

MEJORAMIENTO GENÉTICO DE LA CAÑA DE AZÚCAR MEDIANTE EL USO DE MARCADORES MOLECULARES EN MÉXICO

GENETIC IMPROVEMENT OF SUGARCANE USING MOLECULAR MARKER IN MEXICO

Hilda Victoria Silva-Rojas¹, Paulino Pérez-Rodríguez¹, Carlos Flores-Revilla², Carlos Aaron Rangel-Ortega², Benjamín Cervantes-Romero¹, José Manuel Aguirre-Rayo¹, Camilo Hernández-Juarez¹, Gustavo Alonso Pinto-Hernández² y Roberto Loyo-Joachín²

¹Colegio de Postgraduados, Campus Montecillo, Texcoco, Estado de México CP 56230 México

²Centro de Investigación y Desarrollo de la Caña de Azúcar A.C., Ciudad de México CP 06500 México

Resumen

La caña de azúcar (*Saccharum* spp. híbridos) es uno de los principales cultivos agroindustriales a nivel mundial para la producción de azúcar. Es una gramínea alloploiploide, lo que ha restringido el uso de estudios genéticos clásicos en los programas de mejoramiento de este cultivo. Una de las técnicas que ha dado los mejores resultados para la genotipificación de clones de caña de azúcar es el uso de marcadores DArT (Diversity Array Technology) que permite estimar la diversidad genética de la caña de azúcar basado en la reducción eficiente del genoma. Por tal razón, el objetivo del presente trabajo fue genotipificar y determinar las distancias genéticas de 93 clones procedentes de la Estación de Hibridación del CIDCA en Chiapas. La extracción del DNA se realizó a partir de hojas jóvenes de cada clon, los que se enviaron a la compañía DArT en Australia para el análisis del genoma con una alta profundidad. Los resultados mostraron 52,828 marcadores de tipo “presencia” (1) y “ausencia” (0) de un fragmento de restricción que contiene el marcador de interés. Se eliminaron marcadores de baja calidad, obteniéndose 43,250 marcadores. Una pequeña fracción de los datos presentó datos faltantes, razón por la cual fueron imputados generando muestras aleatorias de la distribución marginal de los genotipos observados. A partir de los marcadores se obtuvo la matriz de distancias Euclidianas (E) y la matriz de relaciones genómicas (G). La matriz E fue utilizada para construir un agrupamiento jerárquico de los individuos. Los clones también se agruparon utilizando partición basada en medioides, la gráfica de siluetas resultó en la formación de seis grupos. Una de las principales contribuciones de la presente investigación fue el conocimiento por primera vez de las distancias genéticas de 93 clones que podrán ser utilizadas en el programa anual de mejoramiento genético del CIDCA en México.

Palabras clave: cruzamiento, DArT, distancias Euclidianas, poliploide, matriz de relaciones genómicas

Abstract

Sugarcane (*Saccharum* spp. hybrids) is one of the world's leading agro-industrial crops for sugar production. It is an alloploid grass, which has restricted the use of classical genetic studies in breeding programs of this crop. One technique that has given the best results for genotyping of sugarcane clones is the use of DArT (Array Diversity Technology) molecular marker to estimate genetic diversity of sugarcane based on efficient genome reduction. For this reason, the objective of the present work was genotyped and to determine the genetic distances of 93 clones from the CIDCA Hybridization Station in Chiapas. DNA extraction was performed from young leaves of each clone, which were sent to DArT Company in Australia for genome analysis with a high depth. The results showed 52,828 "presence" (1) and "absence" (0) type markers of a restriction fragment containing the marker of interest. Low quality markers were eliminated, obtaining 43,250 markers. A small fraction of the data presented missing data, which is why they were imputed generating random samples of the marginal distribution of the observed

genotypes. From the markers the matrix of Euclidean distances (E) and the matrix of genomic relations (G) were obtained. The E matrix was used to construct a hierarchical grouping of individuals. The clones were also grouped using meiotic-based partition, the silhouetted graph resulted in the formation of six clusters. One of the major contributions of this research was the first knowledge of the genetic distances of 93 varieties that may be used in the CIDCA annual breeding program in Mexico.

Keywords: Crosses, DArT, Euclidean distances, polyploid, matrix of genomic relations

Introducción

La caña de azúcar (*Saccharum* spp. híbridos) es uno de los principales cultivos agroindustriales a nivel mundial para la producción de azúcar (Dookun-Saumtally et al., 2012). En una gramínea alloploide originaria de Nueva Guinea, cuyo genoma contiene información genética de *Saccharum officinarum* y *S. spontaneum* con un número variable de cromosomas que varía de $2n = 100$ a 130 . En el complejo '*Saccharum*' se encuentran especies con diferentes niveles de ploidia y con un número variable de cromosomas entre $2n = 20$ a ~ 200 . La alta poliploidia y heterocigosidad debido a la hibridación ha restringido el uso de estudios genéticos clásicos en este cultivo (Premachandran et al., 2011). Sin embargo, esta condición de poliploide le permite tener un gran rango de adaptabilidad a diferentes condiciones ambientales en regiones tropicales y subtropicales de mundo, sembrándose en más de 103 países en los cinco continentes (Mendes de Paula et al., 2014).

