



Revista Industrial
y Agrícola de
Tucumán

ISSN 0370-5404

En línea
1851-3018

Tomo 104 (1):
1-5; 2026



ESTACION EXPERIMENTAL
AGROINDUSTRIAL
OBISPO COLOMBRES
Tucumán | Argentina

Av. William Cross 3150
T4101XAC - Las Talitas.
Tucumán, Argentina.

Trabajo ya publicado

**Trabajo
presentado en
el XXXII ISSCT
Centennial
Congress, 24 al 28
de agosto de 2025,
Cali, Colombia,
traducido al
castellano.**

Fecha de
recepción:
07/10/2026

Fecha de
aceptación:
09/10/2026

Diversidad genética en el banco de germoplasma de la EEAOC y evaluación de una colección núcleo

Perera, María Francisca; Peña Malavera, Andrea Natalia; Noguera, Aldo Sergio; Racedo, Josefina y Ostengo, Santiago

Instituto de Tecnología Agroindustrial del Noroeste Argentino (ITANOA), Estación Experimental Agroindustrial Obispo Colombres (EEAOC) - Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), CCT NOA Sur. Av. William Cross 3150. Las Talitas, Tucumán, R. Argentina, T4101XAC. Email: franciscaperera@eeaoc.org.ar

RESUMEN

El conocimiento de la diversidad genética y la estructura poblacional de los bancos de germoplasma de caña de azúcar proporciona información valiosa para producir progenies con la máxima variabilidad. Para ello, se genotiparon 350 accesiones del banco de germoplasma de la EEAOC con 74.969 loci de SNP bialélicos mediante la tecnología DArT-seq. La diversidad genética y la estructura genética de toda la colección se analizaron mediante un método bayesiano y análisis de coordenadas principales (ACoP), mientras que la variabilidad se estimó mediante análisis de varianza molecular (AMOVA). También, se sugirieron colecciones núcleo de diferentes tamaños. La heterocigosidad media observada (HO) de todos los SNP fue de 0,28 (0 a 1), mientras que su heterocigosidad esperada (HE) varió de 0 a 0,5, con una media de 0,3. El contenido de información polimórfica (PIC) de los SNP varió de 0 a 0,38, con un valor medio de 0,22 y una media del poder discriminante (D_j) de 0,28. El análisis de la estructura genética y el ACoP revelaron que las accesiones se agruparon en tres grupos (I: 71 genotipos; II: 243 genotipos y III: 34 genotipos). El AMOVA sugirió un mayor porcentaje de variación dentro de cada subpoblación (89,75%) que entre ellas (10,25%; $p < 0,0001$), mientras que el valor de Φ_{ST} del 10% indicó una diferenciación genética moderada entre estos tres grupos. Por otro lado, una colección núcleo de 35 genotipos (10%) capturó casi el 96% de la diversidad, mientras que al incluir 104 genotipos (30%) se capturó más del 98% de la diversidad. Los resultados permitirán la explotación eficiente de la diversidad genética del germoplasma de caña de azúcar y el desarrollo de estrategias para obtener nuevos cultivares con características superiores.

Palabras clave: marcadores DArT-seq; estructura poblacional; diversidad capturada.

