ravishankar CR (2016). Genetic variability, heritability and genetic advance in selected clones of sugarcane. *Plant Arch.*, 16:700-704.
36. Gupta PK, Rustgi S, Sharma S, Singh R, et al. (2003). Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. *Mol. Genet. Genomics*, 270:315-323.
37. Jaccard P (1908). Nouvelles recherches sur la distribution florale. *Bull. Soc. Vaudoise Sci. Nat.*, 44:223-270.
38. Japheth EJ, James O, Oliver K (2019). Estimates of genetic parameters and genotype by environment interactions for sugar yield and its components in sugarcane genotypes in Western Kenya. *J. Plant Breed. Crop Sci.*, 11(9):206-212.
39. Johnson HW, Robinson HF, Comstock RE (1955). Estimates of genetic and environmental variability in soybean. *Agron. J.*, 47(7):314-318.
40. Karthik K, Rakshitha S, Kamat DN (2025). Genetic evaluation of climate resilient sugarcane clones under waterlogged condition. *Electron. J. Plant Breed.*, 16(1):38-45.
41. Kumar P, Pandey SS, Kumar B, Kamat DN, et al. (2018). Genetic variability, heritability and genetic advance of quantitative traits in sugarcane. *Int. J. Chem. Stud.*, 6(3):3569-3572.
42. Mali SC, Patel AI (2013). Correlation and heritability studies in sugarcane. *AGRES Int. e-J.*, 2(4):466-471.
43. Masri MI (2015). Genetic trait interrelationships and selection indices for cane yield in sugar cane. *Egypt. J. Plant Breed.*, 19(4):1183-1197.
44. Masri MI, El-Taib ABA, Abu-Ellail FFB (2022). Genetic and phenotypic correlation and path coefficient analysis for traits in sugarcane. *SVU Int. J. Agric. Sci.*, 4(2):53-64.
45. Meade GP, Chen JCP (1977). Cane Sugar Handbook, 10th edn. John Wiley and Sons, New York.
46. Mebrahtom F, Firew M, Eyasu A (2016). Multivariate analysis of sugar yield contributing traits in sugarcane (*Saccharum officinarum* L.) in Ethiopia. *Afr. J. Plant Sci.*, 10(8):145-156.
47. Mulamba NN, Mock JJ (1978). Improvement of yield potential of the Eto Blanco maize (*Zea mays* L.) population by breeding for plant traits. *Egypt. J. Genet. Cytol.*, 7:40-51.
48. Pecetti L, Annicchiarico P, Damania AB (1992). Biodiversity in a germplasm collection of durum wheat. *Euphytica*, 60(3):229-238.
49. Perera MF, Arias ME, Costilla D, Luque AC, et al. (2012). Genetic diversity assessment and genotype identification in sugarcane based on DNA markers and morphological traits. *Euphytica*, 185(3):491-510.
50. Pinto LR, Oliveira KM, Ulian EC, Garcia AAF, et al. (2004). Survey in the sugarcane expressed sequence tag database (SUCEST) for simple sequence repeats. *Genome*, 47(5):795-804.
51. R Core Team (2022). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
52. Rakesh G, Eswara Reddy G, Dinesh A, Swapna N, et al. (2023). Genetic studies in advanced sugarcane mid-late clones through yield and quality traits. *Int. J. Environ. Clim. Change*, 13(10):4535-4542.
53. Sivasubramanian S, Menon M (1973). Heterosis and inbreeding depression in rice. *Madras Agric. J.*, 60(7):1139.
54. Srivastava S, Gupta PS (2008). Inter simple sequence repeat profile as a genetic marker system in sugarcane. *Sugar Tech*, 10:48-52.

