



Revista Industrial
y Agrícola de
Tucumán

ISSN 0370-5404

En línea
1851-3018

Tomo 102 (2):
34-40; 2025



ESTACION EXPERIMENTAL
AGROINDUSTRIAL
OBISPO COLOMBRES
Tucumán | Argentina

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

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

Fecha de
recepción:
05/09/2025

Fecha de
aceptación:
08/09/2025

Precisión de la selección genómica en poblaciones del programa de mejoramiento de caña de azúcar de la EEAOC

Josefina Racedo*, Ezequiel Rossi**, Andrea N. Peña*, Ana C. Ghio***, Diego D. Henriquez***, Jorge V. Díaz***, Aldo S. Noguera*, María F. Perera* y Santiago Ostengo*

* Instituto de Tecnología Agroindustrial del Noroeste Argentino (ITANOA), Estación Experimental Agroindustrial Obispo Colombres (EEAOC) - Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). CCT CONICET NOA Sur. Las Talitas, Tucumán, T4101XAC, República Argentina. ** Instituto de Investigaciones Agrobiotecnológicas (INIAB, CONICET-UNRC). Río Cuarto, Córdoba, Argentina. *** Estación Experimental Agroindustrial Obispo Colombres (EEAOC). Las Talitas, Tucumán, T4101XAC, República Argentina. Email: joracedo@gmail.com

RESUMEN

La selección genómica (SG) es una herramienta prometedora para incrementar la eficiencia en el mejoramiento de caracteres complejos. Se evaluó la precisión de predicción de los valores de mejora estimados a partir de datos genómicos (GEBV, por sus siglas en inglés de *Genomic-estimated breeding value*) en distintas poblaciones de diferentes etapas del Programa de Mejoramiento Genético de la Caña de Azúcar de la Estación Experimental Agroindustrial Obispo Colombres (PMGCA-EEAOC). Los modelos de SG fueron entrenados en tres poblaciones de mejoramiento fenotipadas para caracteres de maduración temprana y genotipadas con marcadores SNP obtenidos por DArTseq. Las poblaciones de mejoramiento consistieron en: i) 264 accesiones del banco de germoplasma (BG); ii) 160 individuos en la segunda etapa clonal (2C); y iii) 47 individuos de la etapa de Ensayos Internos de variedades en campo (I). Para todas las predicciones genómicas se utilizó el modelo de predicción lineal insesgada mejorado por genómica (G-BLUP por sus siglas en inglés de *Genomic best linear unbiased prediction*). La eficiencia de la SG fue evaluada mediante validación interna y cruzada, dependiendo de la población. Para BG, la población de entrenamiento (PE) consistió en 264 accesiones, y las precisiones se obtuvieron mediante correlaciones dentro de la población usando una partición 80:20. Para 2C, la eficiencia se evaluó dentro de la población utilizando una partición 80:20, y el modelo estimado se aplicó para estimar los GEBVs en la población I (validación cruzada). Además, se analizó una PE conformada por la combinación de las poblaciones 2C e I, con precisión obtenida a través de correlaciones dentro de la población (80:20). Las mayores eficiencias de SG para BG se observaron en los caracteres pol y azúcar recuperado ($r = 0,43$ y $0,40$, respectivamente). Para 2C e I, las mayores eficiencias de SG se observaron en el contenido de azúcar, tanto al considerar ambas poblaciones conjuntamente ($r = 0,41$), como cuando la PE consistió solo en 2C ($r = 0,43$). Dado que la efectividad de la SG en programas de mejoramiento depende de los datos fenotípicos y moleculares, el modelo de predicción, y el tamaño y composición de la PE, estos resultados son alentadores para continuar desarrollando estrategias que optimicen la PE con el fin de alcanzar mayores precisiones.

Palabras clave: GEBV, marcadores moleculares, DArT-seq, mejoramiento molecular.

ABSTRACT

Accuracy of genomic selection in populations of the EEAOC sugarcane breeding program

Genomic selection (GS) is a promising breeding tool for improving the efficiency of complex trait breeding. The prediction accuracy of genomic breeding values was assessed across populations at different stages of the Sugarcane Breeding Program at the Estación Experimental Agroindustrial Obispo Colombres (SCBP-EEAOC). GS models were tested in three breeding populations phenotyped for early maturation traits and genotyped using DArTseq SNP markers. The breeding populations consisted of: i) 264 accessions from germplasm bank (GB); ii) 160 individuals in the second clonal stage (2CS); and iii) 47 individuals from infield variety trials stage (IVT). The genomic best linear unbiased prediction