En México, desde su introducción en 1522, se ha sembrado en diferentes climas y condiciones topográficas, aunque tiene un gran rango de adaptación, se requiere de la utilización de variedades que tengan un alto índice de respuesta en cada zona y que permita expresar su máximo potencial productivo de tal manera que sea económicamente rentable para el productor (Aguilar-Rivera et al., 2010).

En el ámbito internacional, México es reconocido por ser el sexto país productor de caña de azúcar, sin embargo el 77% de las aproximadamente 824,000 ha se siembran con cinco variedades (CONADESUCA, 2016; Senties-Herrera et al., 2014). Considerando que estas variedades se obtuvieron hace más de 30 ó 40 años, se están volviendo vulnerables a las condiciones de cambio climático, por lo que se tiene la necesidad de renovar y diversificar las variedades utilizadas en las zonas productoras de este cultivo, que traerá beneficios directos tanto al sector primario como a la industria mejorando los rendimientos de campo y fábrica (Aguilar-Rivera et al., 2010).

El programa de mejoramiento genético para la obtención de nuevas variedades de caña de azúcar estuvo a cargo del Instituto para el Mejoramiento de la Producción de Azúcar (IMPA) quien coordinaba trabajos de mejoramiento genético de la caña hasta 1990. A la fecha estas investigaciones han sido retomadas por el Centro de Investigación y Desarrollo de la Caña de Azúcar A.C., quien cuenta con un banco de germoplasma con más de 2,800 accesiones. Este programa de mejoramiento está enfocado en la producción de nuevas variedades con características agronómicas (buen porte de las plantas, resistentes al acame, resistentes a plagas y enfermedades etc.), genéticas (presencia de genes de rendimiento, producción de azúcar etc.) e industriales (porcentaje de sacarosa, fibras, mieles etc.).

Debido al periodo que se necesita para poder liberar una variedad entre 14 y 16 años se ha buscado alternativas que permitan reducir estos tiempos, por lo que se ha considerado el uso de técnicas modernas basadas en marcadores moleculares. Esta metodología ha sido empleada de forma exitosa en otras especies, e.g. maíz, trigo, cebada, etc.

Estudios en caña de azúcar realizados en Australia (Aitken et al., 2014; Al-Khayri et al., 2015; Berkman et al., 2013; Huang et al., 2010) y Brasil (García et al., 2006; García et al., 2013) han permitido identificar numerosos marcadores asociados con caracteres de importancia agronómica, además del desarrollo de mapas genéticos, entre otros avances (Jaccoud et al., 2001; Manners, 2011). Una de las técnicas que ha dado los mejores resultados para la genotipificación de clones de caña de azúcar es la técnica de Diversity Array Technology (DART) desarrollada por Heller-Uszynska et al. (2010) se ha convertido en una novedosa plataforma para el análisis genético, genotipificación de todo el genoma y estudios de la diversidad genética de caña de azúcar basado en la reducción eficiente del genoma.

Ante este panorama, resulta evidente la posibilidad de complementar las características agronómicas e industriales de la caña de azúcar con la información genética que pueda apoyar los programas de mejoramiento para la obtención de variedades adaptadas a cada región en los que se pueda determinar la resistencia o susceptibilidad a plagas, enfermedades y con un alto contenido de sacarosa.

Por tal razón, el objetivo del presente trabajo fue determinar las distancias genéticas de 93 clones de caña de azúcar procedentes de la Estación de Hibridación del CIDCA para ser utilizados en el programa anual de cruzamientos.

Materiales y Métodos

1.1 Material de estudio

Se colectaron hojas de 93 clones de caña de azúcar establecidos en el Banco de Germoplasma de la Estación de Hibridación del CIDCA, ubicado en Tuxtla Chico, Chiapas (Figura 1).



Figura 1. Regiones productoras de caña de azúcar en México, indicando la ubicación de la Estación de Hibridación del CIDCA AC. en Tuxtla Chico, Chiapas.

2.2 Material en estudio

Se tomó una porción de tejido de hojas de los 93 clones en estudio (Tabla 1) con un guante, se colocó dentro de una bolsa plástica, la que se procedió a etiquetar y colocar de inmediato en un recipiente de unicel con hielo. Los recipientes se trasladaron al laboratorio donde se procedió a lavar la muestra con agua destilada, posteriormente se desinfectó con una solución de hipoclorito de sodio al 1% durante 3 min y se enjuagó tres veces con agua destilada estéril. Las muestras se colocaron en papel toalla estéril para eliminar el exceso de agua, se re-etiquetó y se colocaron en una nueva bolsa de plástico en un congelador a -20°C hasta su procesamiento.