55. Tadesse F, Negi T, Getaneh A, Dilnesaw Z, et al. (2014). Genetic variability and heritability of ten exotic sugarcane genotypes at Wonji Sugar Estate of Ethiopia. *Glob. Adv. Res. J. Phys. Appl. Sci.*, 3(4):1-4.
56. Tena E, Mekbib F, Ayana A (2016). Correlation and path coefficient analyses in sugarcane genotypes of Ethiopia. *Am. J. Plant Sci.*, 7(10):1490-1497.
57. Tena E, Mekbib F, Ayana A (2018). Sugarcane landraces of Ethiopia: germplasm collection and analysis of regional diversity and distribution. *Adv. Agric.*, 2018:7920724.
58. Tesfaye D, Tolera B, Tena E, Tadesse F, et al. (2020). Estimates of components of variances, heritability and genetic advance in sugarcane genotypes at Finchaa and Metahara sugar estates, Ethiopia. *Int. J. Curr. Res. Acad. Rev.*, 8(11):23-30.
59. Tolera B, Gedebo A, Tena E (2023). Variability, heritability and genetic advance in sugarcane (*Saccharum* spp. hybrid) genotypes. *Cogent Food Agric.*, 9(1):2194482.
60. Tolera B, Gedebo A, Tena E (2024). Genetic variability, character association and path analysis in sugarcane genotypes. *Arch. Agron. Soil Sci.*, 70(1):1-15.
61. Varshney RK, Graner A, Sorrells ME (2005). Genic microsatellite markers in plants: features and applications. *Trends Biotechnol.*, 23(1):48-55.
62. Vennela M, Kumar B, Varma DD (2024). Association study of yield and its component traits in sugarcane seedlings. *Electron. J. Plant Breed.*, 15(2):520-525.
63. Ward JH (1963). Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.*, 58(301):236-244.
64. Wei X, Jackson PA, McIntyre CL, Aitken KS, et al. (2006). Associations between DNA markers and resistance to diseases in sugarcane and effects of population substructure. *Theor. Appl. Genet.*, 114:155-164.
65. Wright S (1921). Correlation and causation. *J. Agric. Res.*, 20(3):557-585.
66. Zhang J, Zhang X, Tang H, Zhang Q, et al. (2018). Allele-defined genome of the autopolyploid sugarcane *Saccharum spontaneum* L. *Nat. Genet.*, 50(11):1565-1573.

Table 1: Parentage of the forty-five sugarcane clones evaluated in the present study

Cross	Parentage	No. of progenies	Representative clones
A	Co 98006 × Co 5510	7	CoXK25001, 25010, 25014, 25015, 25017, 25018, 25020
B	CoP 9301 × Co 13001	5	CoXK25022, 25023, 25025, 25026, 25027
C	Co 11001 × B 092	9	CoXK25029, 25030, 25031, 25032, 25034-25038
D	Co 8371 × Co 88216	6	CoXK25039, 25040, 25041, 25044, 25046, 25049
E	BO 155 × Co 11015	18	CoXK25050, 25054, 25059, 25062, 25064, 25069, 25072-25086

Checks: CoP16437 (better check) and CoP09437 (replicated in every block).

Table 2: SSR primer pairs screened in the present study

Sl. No.	Primer	Forward primer (5'→3')	Reverse primer (5'→3')	Ta (°C)	Retained
1	SEM407	GCGAAACTAGCGCTGCTAAA	GGAGGTTTCGGGTACGAGTC	55	Yes
2	SMC1039GC	AGGTGAGAGTTCCTGGCTTTCCA	TGTGCTGGCAAGCCCCTACTT	57	No
3	NKS33	ACAGGAGCGCTTGGAGATTA	GAGCAGAAGGGCTAGAAGCA	55	No
4	NKS42	ACCGATTGTTTCAGTGGGAAG	AACCTAGCAATTTACAAGAGAATTAGA	53	No
5	NKS40	GATGGAGGCTTTGCAATGAT	GCATGTCCCACTGAACTGAA	51	No
6	SEM435	AGGCTGAGAGAGCAAAGAAAAGA	CCTAGGATCCTTCGGGTTTC	51.2	Yes
7	SEM428	CAGGAAGAAACAGTAGGAAGCA	AGGTACTTGCGGTCTTGAT	49.6	No
8	SEM112	CCTTCTGCAGACGAGTTGAA	ACCTGACCAGCAAATCAACA	53.8	Yes
9	CEMB-14A	CCGTCGAGCTGGACTTCACG	CACCTCCACGTCACCGAC	47.6	Yes
10	SEM425	GTGCCACCAGCAGCAAT	TCTCGTAGCTGCTCGACTTC	56	Yes
11	CEMB-5A	GTTGAGGGTGAAGCGGATGG	AGCCTCTGCCACCACTCCTC	51.2	Yes
12	SEM2	ACACCGAGCTGTCCCAAT	GCATCTGATGAGCCTGTGAA	53.9	Yes
13	SEM436	GGTCCCATAACACAAGCA	TGCATGAAGAAGCTCAGGTG	55	No
14	SEM229	AGAACCACAACCACCAGGAG	ACAGTTGAATAGGCCGGATG	55	No
15	SEM369	CGCTTCCATATCTTCTTCTTG	TGACTCTCCGGTCCCTACAC	58	No

Ta = annealing temperature. Seven primer pairs (SEM407, SEM435, SEM425, SEM2, SEM112, CEMB-14A, CEMB-5A) produced clear, reproducible, polymorphic amplification and were retained for diversity analysis.