(G-BLUP) model was used for all genomic predictions. The efficiency of GS was evaluated through intra- and cross-validation, depending on the population. For GB, the training population (TR) consisted of 264 accessions, and accuracies were obtained by correlations within population using an 80:20 split. For 2CS, efficiency was evaluated within the population using an 80:20 split, and the estimated model was applied to estimate genomic estimated breeding values (GEBVs) for population IVT (cross-validation). Additionally, a TR consisting of combined 2CS and IVT populations was analyzed, with accuracy obtained through correlations within the population (80:20). The highest GS efficiencies for GB were observed for pol and sugar-recovered traits ($r = 0.43$ and 0.40 , respectively). For 2CS and IVT, the highest GS efficiencies were observed for sugar content when considering both populations jointly ($r = 0.41$), and when the TR consisted only of 2CS ($r = 0.43$). Since the effectiveness of GS in breeding programs depends on the phenotypic and molecular data, the prediction model, and the size and composition of the TR, these results are encouraging for the continued development of strategies to optimize the TR to achieve better accuracies.

Key words: GEBV, molecular markers, DArT-seq, molecular breeding.

INTRODUCCIÓN

La selección genómica (SG) es una estrategia cada vez más importante que complementa el mejoramiento convencional. A diferencia de la selección asistida por marcadores (SAM), que utiliza un número reducido de marcadores previamente asociados a genes o QTLs (por sus siglas en inglés de *quantitative trait loci*) de interés, la SG utiliza simultáneamente todos los marcadores distribuidos en el genoma para predecir el valor genético o el desempeño fenotípico de los individuos en una población, incluyendo aquellos QTLs con efectos menores (Meuwissen *et al.*, 2001).

Este valor genético, denominado Valor de Mejora Estimado a partir de datos Genómicos (GEBV, por sus siglas en inglés de *Genomic-estimated breeding value*), resulta de aplicar un modelo estadístico predictivo que integra la información de los marcadores moleculares obtenidos mediante genotipado de alta densidad con los datos fenotípicos de los caracteres de interés. Una vez entrenado el modelo, la implementación exitosa de SG en un programa de mejoramiento requiere validación cruzada, tanto dentro de la población como entre poblaciones. Durante la validación, se evalúa la capacidad predictiva del modelo calibrado comparando los valores predichos con los observados en una nueva población. Una buena calibración asegura que los GEBVs se correlacionen suficientemente con los valores de mejora derivados de los datos fenotípicos.

Dada la importancia de incrementar la eficiencia del Programa de Mejoramiento Genético de la Caña de Azúcar de la Estación Experimental Agroindustrial Obispo Colombres (PMGCA-EEAOC), se implementó un proyecto de SG. Inicialmente, se entrenaron distintos modelos de SG con una población de 80 genotipos avanzados del esquema de mejoramiento para caracteres de maduración temprana. Aunque los modelos explicaron una porción significativa de la variabilidad del mérito genético, la eficiencia de SG fue baja debido a la limitada variabilidad genética y al reducido tamaño de la población de entrenamiento (Ostengo, 2020).

En 2016, se amplió la población de entrenamiento a 188 accesiones del banco de germoplasma (BG), incrementando tanto el número de individuos evaluados como la variabilidad genética. Esta población fue fenotipada entre 2016 y 2021 para el número y peso de tallos, la resistencia a enfermedades y la calidad industrial, y además fue genotipada mediante GBS (por sus siglas en inglés de *Genotyping by sequencing*). Los modelos de SG fueron ajustados para estimar los GEBVs individuales, útiles para diseñar futuros cruzamientos (Racedo *et al.*, 2023).

El objetivo del presente estudio fue implementar SG, incluyendo poblaciones de clones en diferentes etapas del esquema de mejora, rutinariamente fenotipadas por el PMGCA-EEAOC, y realizar validaciones intra e interpopulacionales para evaluar la precisión de las estimaciones. Los resultados permitirán comparar distintas estrategias de implementación de SG dentro del programa y determinar cuál es la más precisa, eficiente y factible para acortar el ciclo de mejoramiento y acelerar la tasa de ganancia genética.

MATERIALES Y MÉTODOS

Material vegetal, genotipado y fenotipado

• Población del Banco de Germoplasma (BG)

Se seleccionaron 303 genotipos de caña de azúcar de un total de 789 accesiones del banco de germoplasma del PMGCA-EEAOC (Ostengo *et al.*, 2021). La selección priorizó incluir los individuos más utilizados como progenitores en cruzamientos, añadiendo otros genotipos para capturar la diversidad alélica. Se recolectaron muestras de hojas en campo para realizar extracciones de ADN siguiendo las recomendaciones del servicio de genotipado de alta densidad DArTseq.