Tabla 1. Relación de clones genotipificados mediante marcadores DArT-Seq.

N°	VARIEDAD	N°	VARIEDAD	N°	VARIEDAD
1	ATEMex 96-40	32	COLPOSCTMex 06-39	63	Ja 64-20
2	B 35187	33	COLPOSCTMex 06-474	64	L 60-14
3	B 45181	34	CP 34-79	65	L 68-40
4	C 323-68	35	CP 56-63	66	LCP 81-10
5	CB 36-14	36	CP 59-22	67	LGM 92-156
6	CB 40-77	37	CP 62-374	68	LTMex 93-354
7	CC 83-29	38	CP 70-1133	69	LTMex 94-2
8	CC 84-59	39	CP 70-1527	70	LTMex 96-10
9	CC 85-92	40	CP 72-1210	71	M 112/34
10	CC 86-29	41	CP 72-2086	72	Mex 07-1408
11	CC 87-505	42	CP 73-1547	73	Mex 08-1270
12	CC 92-2198	43	CP 74-2005	74	Mex 09-1358
13	CC 92-2358	44	CP 75-1091	75	Mex 09-1433
14	CC 93-3423	45	CP 75-1632	76	Mex 52-17
15	CC 93-3458	46	CP 78-1610	77	Mex 53-142
16	CC 93-3826	47	CP 80-1743	78	Mex 56-105
17	CC 93-4206	48	CP 82-2043	79	Mex 56-294
18	CC 93-4418	49	CP 84-730	80	Mex 56-476
19	CC 94-5168	50	CP 88-1508	81	Mex 57-1456
20	CC 94-5739	51	CP 89-2143	82	Mex 57-473
21	CC 98-12	52	CP 89-2377	83	Mex 58-418
22	CCSP 89-43	53	CP 92-1401	84	Mex 59-12
23	CG 97-100	54	CYZ 82-154	85	Mex 59-428
24	Cl 47-83	55	EMex 01-323	86	Mex 60-1459
25	Co 421	56	EMex 02-05	87	Mex 61-1428

26	Co 453	57	EMex 05-222	88	Mex 62-631
27	Co 467	58	EMex 05-225	89	Mex 64-406
28	ColMex 02-225	59	HOCP 93-746	90	Mex 64-649
29	COLPOSCTMex 05-204	60	ICA 76-7	91	Mex 65-1418
30	COLPOSCTMex 05-223	61	ICPMex 92-1420	92	Mex 66-1235
31	COLPOSCTMEX 06-2362	62	ITV 92-1424	93	Mex 68-1345

2.2. Extracción de DNA

La extracción de DNA se realizó con el método de Bromuro de Cetiltrimetilamonio (CTAB) /cloroformo/ alcohol isoamílico (Doyle y Doyle, 1987) con ligeras modificaciones.

La concentración de DNA de cada una de las muestras se cuantificó por espectrofotometría en un NanoDrop 2000 UV-Vis (Thermo Scientific, USA). Se consideró que el DNA es de buena calidad cuando las lecturas en ambas absorbancias $A_{260/A280}$ y $A_{260/A230}$, estuvieron en un rango entre 1.8 y 2.0.

2.2 Preparación de las matrices DArT

El DNA se envió a la compañía Diversity Array Technology (DArT) ubicada en la Universidad de Canberra en Australia, quienes desarrollaron esta metodología para la genotipificación de poliploides. En el caso de caña de azúcar cuentan con una biblioteca de DNA compuesta con la información de 768 a 4608 clones.

2.3 Análisis de los marcadores DArT

La preparación de las representaciones genómicas individuales de los genotipos de caña de azúcar “Targets” (=objetivos) para la hibridación con las matrices DArT, se realizó los metadatos, polimorfismos SNP obtenidos en el HiSeq 2500 de la plataforma de Illumina.

Tanto el análisis de la imagen, la identificación del polimorfismo y el registro de datos, se realizaron con el programa DArTSoft especialmente desarrollado para este propósito por Diversity Arrays Technology Pty. Ltd. (WWW.DiversityArrays.com/software.html) como lo mencionó Wenzl et al. (2004) y Akbari et al. (2006).

DArTSoft analizó 52, 828 marcadores para cada muestra, registró los datos de intensidad de las hibridaciones con las matrices, identificó, anotó los polimorfismos y calculó los parámetros de calidad para cada marcador. Los resultados finales fueron analizados utilizando el programa R(R Core Team, 2016, <https://www.R-project.org/>) para determinar los datos faltantes, el mapa de calor y las distancias genéticas, así como algunos análisis de datos fenotípicos y genotípicos en forma conjunta con la finalidad de estudiar el poder predictivo de los modelos basados en marcadores moleculares.