ABSTRACT

Genetic diversity in EEAOC's germplasm bank and assessment of a core collection

Understanding the genetic diversity and population structure of sugarcane germplasm banks provides valuable information to produce progenies with maximum variability. To this end, 350 accessions of the EEAOC's germplasm bank were genotyped at 74,969 biallelic SNP loci using DArT-seq technology. Genotyping data were analyzed to assess the genetic diversity and the genetic structure of all the collection through a Bayesian method and a principal coordinate analysis (PCoA), and its variability through an analysis of molecular variance (AMOVA). Finally, core collections of different sizes were suggested. Across all the accessions, the mean observed heterozygosity (HO) of all SNPs was 0.28 (ranging from 0 to 1) while their expected heterozygosity (HE) ranged from 0 to 0.5, with a mean of 0.3. The polymorphism information content (PIC) of the SNPs varied from 0 to 0.38, with a mean value of 0.22 and a mean of their discrimination power (D_j) of 0.28. Genetic structure analysis and PCoA revealed that accessions clustered in three groups (I: 71 genotypes; II: 243 genotypes and III: 34 genotypes). AMOVA suggested a greater percentage of variation inside each subpopulation (89.75%) than among them (10.25%; $p < 0.0001$), whereas the Φ_{ST} value of 10% indicated a moderate genetic differentiation among these three groups. Different sizes of core collections were tested; a core collection

of 35 genotypes (10%) captured almost 96% of diversity, while including 104 genotypes (30%) more than 98% of diversity was captured. Results will allow the efficient exploitation of genetic diversity of sugarcane germplasm, and the development of strategies to obtain new cultivars with superior characteristics.

Key words: DArT-seq markers; population structure; captured diversity.

INTRODUCCIÓN

En los programas de mejoramiento genético, la variabilidad genética y su manipulación mediante métodos adecuados conducen a la generación de genotipos superiores con características agronómicas interesantes. En este sentido, los bancos de germoplasma desempeñan un papel esencial en la preservación de esta variabilidad genética, aunque su manejo puede ser complejo debido al gran número de accesiones (de Oliveira *et al.*, 2020). A fin de ampliar la base y el acervo genéticos, la redundancia de alelos debe minimizarse mientras que la introducción de nuevas accesiones debe ser optimizada.

El conocimiento de la diversidad genética y la estructura poblacional de los bancos de germoplasma, permite planificar y conservar genotipos para garantizar el empleo de recursos genéticos en los programas de mejoramiento a largo plazo (Kanaka *et al.*, 2023). En general los bancos contienen numerosas accesiones, sin embargo, solo una pequeña proporción de estos recursos se utiliza en la práctica. Por lo tanto, es crucial optimizar el uso de estos recursos agrupando suficientes accesiones en una colección núcleo que maximice la diversidad genética descrita en toda la colección.

En Argentina, existen tres programas de mejoramiento genético de caña de azúcar; el de la EEAOC produce variedades TUC, y su importancia para la provincia de Tucumán (responsable del 70% de la producción total) es tal que ha liberado nueve variedades desde 2009. Además, el último censo varietal reveló que más del 98% del área de caña de azúcar relevada en Tucumán está plantada con variedades liberadas por este programa de mejoramiento (Henríquez *et al.*, 2023). Previamente, la diversidad genética y la estructura poblacional de algunos genotipos del banco de germoplasma se caracterizaron mediante AFLP (Perera *et al.*, 2012) y marcadores TRAP y SSR (Perera *et al.*, 2023). Sin embargo, actualmente los polimorfismos en nucleótidos simples (SNP) son los marcadores empleados por excelencia para estudios de diversidad genética (Yirgu *et al.*, 2023). Considerando lo expuesto, el objetivo del trabajo fue cuantificar la diversidad genética y determinar la estructura genética del banco de germoplasma de la EEAOC mediante la caracterización molecular de 350 individuos con la tecnología DArT-seq y sugerir la composición de una colección núcleo.

MATERIALES Y MÉTODOS

Material vegetal y extracción de ADN

La población incluyó 350 genotipos de caña de azúcar de los 789 que conforman la colección de ger-

moplasma del programa de mejoramiento genético de la EEAOC (Ostengo *et al.*, 2021). Si bien la mayoría se utiliza frecuentemente como progenitores para el cruzamiento, otros se incluyeron para capturar la diversidad alélica.

El ADN se extrajo de tejido foliar molido con nitrógeno líquido siguiendo el protocolo de Diversity Arrays Technology (DArT) (www.diversityarrays.com/orderinstructions/plant-dna-extraction-protocol-for-dart/). La calidad y la cantidad del ADN se verificaron en geles de agarosa al 0,8%, teñidos con gelred.