Los ensayos fenotípicos se llevaron a cabo en el campo experimental de la EEAOC (Las Talitas, departamento Tafí Viejo, Tucumán, Argentina), donde cada genotipo estuvo representado por parcelas de 2 surcos \times 3 metros. Se evaluaron los componentes de calidad industrial

azúcar recuperada (AR %) y Pol % caña, según Diez *et al.* (2000). Para ello, se recolectaron 10 tallos por parcela en la edad de planta, que fueron limpiados, cortados y procesados en laboratorio con trituradora de martillos (95% de celdas abiertas), extrayéndose el jugo mediante prensa hidráulica a $240 \text{ kg cm}^{-2} \text{ min}^{-1}$. La biomasa se estimó mediante el peso de 10 tallos tomados al azar de cada parcela.

- **Población de la segunda etapa clonal (2C)**

Se realizó el genotipado de alta densidad en 200 individuos de la segunda etapa clonal (2C) del esquema de mejora del PMGCA-EEAOC mediante DArTseq. Los datos fenotípicos se recolectaron en ensayos con diseño de bloques completos al azar (DBCA) de 3 surcos \times 3 metros, con dos repeticiones y dos cortes sucesivos en un sitio. Se midieron contenido de sacarosa (CS%), número de tallos por metro y peso de tallo, procesando 10 tallos por parcela central en un molino experimental de laboratorio. El contenido de azúcar se calculó a partir de Brix y Pol % jugo según Fogliata (1995):

$$CS\% = 0,98 \times Pol\% \text{jugo} - 0,28 \times Brix$$

Los datos fenotípicos se analizaron mediante modelos lineales mixtos para obtener los BLUPs por carácter.

- **Población de Ensayos Internos de variedades en campo (I)**

Se genotiparon 50 clones correspondientes a la etapa de Ensayos Internos de variedades (I), siguiendo el procedimiento de DArTseq. Los ensayos fenotípicos se organizaron en parcelas de 3 surcos \times 8 metros, con DBCA y tres repeticiones, evaluándose durante dos cortes consecutivos en dos localidades (El Colmenar y Santa Ana, provincia de Tucumán). Se determinaron contenido de azúcar y componentes de rendimiento (número de tallos por metro y peso de tallo) en caña planta y primera soca. Los datos se analizaron mediante modelos lineales mixtos apropiados para evaluación multiam biental, obteniéndose los predictores genéticos (BLUPs).

- **Caracterización de poblaciones mediante análisis conjunto de datos fenotípicos y genotípicos (Procrustes) y ACoP**

Los caracteres fenotípicos se analizaron mediante análisis de componentes principales (ACP), utilizando distancia euclidiana como medida de disimilitud (Infostat, Di Rienzo *et al.*, 2020). Para la caracterización molecular, se calculó la distancia genética de Nei (1972) a partir de todos los marcadores polimórficos, y se construyeron matrices de distancia para el ordenamiento. La ordenación se realizó mediante análisis de coordenadas principales (ACoP) y escalamiento multidimensional (MDS por sus siglas en inglés de *Multidimensional Scaling*).

El grado de concordancia entre ordenamientos fenotípicos y genotípicos se evaluó mediante análisis de Procrustes generalizado (Gower, 1975), considerando las dos primeras coordenadas principales de cada conjunto de datos.

Además, se visualizaron las relaciones genéticas entre las tres poblaciones mediante ACoP (Infostat, 2020).

- **Datos genotípicos para selección genómica**

Se analizaron las frecuencias alélicas en cada población para excluir marcadores con frecuencia menor a 0,1 (MAF de las siglas en inglés de *minor allele frequency* < 0,1). Se incluyeron marcadores con tasa de llamado superior a 0,8 en los modelos de SG.

- **Modelos de selección genómica y validación cruzada**

Para todas las predicciones genómicas se utilizó el modelo G-BLUP. Los ajustes se realizaron mediante cadenas de Markov Monte Carlo (MCMC), usando Gibbs sampler con 50.000 iteraciones y un burn-in de 5000, implementados con el paquete BGLR en R (de los Campos y Pérez-Rodríguez, 2014; R Core Team, 2024).

La eficiencia de la SG se evaluó mediante validación intra e interpoblacional, según corresponda. Para BG, se utilizó partición 80:20 para obtener las precisiones mediante correlaciones dentro de la población, replicando 100 veces con asignaciones aleatorias independientes para entrenamiento y prueba. Para 2C, se evaluó intrapoblacional con 80:20, y el modelo estimado se aplicó para predecir GEBVs en la población I (validación cruzada). Además, se analizó una PE combinando 2C e I, con precisión obtenida mediante correlaciones dentro de la población (80:20).