Resultados y Discusión

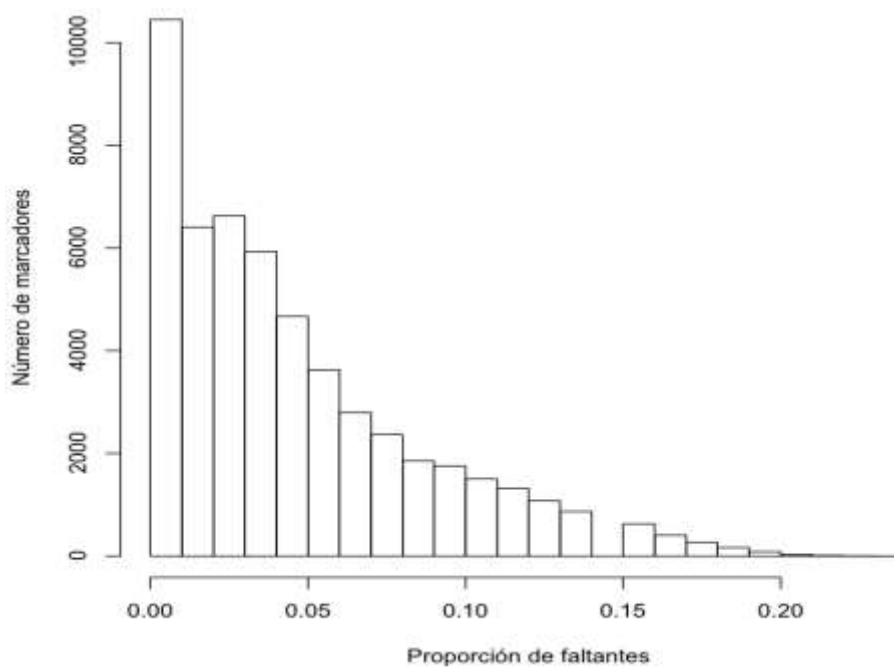
3.1 Caracterización genética

Se obtuvo información genética de 93 clones de caña de azúcar genotipificados con la tecnología DArT. Los marcadores generados son de tipo “presencia” (1) y “ausencia” (0) de un fragmento de restricción que contiene el marcador de interés. En total se obtuvieron 52, 828 marcadores para cada variedad con un porcentaje bajo de marcadores que tuvieron valores faltantes por celda.

La Fig. 2 representa la distribución de frecuencias de valores faltantes para los 52, 828 marcadores, se observa que más de 10,000 marcadores tienen una proporción de entre 0 y 0.01 de valores faltantes o sea el 10% de los marcadores que no tuvieron la lectura del arreglo en ese marcador y solamente el 20% tuvieron menos de 10 marcadores faltantes, lo cual coincide con la moda de la distribución, indicando que la información genotípica obtenida es de muy buena calidad.

Debido a que no se puede trabajar con las tasas de valores faltantes por marcador que son bastante bajas, se realizó una imputación basada en las frecuencias de los marcadores observados. Los marcadores con frecuencia del alelo menor inferiores a 0.05 se eliminaron.

Los genotipos faltantes se imputaron generando muestras aleatorias de la distribución marginal de los genotipos observados, es decir $x_{ij} \sim \text{Bernoulli}(p_j)$, donde p_j representa la frecuencia alélica calculada usando los genotipos no faltantes (Crossa-Iriarte et al., 2010).



Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.00000	0.01075	0.03226	0.04403	0.06452	0.23660

Figura 2. Distribución de valores faltantes.

La Figura 3 representa la distribución de frecuencias del alelo menor (MAF) después de la imputación, donde se observa que cerca de 10,000 marcadores tienen una $MAF < 0.05$, se eliminaron del análisis. Después de imputar y eliminar marcadores cuya MAF es menor del 0.05 se obtuvieron 43,250 marcadores para el análisis de cada variedad. Esta imputación sirve también para la eliminación de marcadores monomórficos que no aportan información relevante para el análisis.