DArT-seq

Todos los individuos fueron genotipados utilizando la plataforma DArT-seq (Yarralumla, Australia), una tecnología de alto rendimiento basada en un método de reducción de la complejidad genómica mediante la selección inteligente de fracciones genómicas que corresponden predominantemente a genes activos. Estas fracciones genómicas se secuenciaron posteriormente mediante plataformas NGS. Los datos obtenidos se filtraron, depuraron excluyendo SNP monomórficos y se clasificaron en tres clases: 0: homocigota de referencia; 1: homocigota para SNP; y 2: heterocigota, dado que no se infirió la dosis alélica.

Nivel de polimorfismo

Se calcularon estadísticas descriptivas del nivel de polimorfismo en todas las accesiones utilizando R: número de alelos efectivos (N_e), frecuencia de alelos nulos (N_A), heterocigosidad observada (H_O), heterocigosidad esperada (H_E), contenido de información polimórfica (PIC) y poder discriminante (D_j).

Análisis de estructura genética

Se determinó mediante un método bayesiano, donde se ejecutaron modelos para $K=2-10$, análisis de coordenadas principales (ACoP) y análisis de varianza molecular (AMOVA) con el software R.

Colecciones núcleo

Se generaron diferentes muestras modificando el tamaño de las colecciones núcleo de 0,1 (10%) a 0,3 (30%), para identificar el subconjunto de genotipos que pudiera capturar la mayor diversidad genética utilizando el paquete R corehunter 3.0.

RESULTADOS Y DISCUSIÓN

Nivel de polimorfismo

Una mejor comprensión de la diversidad genética y de la existencia de estructura poblacional en el ger-

plasma de caña de azúcar disponible para el mejoramiento constituye un primer paso necesario para determinar cómo caracterizar y organizar la variación presente en la colección. Esto permite seleccionar progenitores potenciales para los cruzamientos, orientados a una introgresión eficiente de alelos deseables con el fin de alcanzar los objetivos del programa de mejoramiento (Suman *et al.*, 2012).

A fin de estudiar la diversidad genética, se caracterizaron 350 genotipos del banco de germoplasma del EEAOC mediante DArT-seq. Se obtuvieron 74.969 SNP. En el primer análisis, se identificaron dos genotipos fuera de grupo (L 79-1002 y POJ 2878) y se eliminaron de la matriz original para realizar el análisis nuevamente (Tabla 1).

Según Kanaka *et al.* (2023), los valores de PIC comprendidos entre 0,25 y 0,50 son moderadamente informativos, mientras que los valores inferiores a 0,25 se consideran no informativos. En este estudio, 39.622 SNPs presentaron valores de PIC entre 0,25 y 0,38, por lo que constituyeron los marcadores de mayor utilidad para diferenciar los genotipos.

Tabla 1. Número de alelos efectivos (Ne), frecuencia de alelos nulos (NA), heterocigosidad observada (HO), heterocigosidad esperada (HE), contenido de información polimórfica (PIC) y poder discriminante (Dj) de los 74.969 SNP en los 348 genotipos de caña de azúcar.

Variable	Media	Desvío estándar	Mínimo	Máximo
Ne	1,54	0,40	1,00	2,00
NA	134,60	79,51	0,00	316
HO	0,28	0,33	0,00	1,00
HE	0,30	0,20	0,00	0,50
PIC	0,22	0,15	0,00	0,38
Dj	0,28	0,20	0,00	0,50

Análisis de estructura genética

En el análisis de la estructura genética con el método bayesiano se obtuvo un K óptimo = 3, los genotipos fueron asignados a uno de los tres grupos considerando su mayor probabilidad de pertenecer a cada uno. El grupo I consistió en 71 genotipos, el grupo II: 243 genotipos y el grupo III: 34 genotipos (Figura 1).