RESULTADOS Y DISCUSIÓN

El desbalance entre el gran número de marcadores y el menor número de observaciones fenotípicas puede influir en la precisión del modelo G-BLUP. Aunque G-BLUP incorpora de manera inherente regularización mediante la matriz de relaciones genómicas (G), un exceso de marcadores puede introducir ruido y reducir la precisión predictiva, particularmente si muchos de ellos no están relacionados con el carácter de interés (VanRaden, 2008; Hayes *et al.*, 2009). Estrategias como filtrar marcadores de baja calidad (Wray *et al.*, 2013) o aumentar el número de observaciones fenotípicas pueden ayudar a mitigar este efecto.

En este estudio se obtuvieron 74.969 SNPs (por sus siglas en inglés de *single nucleotide polymorphism*) de todas las poblaciones mediante DArTseq. Tras filtrar los SNPs con MAF < 0,1, se incluyeron 35.869 SNPs para los procedimientos de SG. Algunos genotipos se excluyeron debido a la baja calidad de secuenciación, probablemente por degradación de ADN durante el envío al servicio de genotipado. Como resultado, los modelos de SG se entrenaron en tres poblaciones de mejoramiento fenotipadas para caracteres de maduración temprana: i) 264 accesiones del Banco de Germoplasma (BG); ii) 160 individuos de la Segunda Etapa Clonal (2C); y iii) 47 individuos de la etapa de Ensayos Internos de Variedades (I).

Caracterización de poblaciones mediante análisis conjunto de datos genotípicos y fenotípicos (Procrustes) y ACoP

Para visualizar las relaciones entre los individuos,

las poblaciones fueron caracterizadas mediante un análisis conjunto de datos fenotípicos y genotípicos usando análisis de Procrustes. Este enfoque redujo la dimensionalidad de las matrices de datos de los caracteres fenotípicos y genotípicos a dos ejes. Para la población BG, el ACoP explicó el 7% de la variabilidad total observada con 14.725 SNPs polimórficos en los ejes 1 y 2, mientras que el ACP de los datos fenotípicos explicó el 100% de la variabilidad en los dos primeros ejes (Figura 1). Para la

población 2C, el análisis ACoP explicó el 11% de la variabilidad total observada con 15.600 SNPs polimórficos en los ejes 1 y 2, mientras que el ACP de los datos fenotípicos representó el 75% de la variabilidad en los dos primeros ejes (Figura 2). Para la población I, el ACoP explicó el 13% de la variabilidad total observada con 17.039 SNPs polimórficos en los ejes 1 y 2, mientras que el ACP de los datos fenotípicos explicó el 89% de la variabilidad en los dos primeros ejes (Figura 3).

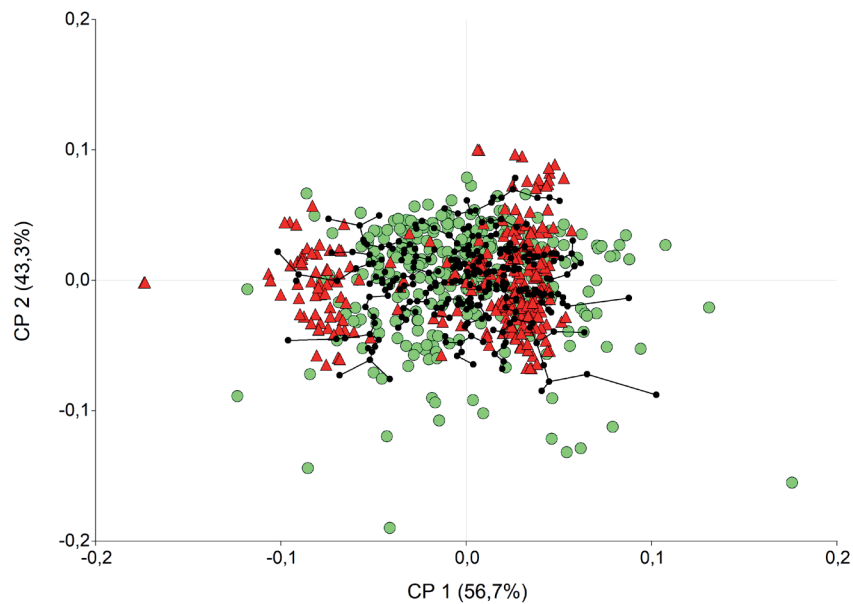


Figura 1. Proyección de las coordenadas CP1 y CP2 con el Árbol de Expansión Mínima para la configuración de consenso (línea negra), configuración fenotípica (círculos verdes) y configuración molecular (triángulos rojos), derivada del Análisis Generalizado de Procrustes de la población del Banco de Germoplasma (BG).

