



Journal of Basic & Applied Genetics

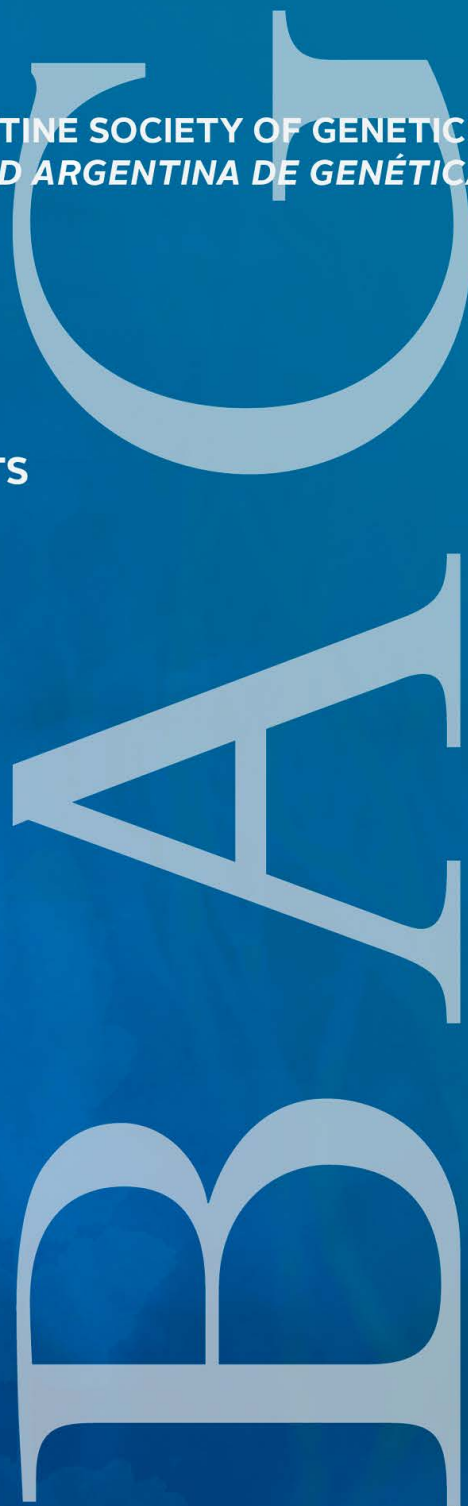
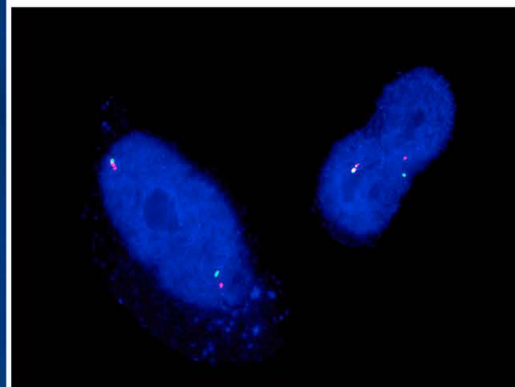
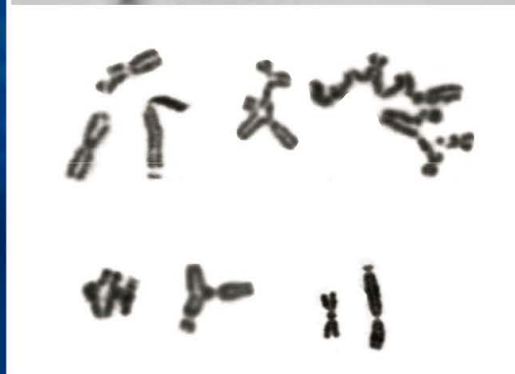
(Formerly MENDELIANA)

JOURNAL OF THE ARGENTINE SOCIETY OF GENETICS
REVISTA DE LA SOCIEDAD ARGENTINA DE GENÉTICA



Cited by
BIOLOGICAL ABSTRACTS
GENETICS ABSTRACTS
SISTEMA LATINDEX
THOMSON REUTERS
SCOPUS

Included in **SciELO**



COMITÉ EDITORIAL

Editor General:

Dra. Elsa L. Camadro

Área de Investigación en Agronomía
EEA Balcarce, Instituto Nacional de Tecnología
Agropecuaria (INTA), Fac. de Ciencias Agrarias,
Universidad Nacional de Mar del Plata (UNMDP, y
Consejo Nacional de Investigaciones Científicas y
Técnicas (CONICET)
Balcarce, Argentina
camadro.elsa@inta.gob.ar