El grupo II agrupó la mayoría de los genotipos, principalmente aquellos de diferentes orígenes (CB, CO, CP, FAM, NA, NCO, R, SP, US y VESTA). Los genotipos

TUC, HO, HOCP, L y LCP se distribuyeron en los tres grupos. El programa de mejoramiento de la EEAOC ha mantenido un intercambio activo de variedades con programas de mejoramiento de caña de azúcar de Florida (USDA-ARS, Sugarcane Field Station, Canal Point, variedades CP) y de Luisiana (Louisiana State University Agricultural Center, variedades L; y USDA-ARS, Sugarcane Research Unit, Houma, variedades HO) (Ostengo *et al.*, 2021). Esto podría explicar el número de alelos compartidos entre los genotipos TUC y aquellos provenientes de Estados Unidos.

Posteriormente, se representó el patrón de diversidad de la colección a través de un ACoP, que reveló que los genotipos se agruparon en tres grupos al igual que en el análisis de estructura genética.

Considerando los tres grupos detectados, se realizó un AMOVA que sugirió un mayor porcentaje de variación dentro de cada subpoblación (89,75%) que entre ellas (10,25%; $p < 0,0001$). El índice Φ_{ST} (10%) indicó una diferenciación genética moderada entre los tres grupos.

Colecciones núcleo

Una colección núcleo, conformada por un número reducido de genotipos no redundantes y representativa de la máxima diversidad genética de todo el banco de germoplasma, facilita su utilización en programas de mejoramiento. De este modo, es posible gestionar los recursos de manera más eficiente sobre una población más pequeña pero genéticamente diversa, al tiempo que se maximiza la ganancia genética (Fickett *et al.*, 2020).

Para detectar una colección núcleo que capture la mayor parte de la diversidad genética en un número menor de genotipos, se evaluaron tres tamaños diferentes. Una colección de 35 genotipos (10%) capturó casi el 96% de la diversidad (Tabla 2), pero la mayoría de los genotipos en la colección núcleo más pequeña estuvieron incluidos en las más grandes (Figura 2).

Tabla 2. Diversidad capturada por diferentes tamaños de colecciones núcleo.

Tamaño	Número de genotipos	Número de alelos	Diversidad capturada
0,1	35	137.040	95,78%
0,2	70	139.013	97,16%
0,3	104	140.554	98,24%
1,0	348	143.063	100%

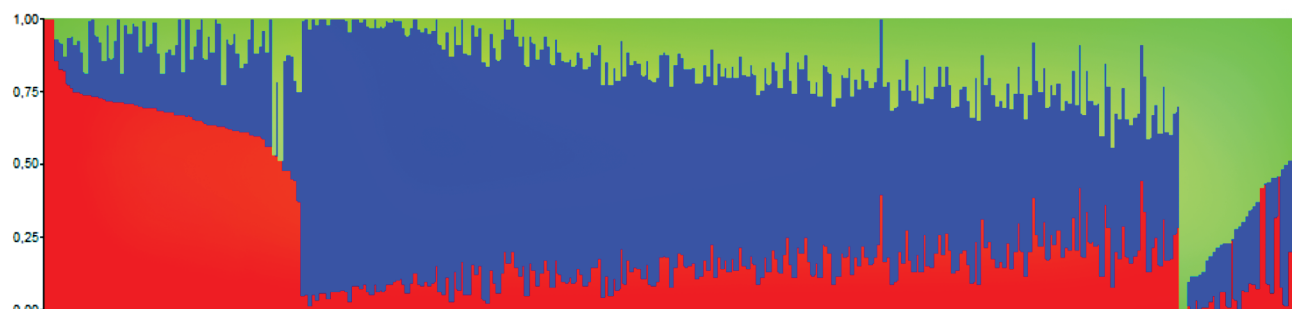


Figura 1. Estructura genética de 348 accesiones de caña de azúcar obtenida con 74.969 SNPs. Cada genotipo está representado por una sola línea vertical, dividida en colores según su pertenencia estimada en cada grupo. En rojo, grupo I; en azul, el grupo II; y grupo III en color verde.