Table 3: Analysis of variance (mean squares) for nineteen characters in forty-seven sugarcane treatments evaluated in an augmented design

Character	Block (df=2)	Treatment (df=46)	Check (df=1)	Test (df=44)	Error (df=2)
Germination (%)	10.6*	14.2*	15.4*	14.5*	0.245
Shoots 120 DAP (‘000/ha)	40.6*	54.4*	52.3*	55.7*	1.01
Shoots 240 DAP (‘000/ha)	324**	117*	113*	119*	2.17
Brix 8 month (%)	0.735	3.86*	3.84*	3.95*	0.0800
Purity 8 month (%)	0.0673	1.73*	1.13*	1.78*	0.0317
Sucrose 8 month (%)	0.475	3.39*	3.29*	3.46*	0.0648
Brix 10 month (%)	2.19*	4.96*	4.86*	5.08*	0.0800
Purity 10 month (%)	0.764*	1.83*	1.13*	1.89*	0.0317
Sucrose 10 month (%)	1.62*	4.36*	4.20*	4.46*	0.0841
Brix 12 month (%)	4.64*	5.24*	5.42*	5.36*	0.0800
Purity 12 month (%)	0.706	1.48*	0.882*	1.53*	0.0467
Sucrose 12 month (%)	1.80	5.09*	4.97*	5.21*	0.0968
Millable canes (‘000/ha)	24.7*	50.3*	50.7*	51.4*	0.934
Cane diameter (cm)	0.0752*	0.127*	0.123*	0.130*	0.00247
Plant height (cm)	29.0	599*	580*	613*	10.2
Single cane weight (kg)	0.00277	0.0139*	0.0138*	0.0142*	0.000242
Cane yield (t/ha)	122*	153*	155*	156*	2.86
Commercial cane sugar (%)	2.11*	2.45*	2.54*	2.50*	0.0450
CCS yield (t/ha)	1.93*	4.42*	4.27*	4.52*	0.0841

* Significant at $P \leq 0.05$; ** significant at $P \leq 0.01$. DAP = days after planting.

Table 4: Descriptive statistics for nineteen characters based on adjusted means of forty-five test clones and two check varieties

Character	Mean	SD	Min	Max	CV (%)
Germination (%)	31.95	3.74	25.58	41.48	11.70
Shoots 120 DAP (‘000/ha)	97.72	7.32	83.29	116.43	7.49
Shoots 240 DAP (‘000/ha)	143.15	10.72	123.15	169.51	7.49
Brix 8 month (%)	17.26	1.95	13.40	20.90	11.30
Purity 8 month (%)	85.50	1.31	83.12	87.42	1.53
Sucrose 8 month (%)	14.81	1.83	11.20	18.27	12.34
Brix 10 month (%)	19.02	2.21	14.80	23.70	11.63
Purity 10 month (%)	86.87	1.34	83.28	88.83	1.55
Sucrose 10 month (%)	16.57	2.07	12.49	20.84	12.51
Brix 12 month (%)	20.60	2.27	15.50	24.40	11.03
Purity 12 month (%)	88.69	1.21	86.00	90.30	1.37
Sucrose 12 month (%)	18.33	2.24	13.59	21.97	12.22
Millable canes (‘000/ha)	99.58	7.04	88.41	118.06	7.07
Cane diameter (cm)	2.82	0.353	2.05	3.52	12.54
Plant height (cm)	199.80	24.31	157.83	265.33	12.17

Character	Mean	SD	Min	Max	CV (%)
Single cane weight (kg)	0.971	0.117	0.76	1.21	12.04
Cane yield (t/ha)	96.09	12.27	73.97	127.69	12.77
Commercial cane sugar (%)	12.69	1.55	9.25	15.34	12.24
CCS yield (t/ha)	12.19	2.09	7.80	16.34	17.12