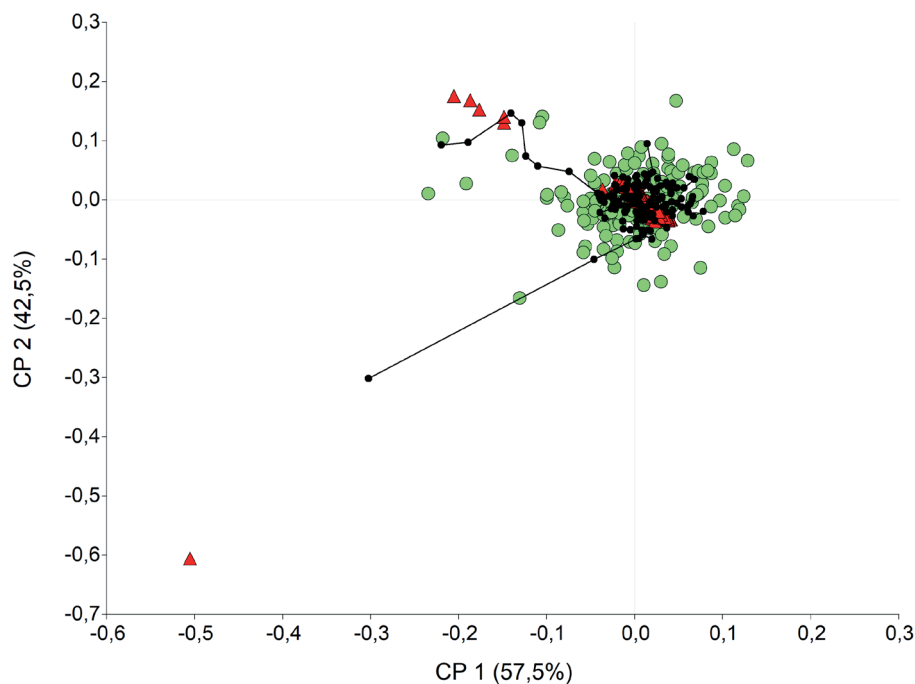


Figura 2. Proyección de las coordenadas CP1 y CP2 con el Árbol de Expansión Mínima para la configuración de consenso (línea negra), configuración fenotípica (círculos verdes) y configuración molecular (triángulos rojos), derivada del Análisis Generalizado de Procrustes de la población de la Segunda Etapa Clonal (2C).

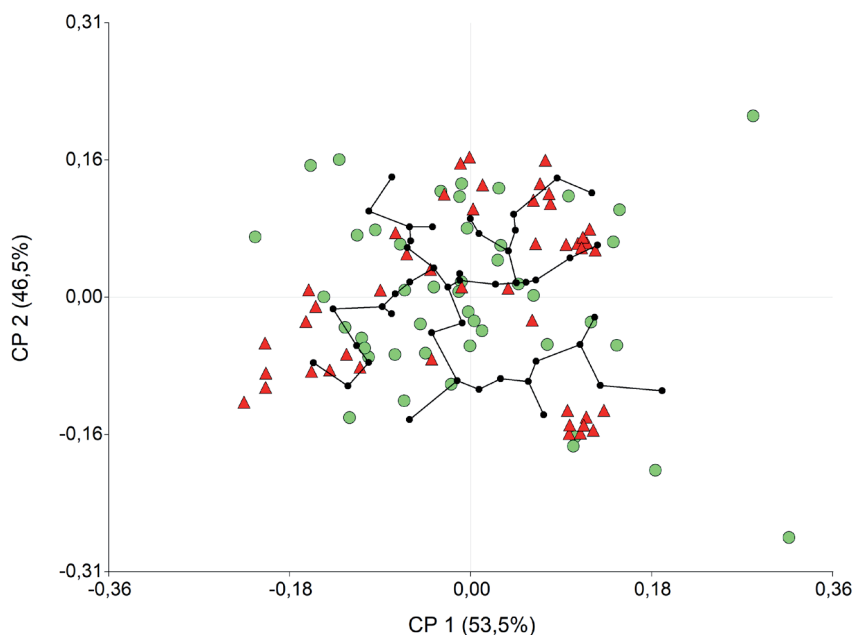


Figura 3. Proyección de las coordenadas CP1 y CP2 con el Árbol de Expansión Mínima para la configuración de consenso (línea negra), configuración fenotípica (círculos verdes) y configuración molecular (triángulos rojos), derivada del Análisis Generalizado de Procrustes de la población de Ensayos Internos de Variedades (I).