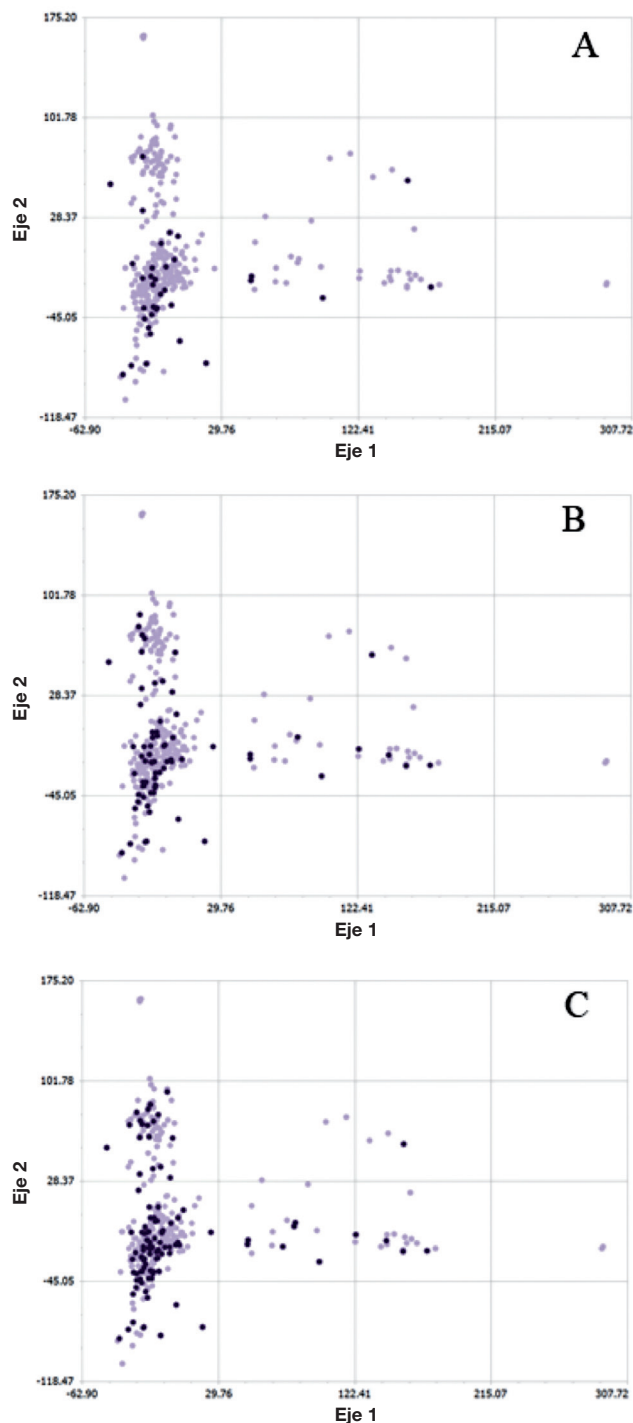


Figura 2. ACoP de 348 accesiones de caña de azúcar obtenido con 74.969 SNP. **A:** Colección núcleo con 35 genotipos (10%). **B:** Colección núcleo con 70 genotipos (20%) y **C:** Colección núcleo con 104 genotipos (30%). Los genotipos de las colecciones núcleo se indican en morado oscuro.