Table 5: Genetic variability parameters for nineteen characters in sugarcane

Character	GCV (%)	PCV (%)	GAM (%)	Class
Germination (%)	11.83	11.93	24.20	High
Shoots 120 DAP	7.57	7.64	15.47	Medium
Shoots 240 DAP	7.56	7.63	15.45	Medium
Brix 8 month (%)	11.39	11.51	23.27	High
Purity 8 month (%)	1.55	1.56	3.16	Low
Sucrose 8 month (%)	12.45	12.57	25.45	High
Brix 10 month (%)	11.75	11.85	24.05	High
Purity 10 month (%)	1.57	1.58	3.21	Low
Sucrose 10 month (%)	12.63	12.75	25.81	High
Brix 12 month (%)	11.16	11.24	22.84	High
Purity 12 month (%)	1.37	1.39	2.79	Low
Sucrose 12 month (%)	12.34	12.45	25.21	High
Millable canes per hectare	7.14	7.20	14.59	Medium
Cane diameter (cm)	12.65	12.78	25.85	High
Plant height (cm)	12.29	12.39	25.13	High
Single cane weight (kg)	12.16	12.27	24.87	High
Cane yield (t/ha)	12.89	13.01	26.35	High
Commercial cane sugar (%)	12.35	12.47	25.26	High
CCS yield (t/ha)	17.28	17.44	35.31	High

GCV/PCV = genotypic/phenotypic coefficient of variation; GAM = genetic advance as per cent of mean.

Table 6: Sugarcane clones showing significant superiority over the better check CoP16437 for cane yield, commercial cane sugar and/or CCS yield

Clone	CY (t/ha)	% over check	CCS (%)	% over check	CCSY (t/ha)	% over check
CoXK25036	127.69	26.44 (+)	12.33	-7.78 (ns)	16.34	25.40 (+)
CoXK25079	100.78	-0.21 (ns)	13.66	2.17 (ns)	16.24	24.64 (+)
CoXK25017	117.75	16.60 (+)	10.80	-19.22 (-)	15.66	20.18 (+)
CoXK25010	116.03	14.89 (+)	9.25	-30.82 (-)	15.55	19.34 (+)
CoXK25082	103.79	2.77 (ns)	13.78	3.07 (ns)	15.15	16.27 (+)
CoXK25046	98.19	-2.77 (ns)	11.64	-12.94 (-)	14.86	14.04 (+)
CoXK25034	119.24	18.07 (+)	14.26	6.66 (ns)	14.83	13.81 (+)
CoXK25039	114.80	13.67 (+)	10.30	-22.96 (-)	14.78	13.43 (ns)
CoXK25020	105.51	4.48 (ns)	15.34	14.73 (+)	12.67	-2.76 (ns)
CoXK25072	73.97	-26.76 (-)	15.29	14.36 (+)	10.28	-21.11 (-)
CoXK25035	73.97	-26.76 (-)	15.34	14.73 (+)	8.49	-34.84 (-)
CoXK25059	93.66	-7.26 (ns)	14.89	11.37 (+)	9.30	-28.63 (-)
CoP16437 (check)	100.99	—	13.37	—	13.03	—

CY = cane yield; CCS = commercial cane sugar; CCSY = CCS yield. (+) significantly superior, (–) significantly inferior, (ns) at par with check, all at $P \leq 0.05$. Critical difference: $\pm 10.2\%$ (CY), $\pm 9.7\%$ (CCS), $\pm 13.5\%$ (CCSY).

Table 7: Ranking of the fifteen highest-performing treatments by ranked summation index (RSI) and standardized selection index (SI)

Rank	Genotype	RSI	Selection index (SI)
1	CoXK25034	140	21.661
2	CoXK25036	270	18.438
3	CoXK25082	189	17.514
4	CoXK25031	239	17.325
5	CoXK25020	235	15.086
6	CoXK25032	267	13.694
7	CoXK25069	284	11.632
8	CoXK25018	293	11.570
9	CoXK25062	311	10.897
10	CoXK25078	315	8.143
11	CoXK25079	336	8.029
12	CoXK25017	367	7.848
13	CoP16437 (check)	312	7.681
14	CoXK25059	385	6.269
15	CoXK25074	353	6.207

Table 8: Pearson's correlation coefficients among selected characters based on adjusted means of forty-seven treatments (n = 47)

Character pair (with Cane Yield)	r	Character pair (with CCS Yield)	r
Millable canes/ha	0.86**	Cane diameter	0.86**
Shoots at 240 DAP	0.82**	Millable canes/ha	0.85**
Single cane weight	0.79**	Shoots at 240 DAP	0.84**
Germination	0.77**	Cane yield	0.83**
Plant height	0.77**	Germination	0.83**
Shoots at 120 DAP	0.76**	Plant height	0.83**
Cane diameter	0.74**	Single cane weight	0.83**
Brix 10 month	–0.27 (ns)	Commercial cane sugar	–0.36*
Commercial cane sugar	–0.25 (ns)	Sucrose 8 mo / Brix 10 mo	–0.35* each
Purity 10 month	–0.00 (ns)	Brix 8 mo / Purity 12 mo	–0.33* each

* Significant at $P \leq 0.05$ ($r \geq 0.288$); ** significant at $P \leq 0.01$ ($r \geq 0.372$); $n = 47$. The full 19×19 correlation matrix is available from the corresponding author on request.