Editores Asociados:

Citogenética Animal

Dra. Liliana M. Mola

Departamento de Ecología, Genética y Evolución,
Fac. de Ciencias Exactas y Naturales, Universidad de
Buenos Aires (UBA), y CONICET
Buenos Aires, Argentina
limola@ege.fcen.uba.ar; lilimola@yahoo.com.ar

Citogenética Humana

Dra. Silvia Ávila

Universidad Nacional de COMAHUE
Hospital Castro Rendón
Neuquén, Argentina
silvia347@gmail.com

Citogenética Vegetal

Dra. Liliana M. Mola

Departamento de Ecología, Genética y Evolución,
Fac. de Ciencias Exactas y Naturales, UBA,
y CONICET
Buenos Aires, Argentina
limola@ege.fcen.uba.ar; lilimola@yahoo.com.ar

Dr. José Guillermo Seijo

Instituto de Botánica del Nordeste,
Universidad Nacional del Nordeste (UNNE),
y CONICET
Corrientes, Argentina
seijo@agr.unne.edu.ar

Genética de Poblaciones y Evolución

Dr. Jorge Cladera

Instituto de Genética "Ewald Favret",
Centro de Investigación en Ciencias. Veterinarias y
Agronómicas,
INTA
Castelar, Argentina
cladera.jorge@inta.gob.ar

Dra. Noemí Gardenal

Fac de Ciencias Exactas, Físicas y Naturales,
Universidad Nacional de Córdoba (UNC),
y CONICET
Córdoba, Argentina
ngardenal@unc.edu.ar

Dr. Juan César Vilardi

Departamento de Ecología, Genética y Evolución,
Fac. Ciencias Exactas y Naturales, UBA,
y CONICET
Buenos Aires, Argentina
vilardi@bg.fcen.uba.ar

Genética Humana y Genética Médica

Dr. Santiago Lippold

Centro de Educación Médica e Investigaciones Clínicas
(CEMIC)
Buenos Aires, Argentina
sel1@fibertel.com.ar

Genética Médica, Humana y Citogenética

Dra. María Inés Echeverría

Instituto de Genética, Facultad de Ciencias Médicas,
Universidad Nacional de Cuyo (UNCu)
Mendoza, Argentina
miecheve@fcm.uncu.edu.ar

Dra. Silvia Ávila

Universidad Nacional de COMAHUE
Hospital Castro Rendón
Neuquén, Argentina
silvia347@gmail.com

Genética Molecular (Animal)

Dr. Guillermo Giovambattista

Instituto de Genética Veterinaria (IGEVET), Fac. de
Ciencias Veterinarias,
Universidad Nacional de La Plata (UNLP) y CONICET
La Plata, Argentina
ggiovam@fcv.unlp.edu.ar

Genética Molecular (Vegetal)

Dr. Alberto Acevedo

Centro de Investigación de Recursos Naturales,
INTA
Castelar, Argentina
acevedo.alberto@inta.gob.ar

Dr. Andrés Zambelli

Unidad de Negocios Nutrisun–Advanta Semillas SAIC
Balcarce, Argentina
andres.zambelli@advantaseeds.com

Genética y Mejoramiento Animal

Ing. (M. Sc.) Carlos A. Mezzadra

Área de Investigación en Producción Animal,
EEA Balcarce, INTA y
Fac. de Ciencias Agrarias, UNMDP
Balcarce, Argentina
mezzadra,carlos@inta.gob.ar

FOTOGRAFÍAS Y AUTORES

Dra. Liliana A. Picardi

*Cátedra de Genética, Fac. de Ciencias Agrarias,
Universidad Nacional de Rosario (UNR)
Zavalla, Argentina
lpicardi@fcagr.unr.edu.ar*

Genética y Mejoramiento Genético Vegetal

Dr. Miguel A. Di Renzo

*Facultad de Agronomía y Veterinaria,
Universidad Nacional de Río Cuarto (UNRC)
Córdoba, Argentina
mdirenzo@ayv.unrc.edu.ar*

Dr. Ricardo W. Masuelli

*EEA La Consulta, INTA
Fac. de Ciencias Agrarias, Universidad Nacional de Cuyo
(UNCu), y CONICET,
Mendoza, Argentina
rmasuelli@fca.uncu.edu.ar*

Dra. Mónica Poverene

*Departamento de Agronomía, Universidad Nacional del
Sur (UNS), y CONICET
Bahía Blanca, Argentina
poverene@criba.edu.ar*