El análisis de Procrustes reveló correlaciones moderadas entre los datos moleculares y fenotípicos en las tres poblaciones analizadas, con la correlación más alta en la población 2C (0,631), seguida de la población BG (0,595) y la población I (0,547). Estos resultados indican una congruencia razonable entre los marcadores moleculares y los caracteres fenotípicos, con solo diferencias menores observadas entre poblaciones. Esta congruencia moderada sugiere que las relaciones entre los individuos se reflejan de manera similar tanto en los datos fenotípicos como genotípicos, indicando que ambos tipos de datos capturan patrones de variabilidad comparables (Bramardi *et al.*, 2005). En consecuencia, esta alineación implica que las mediciones fenotípicas podrían reflejar efectivamente

información subyacente similar a la de los datos moleculares, a pesar de las diferencias en la precisión de la medición fenotípica entre poblaciones. La consistencia general sugiere que los datos moleculares podrían predecir razonablemente la variación fenotípica en estas poblaciones.

La precisión de las predicciones de selección genómica depende no solo del tamaño de la población de entrenamiento, sino también de su diversidad, de su estructura genética y de las relaciones entre los individuos de las poblaciones de entrenamiento y de prueba (Berro *et al.*, 2019). Para visualizar las relaciones genéticas entre las tres poblaciones estudiadas, se realizó un ACoP (Figura 4). Nuestros resultados sugieren que la población BG podría entrenar de manera efectiva modelos para

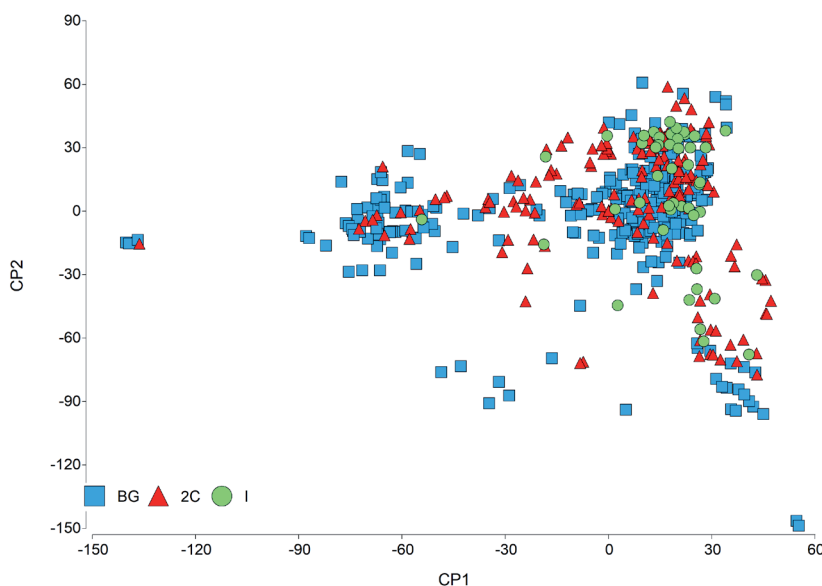


Figura 4. Análisis de Coordenadas Principales (APoC). Se utilizaron 35.869 SNPs derivados de DArT-seq para visualizar las relaciones genéticas entre las poblaciones de caña de azúcar BG, 2CS e I estudiadas.

predecir GEBVs en las poblaciones 2C e I, dado que su amplia diversidad alélica abarca ambas poblaciones objetivo. Sin embargo, aunque BG presenta una diversidad genética extensa, las subestructuras subyacentes deberían considerarse en análisis futuros. Por el contrario, predecir GEBVs de I usando como entrenamiento a la 2C probablemente produciría resultados más débiles, porque pocos individuos de 2CS estaban relacionados con los de I, reduciendo el poder predictivo del modelo (Habier *et al.*, 2007). Estos hallazgos resaltan la importancia de seleccionar una PE que equilibre de manera óptima diversidad, estructura genética y parentesco para mejorar la precisión de la predicción genómica.

Selección genómica

Para BG, la PE consistió en 264 accesiones, y las precisiones se obtuvieron usando una partición 80:20 (Tabla 1).

Para 2C, la eficiencia se evaluó dentro de la población usando una partición 80:20 (Tabla 2), y el modelo estimado se aplicó para estimar los GEBVs de la población I (validación cruzada) (Tabla 3). Además, se analizó una PE conformada por la combinación de las poblaciones 2C e I, obteniéndose la precisión mediante correlaciones dentro de la población (partición 80:20) (Tabla 4).

Las mayores eficiencias de SG para BG se observaron en los caracteres Pol % y azúcar recuperada ($r = 0,43$ y $0,40$, respectivamente; Tabla 1). Para 2C e I, las mayores eficiencias de SG se observaron en el carácter Contenido de azúcar cuando ambas poblaciones se analizaron conjuntamente ($r = 0,41$; Tabla 4) y cuando la PE consistió únicamente en 2C ($r = 0,43$; Tabla 2).