Table 9: Direct and indirect effects of seven yield components on cane yield

Character	GERM	SH120	SH240	NMC	DIA	PH	SCW	r with CY
GERM	0.157	0.101	0.244	0.513	–0.073	–0.130	–0.038	0.774**
SH120	0.116	0.137	0.232	0.512	–0.074	–0.128	–0.038	0.758**
SH240	0.131	0.109	0.292	0.530	–0.080	–0.126	–0.039	0.818**
NMC	0.126	0.110	0.242	0.640	–0.077	–0.138	–0.041	0.862**
DIA	0.121	0.106	0.246	0.516	–0.095	–0.120	–0.039	0.735**
PH	0.130	0.112	0.235	0.563	–0.073	–0.156	–0.037	0.775**
SCW	0.129	0.112	0.242	0.557	–0.078	–0.124	–0.047	0.791**

GERM = germination; SH120/SH240 = shoots at 120/240 DAP; NMC = millable canes; DIA = cane diameter; PH = plant height; SCW = single cane weight. $R^2 = 0.7880$; residual effect = 0.4604. ** $P \leq 0.01$; $n = 47$.

Table 10: Composition of the four clusters obtained by Ward's minimum-variance method

Cluster	n	Treatments
I	8	CoXK25026, 25038, 25040, 25046, 25050, 25064, 25080, 25083
II	8	CoXK25010, 25017, 25029, 25031, 25036, 25039, 25049, 25079
III	13	CoXK25001, 25014, 25023, 25030, 25037, 25041, 25054, 25073, 25077, 25084, 25085, 25086, CoP09437 (check)
IV	18	CoXK25015, 25018, 25020, 25022, 25025, 25027, 25032, 25034, 25035, 25044, 25059, 25062, 25069, 25072, 25074, 25078, 25082, CoP16437 (check)

Table 11: Cluster means for selected yield and quality characters

Character	Cluster I	Cluster II	Cluster III	Cluster IV	Overall mean
Germination (%)	33.66	36.85	28.76	31.17	31.90
Millable canes ('000/ha)	101.04	109.52	93.70	98.77	99.58
Cane diameter (cm)	2.89	3.30	2.52	2.79	2.82
Plant height (cm)	203.50	235.62	180.92	195.89	199.81
Cane yield (t/ha)	97.10	112.26	87.68	94.53	96.09
Brix 12 mo (%)	17.88	19.38	20.22	22.62	20.60
Commercial cane sugar (%)	11.20	11.55	12.43	14.05	12.69
CCS yield (t/ha)	12.66	15.18	10.73	11.71	12.19

Table 12: Allele number and diversity parameters of seven EST-SSR markers

Marker	Na	Genotypes scored	PIC	Nei's H	Ne	Shannon's I	Rp	MI
SEM435	5	23/24	0.725	0.762	4.195	1.520	2.174	3.626
CEMB-14A	5	17/24	0.705	0.747	3.959	1.467	2.000	3.527
SEM2	4	24/24	0.630	0.685	3.173	1.248	2.083	2.518
CEMB-5A	3	24/24	0.527	0.602	2.510	0.998	1.917	1.581
SEM425	3	23/24	0.521	0.586	2.416	0.986	1.739	1.563
SEM407	3	24/24	0.449	0.500	2.000	0.868	1.333	1.347
SEM112	3	24/24	0.432	0.483	1.935	0.839	1.250	1.295
Mean	3.71	—	0.570	0.624	2.884	1.132	1.785	2.208

Na = number of allele bands; PIC = polymorphism information content (Botstein et al., 1980); Ne = effective number of alleles; Rp = resolving power; MI = marker index.

Table 13: UPGMA groups of twenty-four sugarcane genotypes based on Jaccard's distance computed from twenty-six EST-SSR allele bands

Group	n	Genotypes
I	4	CoXK25082, CoXK25083, CoP16437 (check), CoP09437 (check)
II	3	CoXK25010, CoXK25017, CoXK25020
III	4	CoXK25036, CoXK25046, CoXK25059, CoXK25062
IV	6	CoXK25001, CoXK25031, CoXK25037, CoXK25050, CoXK25054, CoXK25079
V	5	CoXK25027, CoXK25034, CoXK25039, CoXK25044, CoXK25049
Unattached	2	CoXK25023, CoXK25029

Cophenetic correlation $r = 0.794$; node support from 100 bootstrap replicates; deepest division (Group I vs. remainder) supported in 100/100 replicates.