Mutagénesis

Dr. Alejandro D. Bolzán

*Laboratorio de Citogenética y Mutagénesis,
Instituto Multidisciplinario de Biología Celular (IMBICE),
y CONICET
La Plata, Argentina
abolzan@imbice.gov.ar; abolzan64@gmail.com*

Mutaciones Inducidas en Mejoramiento Vegetal

Ing. Agr. (M.Sc.) Alberto R. Prina

*Instituto de Genética "Ewald Favret", Centro de
Investigación en Ciencias Veterinarias y Agronómicas,
INTA
Castelar, Argentina
prina.alberto@inta.gob.ar*

Consultor Estadístico:

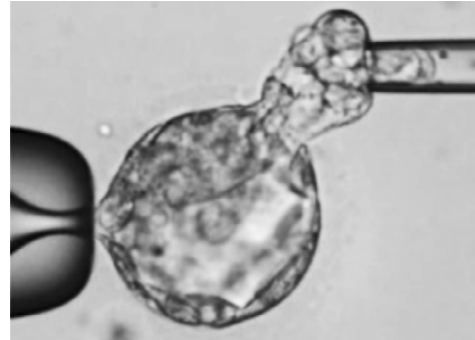
Ing. Agr. Francisco J. Babinec

*EEA Anguil, INTA, y
Fac. de Agronomía, Universidad Nacional de La Pampa
(UNLPam)
La Pampa, Argentina
babinec.francisco@inta.gob.ar*

Secretaría de Redacción:

Dra. María de las Mercedes Echeverría

*Fac. de Ciencias Agrarias, UNMdP
Balcarce, Argentina
echeverria.maria@inta.gob.ar*



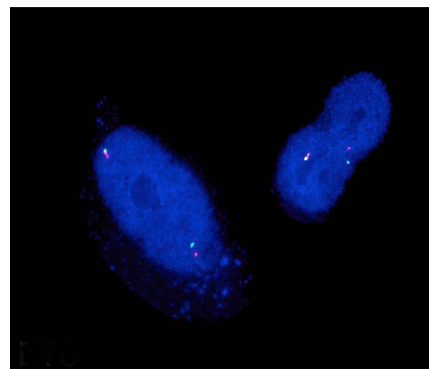
Inyección Intracitoplasmática del Espermatozoide (ICSI)
en humanos

Autor: M. E. Ducatelli



Aberraciones cromosómicas estructurales inducidas por
drogas antitumorales en fibroblastos humanos

Autor: M. De Campos-Nebel *et al.*



Núcleos de fibroblastos humanos con rearrreglos del
gen MLL inducidos por drogas antitumorales

Autor: M. González Cid

ÍNDICE

7 - 15

Article 1 - research

MECHANISMS FOR THE EPIGENETIC REGULATION OF MITOCHONDRIAL DNA
MECANISMOS DE REGULACIÓN EPIGENÉTICA DEL ADN MITOCONDRIAL

Sierra-Delgado J.A., Contreras-García G.A.

17 - 24

Article 2 - research

PREIMPLANTATIONAL GENETIC DIAGNOSIS (PGDs) IN CARRIERS OF BALANCED STRUCTURAL REARRANGEMENTS BY ARRAY COMPARATIVE GENOMIC HYBRIDIZATION (aCGH)

DIAGNÓSTICO GENÉTICO PREIMPLANTATORIO (DGP) EN PORTADORES DE REARREGLOS ESTRUCTURALES BALANCEADOS POR aCGH (ARRAY COMPARATIVE GENOMIC HYBRIDIZATION)

Ducatelli M.E., A. Grazziotin Mondadori, F. Coco, N. Neuspiller, F.L. Gismondi, R. Coco.

25 - 35

Article 3 - research

PERSISTENT GENETIC DAMAGE INDUCED BY TOPOISOMERASE II POISONS IN NORMAL HUMAN FIBROBLASTS: GENERATION OF CHROMOSOME INSTABILITY

DAÑO GENÉTICO PERSISTENTE INDUCIDO POR VENENOS DE TOPOISOMERASA II EN FIBROBLASTOS HUMANOS NORMALES: GENERACIÓN DE INESTABILIDAD CROMOSÓMICA

de Campos-Nebel M., Palmitelli M., Acevedo S., González-Cid M.