Estas precisiones son consistentes con las reportadas en estudios previos de SG en caña de azúcar (Gouy *et al.*, 2013; Deomano *et al.*, 2020; Hayes *et al.*, 2022; Shahi *et al.*, 2025). Este enfoque es particularmente relevante para caracteres como contenido de azúcar y Pol %, donde los efectos genéticos aditivos controlan una proporción significativa de la variabilidad genética.

Las precisiones para caracteres relacionados con el rendimiento de biomasa (tallos por metro y peso de tallo) fueron menores en todos los casos. En estos caracteres, la varianza genética no aditiva contribuye significativamente a la varianza genética total y está fuertemente influenciada por factores ambientales.

El éxito de la SG en programas de mejoramiento depende de múltiples factores, entre ellos la calidad y exhaustividad de los datos fenotípicos, el modelo de predicción empleado y el número y tipo de marcadores moleculares, así como el tamaño y la composición de la PE. Además, la relación genética entre la PE y la población de prueba es uno de los determinantes más importantes para maximizar la capacidad predictiva. En el caso de BG, en comparación con nuestro estudio previo (Racedo *et al.*, 2023), aunque el tamaño de la población es mayor, las precisiones de predicción no mostraron mejoras. Esto subraya que un requisito clave para alcanzar predicciones confiables es la calidad de los datos fenotípicos. No obstante, ampliar el tamaño de las parcelas y las repeticiones para la totalidad de la colección de germoplasma (789 genotipos) resultaría poco práctico. A partir de estos resultados, se definirá una subpoblación de genotipos del banco de germoplasma que será evaluada en campo me-

Tabla 1. Valores medios de los coeficientes de correlación ($N=100$) y desviación estándar para el modelo de SG G-BLUP aplicado a la población del Banco de Germoplasma (BG) de caña de azúcar para contenido temprano de sacarosa y peso de tallo. La precisión se evaluó dentro de la población (80:20).

Variable	Media	Desvío estándar
Pol%	0,43	0,13
Azúcar recuperable	0,40	0,13
Peso del tallo	0,38	0,11

Tabla 2. Valores medios de los coeficientes de correlación ($N=100$) y desviación estándar para validación intra-poblacional con partición 80:20 usando el modelo G-BLUP de SG aplicado a la población de la Segunda Etapa Clonal (2CS) (160 genotipos) de caña de azúcar para caracteres de rendimiento.

Variable	Media	Desvío estándar
Contenido de azúcar	0,43	0,14
Tallos por metro	0,26	0,21
Peso del tallo	0,01	0,11

Tabla 3. Valores medios de los coeficientes de correlación ($N=100$) y desviación estándar para validación cruzada inter-poblacional considerando la población de la Segunda Etapa Clonal (2CS) como conjunto de entrenamiento (160 genotipos) y la población de Ensayos Internos de Variedades (I) como conjunto de prueba (47 genotipos). Los GEBVs estimados para I se correlacionaron con sus valores fenotípicos. El modelo G-BLUP de selección genómica se implementó para caracteres de rendimiento.

Variable	Media	Desvío estándar
Contenido de azúcar	0,04	0,04
Tallos por metro	0,06	0,06
Peso del tallo	0,05	0,06

Tabla 4. Valores medios de los coeficientes de correlación ($N=100$) y desviación estándar para validación intra-poblacional con partición 80:20 usando el modelo G-BLUP de selección genómica aplicado a las poblaciones combinadas Segunda Etapa Clonal (2C) e Interna (I) (total de 207 genotipos) de caña de azúcar para caracteres de rendimiento.

Variable	Media	Desvío estándar
Contenido de azúcar	0,41	0,12
Tallos por metro	0,09	0,14
Peso del tallo	0,03	0,12

dante un diseño experimental adecuado, con el objetivo de optimizar la población de entrenamiento para futuras predicciones.

La precisión obtenida en la 2C para el rendimiento de azúcar es prometedora; sin embargo, cuando los GEBVs de I se estimaron usando el modelo entrenado en la 2C, las precisiones fueron bajas, indicando que no es posible predecir el desempeño de clones avanzados basado en el modelo entrenado en la etapa anterior del esquema de mejoramiento.

CONCLUSIONES

Basado en estudios previamente reportados, los valores de precisión que obtuvimos son alentadores para la implementación de SG en el PMGCA de la EEAOC. Los resultados son prometedores para el desarrollo continuo de estrategias que optimicen la metodología y logren mayores precisiones.