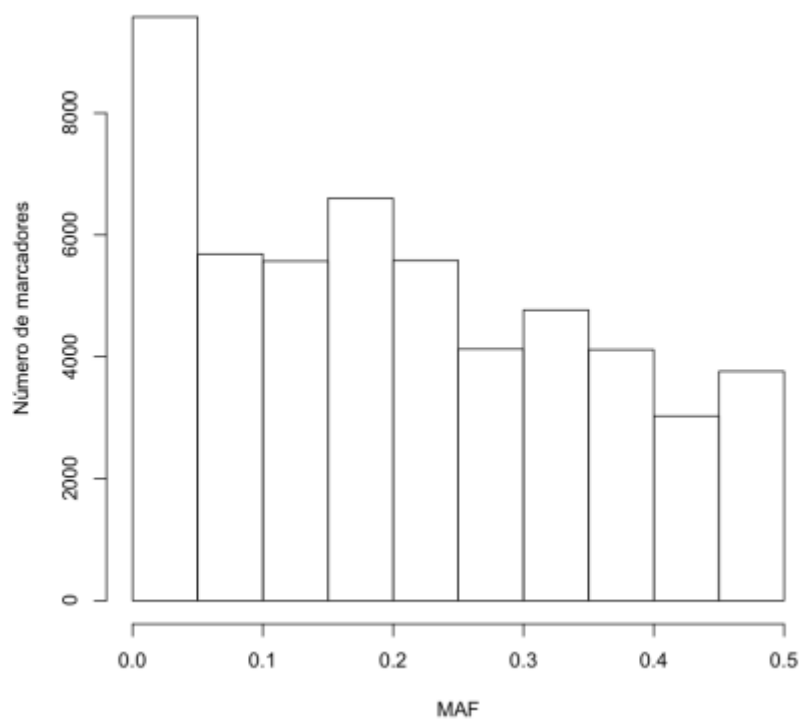


Figura 3. Distribución de frecuencias del alelo menor (MAF).

Después de imputar y eliminar marcadores en base a MAF se obtuvo la matriz de relaciones genómicas G , que es la base para los estudios de las relaciones entre un individuo con los 93 clones, misma que puede ser utilizada para estudios de la diversidad de la población de interés o en predicción genómica a partir de los progenitores. La matriz puede calcularse fácilmente usando la expresión siguiente:

$$G = ZZ^T/p, \dots(1)$$

donde Z es la matriz de marcadores de dimensiones $n=93$ individuos y $p=43,250$, que se obtiene al centrar y estandarizar las columnas de la matriz de marcadores con 0's y 1's (López-Cruz et al., 2015).

La Figura 4 presenta un mapa de calor (heatmap) de los 93 clones que se obtiene usando la matriz G , los mapas de calor son adecuados para visualizar grandes cantidades de datos multidimensionales. Cada cuadro se ilumina dependiendo del valor en la matriz. El mapa de calor se combinó con la agrupación jerárquica en clúster, que es una forma de disponer los elementos en una jerarquía según la distancia o

la similitud entre ellos. El resultado del cálculo de la agrupación jerárquica en clúster se muestra en un mapa de calor como dendrograma.