* $P \leq 0.05$ ** $P \leq 0.01$

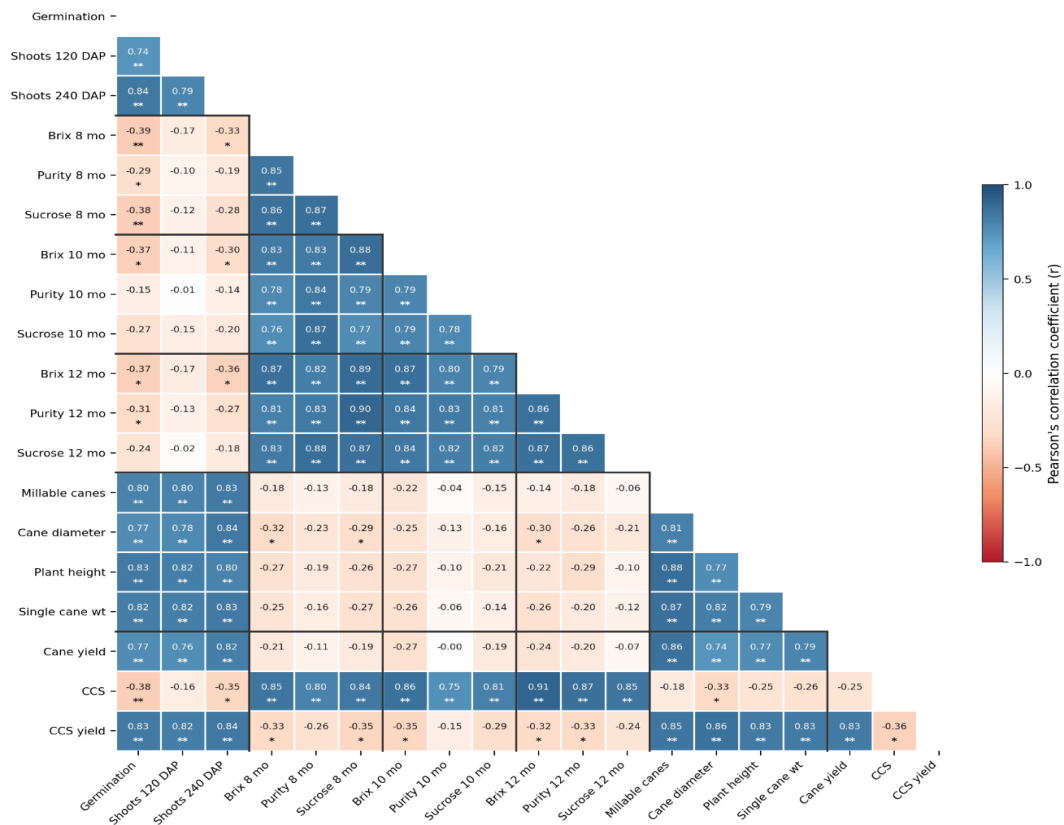
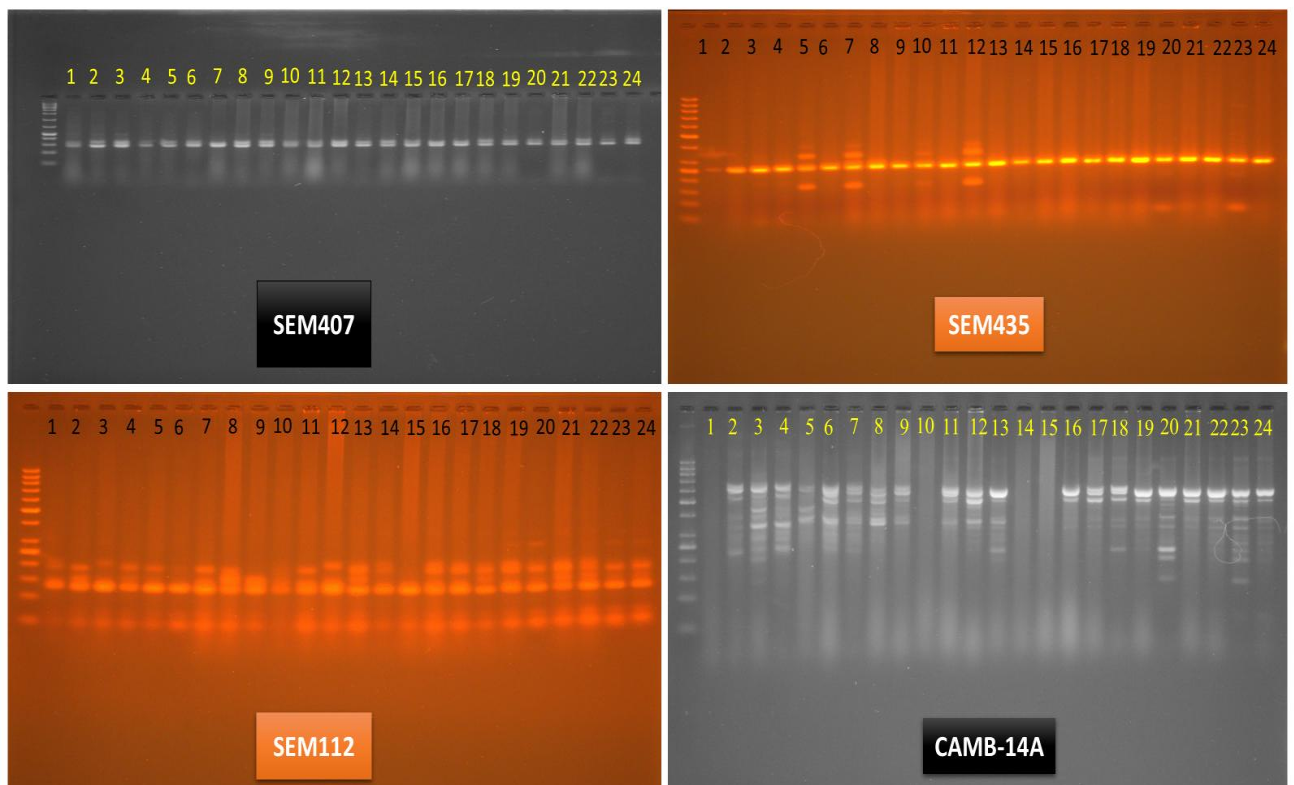