Usando el modelo G-BLUP implementado, se obtuvieron GEBVs para un total de 510 individuos, proporcionando información valiosa para el PMGCA de la EEAOC. Estos GEBVs pueden asistir en la caracterización de líneas parentales, facilitando el diseño de cruzamientos más efectivos y contribuyendo al desarrollo de poblaciones mejoradas.

BIBLIOGRAFÍA CITADA

- Berro, I.; B. Lado; R. S. Nalin; M. Quincke and L. Gutiérrez. 2019.** Training population optimization for genomic selection. *Plant Genome* 12: 190028.
- Bramardi, S. J.; G. P. Bernet; M. J. Asíns and E. A. Carbonell. 2005.** Simultaneous agronomic and molecular characterization of genotypes via the generalised procrustes analysis: an application to cucumber. *Crop Science* 45: 1603–1609.
- de los Campos, G. and P. Pérez-Rodríguez. 2014.** Bayesian generalized linear regression. R package version 1(4).
- Deomano, E.; P. Jackson; X. Wei; K. Aitken; R. Kota and P. Pérez-Rodríguez. 2020.** Genomic prediction of sugar content and cane yield in sugar cane clones in different stages of selection in a breeding program, with and without pedigree information. *Molecular Breeding* 40: 38.
- Diez, O.; S. Zossi; E. R. Chavanne y G. Cárdenas. 2000.** Calidad industrial de las cañas de azúcar de maduración temprana LCP85-384 y LCP85-376 en Tucumán. Análisis de sus principales constituyentes físico-químicos. *Revista Industrial y Agrícola de Tucumán* 77 (2): 39–48.
- Fogliata, F. A. 1995.** Composición química y vegetativa. In: Felipe, L. E. (ed.), *Agronomía de la caña de azúcar*. Argentina: El Graduado, pp. 163–209.
- Gower, J. C. 1975.** Generalized Procrustes analysis. *Psychometrika* 40: 33–51.
- Gouy, M.; Y. Rousselle; D. Bastianelli et al. 2013.** Experimental assessment of the accuracy of genomic selection in sugarcane. *Theoretical and Applied Genetics* 126: 2575–2586.
- Habier, D.; R. L. Fernando and J. C. M. Dekkers. 2007.** The impact of genetic relationship information on genome-assisted breeding values. *Genetics* 177: 2389–2397.
- Hayes, B. J.; P. J. Bowman; A. J. Chamberlain and M. E. Goddard. 2009.** Genomic selection in dairy cattle: Progress and challenges. *Journal of Dairy Science* 92: 433–444.
- Hayes, B. J.; X. We; P. Joyce et al. 2022.** Accuracy of genomic prediction of complex traits in sugarcane. *Theoretical and Applied Genetics* 134: 1455–1462.
- Infostat. Di Rienzo, J. A.; F. Casanoves; M. G. Balzarini; L. Gonzalez; M. Tablada y C. W. Robledo. 2020.** Infostat versión 2020. Grupo Infostat, FCA, Universidad Nacional de Córdoba, Argentina.
- Meuwissen, T. H.; B. J. Hayes and M. E. Goddard. 2001.** Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157: 1819–1829.
- Nei, M. 1972.** Genetic distance between populations. *American Naturalist* 106: 283–292.
- Ostengo, S. 2020.** Aproximaciones metodológicas para la caracterización de la calidad industrial de genotipos de caña de azúcar. PhD thesis. National University of Tucumán, Argentina.
- Ostengo, S.; G. Serino; M. F. Perera et al. 2021.** Sugarcane breeding, germplasm development and supporting genetics research in Argentina. *Sugar Tech* 24: 166–180.
- R Core Team. 2024.** R: a language and environment for statistical computing. R package version 4.0.3. Accessed 10 Jun 2024. <http://www.r-project.org>
- Racedo, J.; E. A. Rossi; M. Aybar Guchea et al. 2023.** Genomic selection for traits of interest in the EEAOC sugarcane breeding program. *Proceedings of the International Society of Sugar Cane Technologists* 31: 775–781.
- Shahi, D.; J. Todd; K. Gravois et al. 2025.** Exploiting historical agronomic data to develop genomic prediction strategies for early clonal selection in the Louisiana sugarcane variety development program. *The Plant Genome* 18: e20545.
- VanRaden, P. M. 2008.** Efficient methods to compute genomic predictions. *Journal of Dairy Science* 91: 4414–4423.
- Wray, N. R.; J. Yang; B. J. Hayes; A. L. Price; M. E. Goddard and P. M. Visscher. 2013.** Pitfalls in predicting complex traits from SNPs. *Nature Reviews Genetics* 14: 507–515.