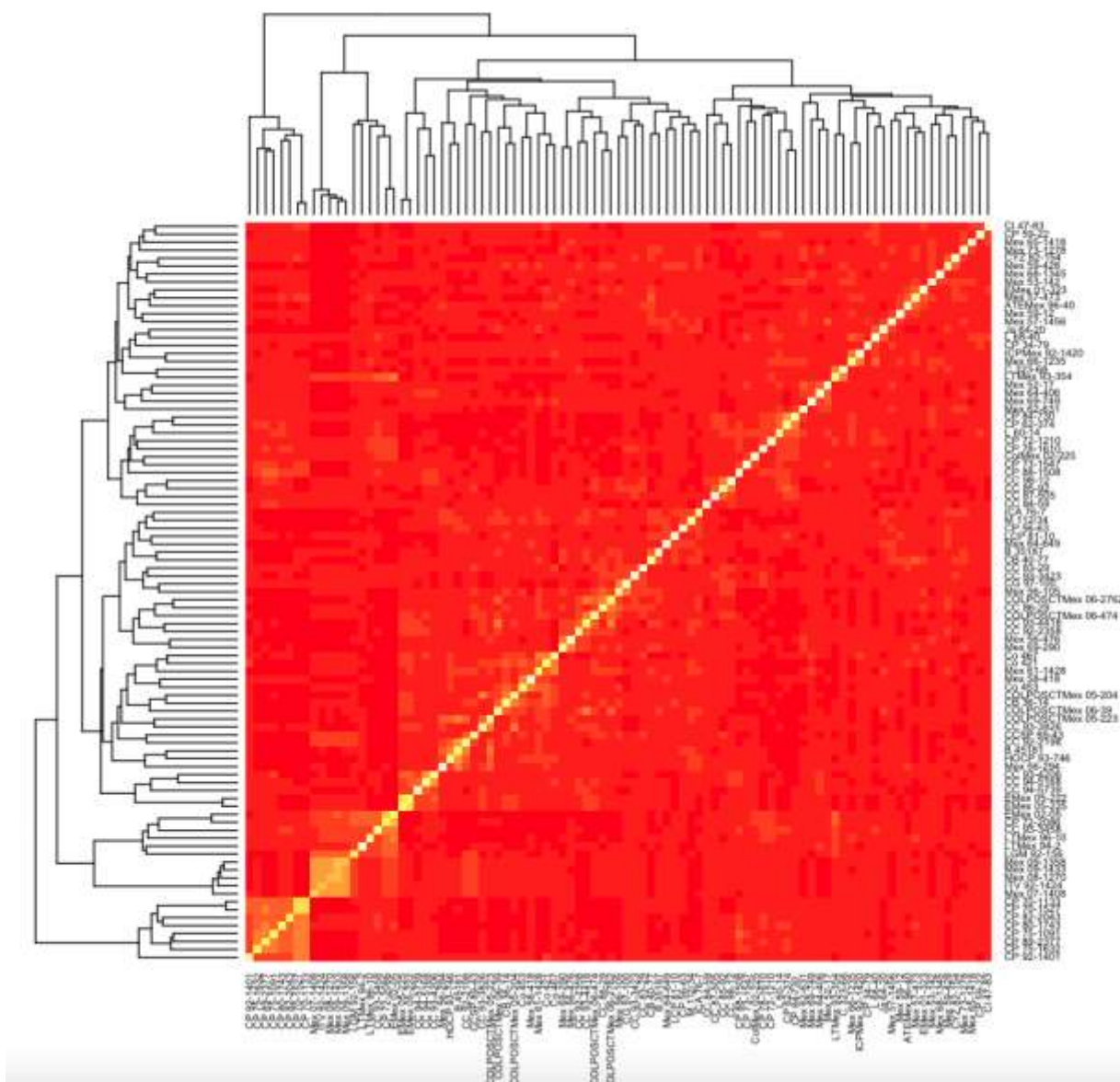


Figura 4. Mapa de calor para los 93 clones de caña de azúcar procedente de la Estación de Hibridación del CIDCA AC. analizadas mediante la técnica de Diversity Array Technology (DArT).

Basada en la matriz de marcadores codificada con 0/1 se realizó también un agrupamiento de los 93 clones. Primero se calculó una matriz de distancias Euclidianas y luego se hizo un agrupamiento jerárquico utilizando encadenamiento completo, de tal manera que el dendrograma muestra las distancias entre los individuos con base al marcador.

Los clones también se agruparon utilizando partición basada en medioides (Kaufman and Rousseuw, 2005). La Figura 5 muestra la gráfica de siluetas resultante para seis grupos. El ancho promedio de silueta mide la calidad del agrupamiento, valores cercanos a 1 indican que una variedad está bien clasificada y valores negativos que la variedad está mal clasificada.

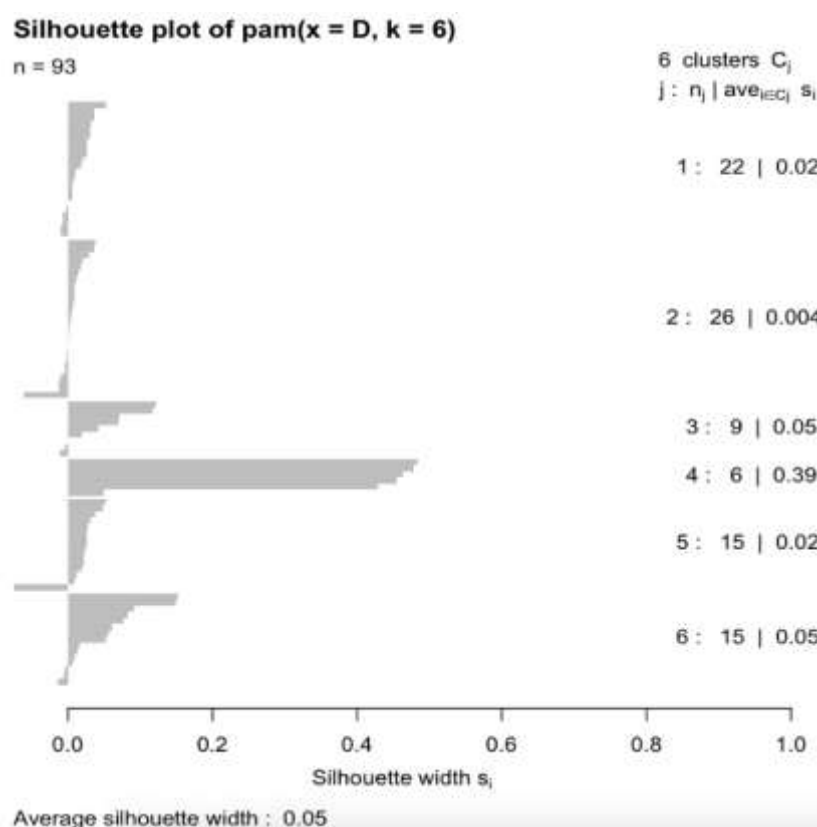


Figura 5. Diagrama de siluetas para agrupamientos resultante usando partición basada en medioides.

La Figura 6 muestra una gráfica de los dos primeros componentes principales basada en la matriz de distancias Euclidianas, muestra también óvalos que incluyen los clones pertenecientes a cada uno de los grupos, mismos que se identifican con diferentes tipos de símbolos en la gráfica. Los dos primeros componentes principales explican solo el 11% de la variabilidad total.