Figure 1: Heat map of Pearson's correlation coefficients among nineteen phenotypic characters, based on adjusted means of forty-seven sugarcane treatments. Cell values are correlation coefficients (r); color intensity denotes the sign and magnitude of association (blue = positive, red = negative), and * / ** denote significance at $P \leq 0.05$ / $P \leq 0.01$.

Agarose gel electrophoresis profiles of SSR markers



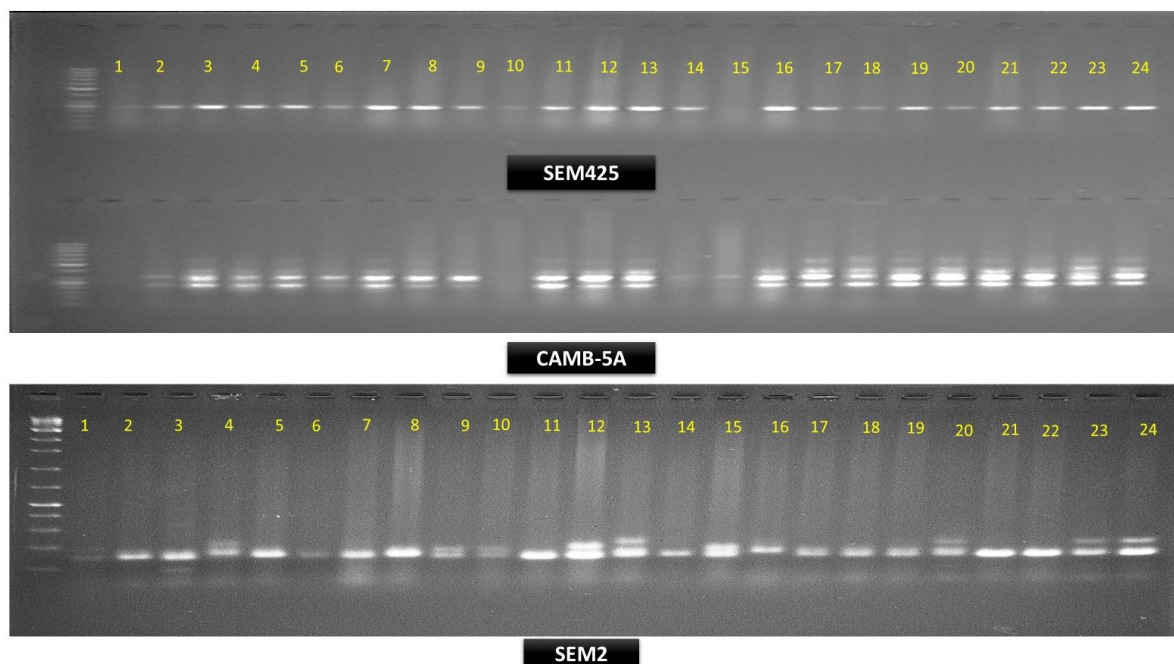


Figure 2: Agarose gel electrophoresis (gel documentation) profiles of the seven polymorphic EST-SSR markers used to genotype the twenty-two sugarcane clones and two check varieties: (a) SEM407, (b) SEM435, (c) SEM112, (d) CEMB-14A, (e) SEM425 / CEMB-5A (two panels from the same gel), (f) SEM2. Lane order follows Table 2; M = 50 bp DNA ladder.