En un estudio previo, 47 genotipos del banco de germoplasma de la EEAOC, también incluidos en el presente estudio, se caracterizaron utilizando marcadores TRAP y SSR (Perera *et al.*, 2023). La mayoría de estos genotipos no quedaron incluidos en las colecciones núcleo sugeridas; solo un genotipo de los 47 se incluyó en la colección núcleo del 20% (TUC 96-52), 3 en la co-

lección núcleo 30% (TUC 00-23, TUC 01-2 y TUC 02-22), mientras que 4 se incluyeron en ambas colecciones (LCP 85-376, TUC 00-19, TUC 07-21 y TUC95-46). Estas colecciones núcleo serán de utilidad para que el programa de mejoramiento incluya nuevos progenitores en su esquema de cruzamientos, en ese sentido genotipos de diferentes orígenes como CB, CO, R y SUMATRA pueden aportar nuevos alelos, además individuos diferentes a los utilizados de los programas CP, FAM, HO, HOCP, L, NA y TUC podrían ser fuente de variabilidad. Sin embargo, debe considerarse que la posible asociación de alelos desfavorables con genes de interés podría dificultar la utilización eficiente en los programas de mejoramiento de los genotipos incluidos en las colecciones núcleo. Por ello, la caracterización fenotípica de estos individuos debe utilizarse para identificar aquellos genotipos de particular interés para determinados caracteres. Estos resultados permitirán realizar una introgresión más eficiente de alelos de interés (Fickett *et al.*, 2020).

CONCLUSIÓN

La implementación de herramientas moleculares permite un uso más racional de la diversidad genética de los bancos de germoplasma para diseñar cruzamientos que permitan la introgresión eficiente de alelos favorables, aumentando la probabilidad de obtener variedades más productivas y/o resistentes.

AGRADECIMIENTOS

Los autores agradecen al Dr. Ing. Agr. Eduardo Romero y al Ing. Agr. Sergio Casen, de la Sección Agronomía de Caña de Azúcar por su colaboración en el muestreo de RAC y sorgo fibroso de los diferentes campos cañeros y lotes de ensayos de la provincia de Tucumán.

BIBLIOGRAFÍA CITADA

- de Oliveira GL, de Souza AP, de Oliveira FA, Zucchi MI, de Souza LM, Moura MF, 2020. Genetic structure and molecular diversity of Brazilian grapevine germplasm: Management and use in breeding programs. PLoS ONE 15: e0240665,
- Fickett ND, Ebrahimi L, Parco AP, *et al.* 2020. An enriched sugarcane diversity panel for utilization in genetic improvement of sugarcane. Scientific Reports 10: 13390.
- Henriquez DD, Medina P, Aybar Guche M, *et al.*, 2023. Relevamiento de la distribución varietal y del uso de otras tecnologías en el cultivo de caña de azúcar en la provincia de Tucumán: campaña 2022/2023. Reporte Agroindustrial 282.
- Kanaka KK, Sukhija N, Goli RC, *et al.*, 2023. On the concepts and measures of diversity in the genomics era, Current Plant Biology 33: 100278.
- Ostengo S, Serino G, Perera MF, *et al.*, 2021. Sugarcane breeding, germplasm development and supporting genetic research in Argentina. Sugar Tech 24:166-180.

Perera MF, Arias ME, Costilla D, et al., 2012. Genetic diversity assessment and genotype identification in sugarcane based on DNA markers and morphological traits. *Euphytica* 185: 491-510.

Perera MF, Ostengo S, Peña Malavera AN, et al., 2023, Genetic diversity and population structure of *Saccharum* hybrids. *PLoS ONE* 18: e0289504.

Suman A, Ali K, Arro J, Parco AS, Kimbeng CA, Baisakh N. 2012. Molecular diversity among members of the

Saccharum complex assessed using TRAP markers based on lignin-related genes. *Bioenergy Research* 5: 197-205.

Yirgu M, Kebede M, Feyissa T, et al., 2024. Single nucleotide polymorphism (SNP) markers for genetic diversity and population structure study in Ethiopian barley (*Hordeum vulgare* L.) germplasm. *BMC Genome* 24: 7.