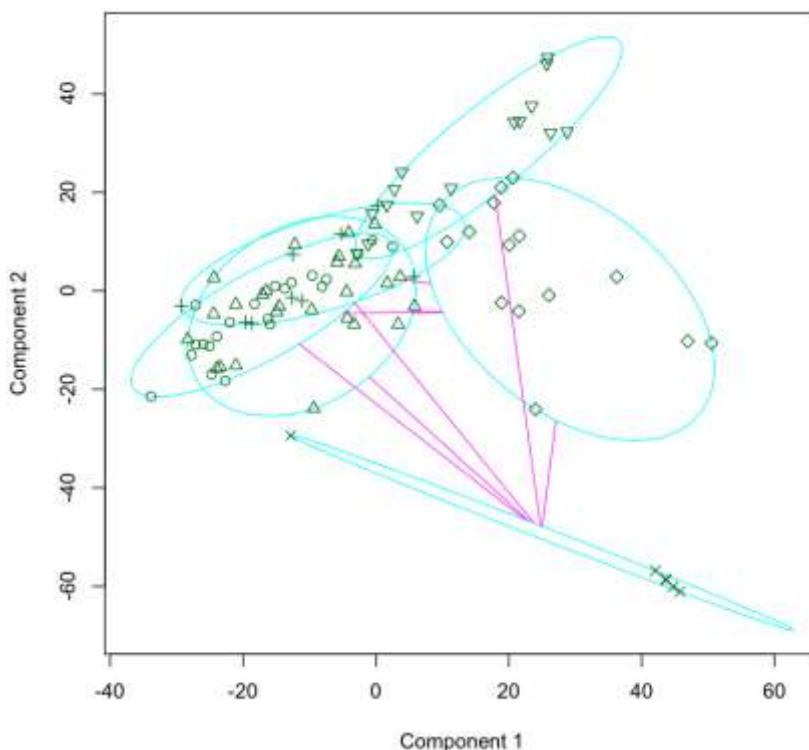


Figura 6. Componente principal 1 vs Componente principal 2 explican el 11.55% de la variabilidad, asimismo el agrupamientos para los 93 clones de caña de azúcar.

El uso de marcadores moleculares DArT desarrollado para especies poliploides como caña de azúcar, permitió realizar el análisis genético de los clones analizados. Con estas herramientas se podrán determinar las distancias genéticas de otros individuos que se encuentran en el banco de germoplasma del CIDCA AC., el conocimiento de la información genética más distante permitirán en algunos casos obtener híbridos con las mejores características agronómicas, genéticas, sanitarias y agroindustriales que puedan ser evaluados en las diferentes regiones productoras del país.

Conclusiones

El uso marcadores moleculares DArT desarrollado para especies poliploides como caña de azúcar, permitió realizar el análisis genético de los 93 clones analizados. Con estas herramientas se determinaron las distancias genéticas entre los clones en estudio, los que serán utilizados en el programa anual de cruzamientos en la Estación de Hibridación del CIDCA AC.

Referencias

Aguilar-Rivera, N.; Galindo-Mendoza, G.; Contreras-Servin, C.; Fortanelli-Martínez, J. (2010), Competitiveness and productivity of Mexico's sugar mills. *Theoria* 19, 7-30.