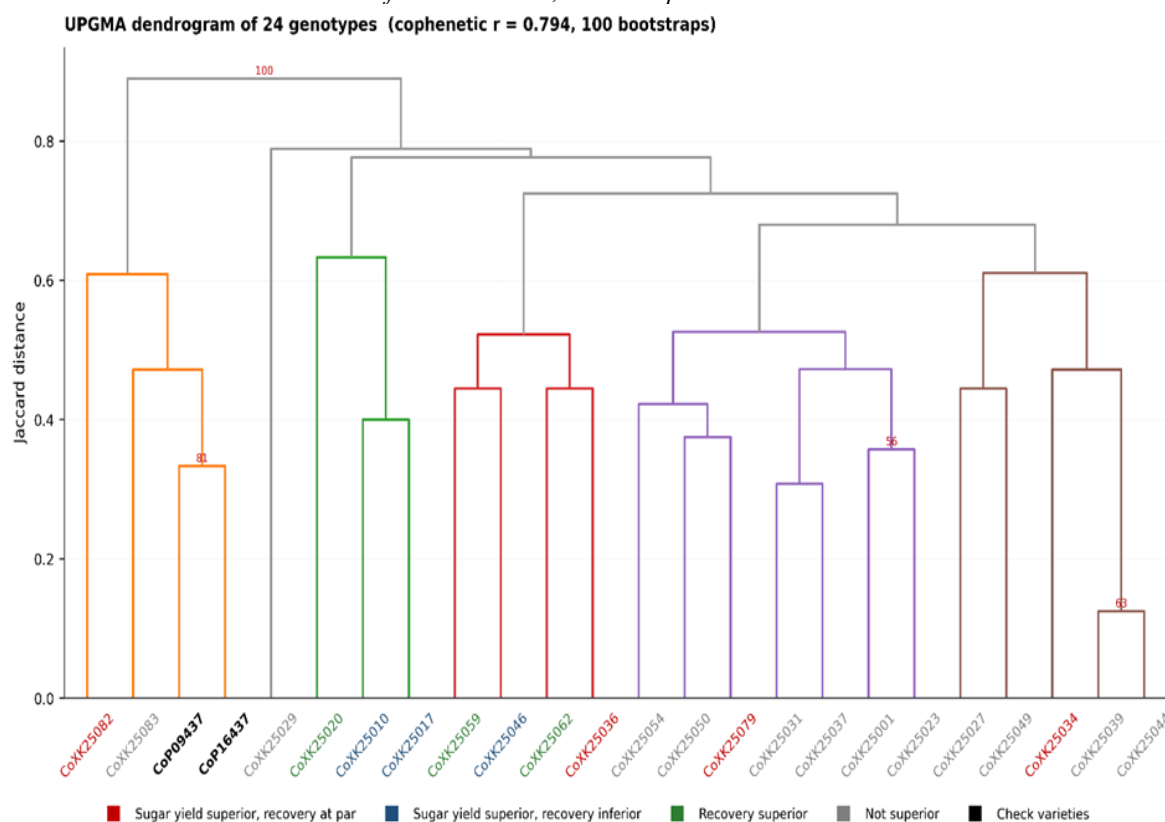


Figure 3: UPGMA dendrogram of the twenty-four sugarcane genotypes based on Jaccard's genetic distance computed from twenty-six EST-SSR allele bands (cophenetic correlation $r = 0.794$; bootstrap support from 100 replicates shown at key nodes). Branch colours denote the five UPGMA groups of Table 13 (orange = Group I, green = Group II, red = Group III, purple = Group IV, brown = Group V); grey branches mark CoXK25023 and CoXK25029, which remained unattached to any group at the 0.65 fusion-distance cut. Genotype-label colours indicate yield/recovery performance class (red = sugar-yield superior with recovery at par, blue = sugar-yield superior with recovery inferior, green = recovery superior, grey = not superior, black = check varieties).