- Aitken, K. S.; McNeil, M. D.; Hermann, S.; Bundock, P. C.; Kilian, A.; Heller-Uszynska, K.; Henry, R. J.; Li, J. (2014), A comprehensive genetic map of sugarcane that provides enhanced map coverage and integrates high-throughput Diversity Array Technology (DArT) markers. *BMC Genomics* 15, 152. DOI: 10.1186/1471-2164-15-152.
- Akbari, M.; Wenzl, P.; Caig, V.; Carling, J.; Xia, L.; Yang, S.; Uszynski, G.; Mohler, V.; Lehmensiek, A.; Kuchel, H.; Hayden, M. J.; Howes, N.; Sharp, P.; Vaughan, P.; Rathmell, B.; Huttner, E.; Kilian, A. (2006), Diversity Arrays Technology (DArT) for high-throughput profiling of the hexaploid wheat genome. *Theor Appl Genet* 113, 1409–1420.
- Al-Khayri, J. M.; Mohan, J. S.; Johnson, D. V. (2015), *Advances in Plant Breeding Strategies: Breeding, Biotechnology and Molecular Tools*. Springer International Publishing. 645pp.
- Berkman, P. J.; Casu, R. E.; Stiller, J.; Rae, A. L.; Aitken, K. S. (2013), Towards the sugarcane genome sequence and an understanding of polyploid. *Proc. Int. Soc. Sugar Cane Technology* 28, 1-10.
- Dookun-Saumtally, A.; Saumtally, S.; Joomun, N. (2012), Biotechnological applications to sugarcane pathogens. *Functional Plant Science and Biotechnology* 6, 1-11.
- CONADESUCA. (2016), Sistema de Información de la Investigación en la Agroindustria de la Caña de Azúcar. Consultada el 02 de octubre de 2016.
<http://www.gob.mx/conadesuca/acciones-y-programas/sistema-infocana?idiom=es>
- Crossa, J.; de los Campos, G.; Pérez-Rodríguez, P.; Gianola, D.; Burgueño, J.; Araus, J. L.; Makumbi, D.; Singh, R. P.; Dreisigacker, S.; Yan, J.; Arief, V.; Banziger, M.; Braun, H.-J. (2010), Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. *Genetics* 186, 713–724.
- Doyle, J. J.; Doyle, J. L. (1987), A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull* 19, 1–15.
- Garcia, A. A.; Kido, E. A.; Meza, A. N.; Souza, H. M.; Pinto, L. R.; Pastina, M. M.; Leite, C. S.; Silva, J. A.; Ulian, E. C.; Figueira, A.; Souza, A. P. (2006), Development of an integrated genetic map of a sugarcane (*Saccharum* spp.) commercial cross, based on a maximum-likelihood approach for estimation of linkage and linkage phases. *Theor Appl Genet* 112, 298-314.
- Garcia, A. A.; Mollinari, M.; Marconi, T. G.; Serang, O. R.; Silva, R. R.; Vieira, M. L.; Vicentini, R.; Costa, E. A.; Mancini, M. C.; Garcia, M. O.; Pastina, M. M.; Gazaffi, R.; Martins, E. R.; Dahmer, N.; Sforça, D. A.; Silva, C. B.; Bundock, P.; Henry, R. J.; Souza, G. M.; van Sluys, M. A.; Landell, M. G.; Carneiro, M. S.; Vincentz, M. A.; Pinto, L. R.; Vencovsky, R.; Souza, A. P. (2013), SNP genotyping allows an in-depth characterisation of the genome of sugarcane and other complex autopolyploids. *Nature: Scientific Reports* 3, Article number: 3399. DOI: 10.1038/srep03399.
- Heller-Uszynska, K.; Uszynski, G.; Huttner, E.; Evers, M.; Carlig, J.; Caig, V.; Aitken, K.; Jackson, P.; Piperidis, G.; Cox, M.; Gilmour, R.; D'Hont, A.; Buterfield, M.; Glaszmann, J. C.; Kilian, A. (2010), Diversity arrays technology effectively reveals DNA polymorphism in a large and complex genome of sugarcane. *Mol Breed* 28, 37-55.
- Huang, E.; Aitken, K.; George, A. (2010), Association Studies. pp. 43-68. In: Henry, R. J., and Kole, C. (eds.). *Genetics, Genomics and Breeding of Sugarcane*. CRC Press. Science Publishers, USA. 300pp.
- Jaccoud, D.; Peng, K.; Feinstein, D.; Kilian, A. (2001), Diversity arrays: a solid state technology for sequence information independent genotyping. *Nucleic Acids Res* 29 (e25), 1-7.
- Jackson, P. A.; Morgan, T. E. (2003), Early stage selection for commercial cane sugar (CCS) in sugarcane clones: effects of time of sampling and irrigation. *Crop and Pasture Science* 54, 389-396.

- Kaufman, L.; Rousseeuw, P. J. (2005), *Finding Groups in Data: An Introduction to Cluster Analysis*. Wiley, New York. 342pp.
- López-Cruz, M.; Crossa-Iriarte, J.; Bonnet, D.; Dreisigacker, S.; Poland, J.; Jannink, L-L.; Singh, R. P.; Autrique, E.; de los Campos, G. (2015), Increased prediction accuracy in wheat breeding trials using a marker \times environment interaction genomic selection model. *G3:Genes|Genome|Genetics* doi:10.1534/g3.114.016097.
- Manners, J. M. (2011), Functional Genomics of Sugarcane. pp. 91-146. In: Jean-Claude Kader and M. Delseny (eds.). *Advances in Botanical Research*. Volume 60. Academic Press, USA. 506pp.
- Mendes de Paula, T. O.; Marinho, C. D.; Souza, V.; Barbosa, M. H.; Peternelli, L. A.; Kimbeng, C. A.; Zhou, M. M. (2014), Relationships between methods of variety adaptability and stability in sugarcane. *Genetic and Molecular Research* 13, 4216-25.
- Premachandran, M. N.; Prathima, P. T.; Lekshmi, M. (2011), Sugarcane and polyploidy, a review. *J Sugarcane Res* 1, 1-15.
- R Core Team. (2016), *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Senties-Herrera, H.; Gómez- Merino, F. C. (2014), Nuevas directrices en mejoramiento genético de caña de azúcar. *Colegio de Postgraduados Campus Córdoba, México. Agroproductividad* 7, 9-15.
- Wenzl, P.; Carling, J.; Kudrna, D.; Jaccoud, D.; Huttner, E.; Kleinhofs, A.; Kilian, A. (2004), Diversity Arrays Technology (DArT) for whole-genome profiling of barley. *Proc Natl Acad Sci USA* 101, 9915–9920.