37 - 48

Article 4 - research

ANCESTRAL GENETIC COMPOSITION OF A HUMAN POPULATION FROM THE COLOMBIAN SOUTHWEST USING AUTOSOMAL AIM-INDELS
COMPOSICIÓN GENÉTICA ANCESTRAL DE UNA POBLACIÓN HUMANA DEL SUDOESTE DE COLOMBIA USANDO AIM-INDELS AUTOSÓMICOS

Urbano L., Portilla E.C., Builes J.J., Gusmão L., Sierra-Torres C.H.

49 - 58

Article 5 - research

STATISTICAL MODELS FOR PHENOTYPE-GENOTYPE ASSOCIATION STUDIES IN GENETICALLY STRUCTURED POPULATIONS
MODELOS ESTADÍSTICOS PARA ESTUDIOS DE ASOCIACIÓN FENOTIPO-GENOTIPO EN POBLACIONES GENÉTICAMENTE ESTRUCTURADAS

Peña Malavera A., Gutierrez L., Balzarini M.

59 - 60

Obituary

TIRANTI, IVÁN NICOLÁS

MECANISMOS DE REGULACIÓN EPIGENÉTICA DEL ADN MITOCONDRIAL

MECHANISMS FOR THE EPIGENETIC REGULATION OF MITOCHONDRIAL DNA

Sierra-Delgado J.A.^{1,2*}, Contreras-García G.A.³

¹ Médica, Candidata a Máster en Ciencias Básicas Biomédicas, Grupo de Genética Humana, Universidad Industrial de Santander, Bucaramanga, Santander, Colombia.

² Médica, Grupo de Investigación en Medicina Genómica y Metabolismo FCV, Fundación Cardiovascular de Colombia, Bucaramanga, Santander, Colombia.

³ Médico Genetista, especialista en Bioética, Grupo de Genética Humana, Universidad Industrial de Santander, Departamento de Pediatría, Hospital Universitario de Santander, Bucaramanga, Santander, Colombia.

*Corresponding author: andreasirrad@outlook.com

ABSTRACT

The mitochondrion is one of the most important cellular organelles for the right functioning of the eukaryotic cell. It plays a fundamental role in traditional epigenetics, because it takes part in one-carbon metabolism and the folate cycle. However, in mammals, the epigenetic mechanisms regulating the synthesis and expression of mitochondrial DNA (mtDNA), a field of study called mitoepigenetics, are understudied. The mtDNA is a circular DNA molecule of 16.6 kb without introns, which codes for 14 proteins, two rRNA and 22 tRNA. The mtDNA is packaged in nucleoids, composed of histone analogs called TFAM (Transcription factor A of Mitochondria). In mammals, evidence of epigenetic regulation of mtDNA has been known since 1984; however, the finding remained controversial until 2011, when the translocation of nuclear DNMT1 to the mitochondria was proved. Since then, there have been reports of other epigenetic mechanisms, which include the presence of 5-methylcytosines in CpG and Non-CpG pattern, 5-hidroximetilcytosines (5HmC), nucleoids remodeling and interference RNA. Similarly, these regulation patterns have been related to aberrant levels of transcription of mitochondrial genes. Association of this aberrant transcription with pathologies like cancer, NASH (not-alcoholic steatopathy), neurodegenerative diseases and diabetic retinopathy, but also with environmental changes and even ageing, have been described. According to these results and the recent advances in the field, mitoepigenetics should be considered as a fascinating and growing field of study, because it might contribute to the understanding of the pathophysiology of complex diseases .

Key words: Epigenetics, Genetic, DNA methylation, mitochondrial DNA, mitochondria.

RESUMEN

La mitocondria es uno de los organelos celulares más importantes para el correcto funcionamiento de la maquinaria celular eucariota. Dentro de la epigenética tradicional cumple un papel primordial al formar parte del metabolismo de un carbono y el ciclo del folato; sin embargo, en mamíferos existen mecanismos epigenéticos específicos poco estudiados, que regulan la síntesis y expresión del ADN mitocondrial (ADNmt), campo denominado "mitoepigenética". El ADNmt es un ADN circular de 16.6 kb, sin intrones que codifica 14 proteínas, 2 ARNr y 22 ARNt. Se encuentra empaquetado en nucleoides, compuestos por proteínas análogas a las histonas llamadas factor de transcripción mitocondrial A (TFAM). Existe evidencia de regulación epigenética del ADNmt en mamíferos desde 1984, sin embargo, su hallazgo permaneció controversial hasta el 2011, cuando se demostró la translocación de DNMT1 desde el núcleo hacia la mitocondria. Desde entonces se han descrito otros mecanismos entre los que se encuentran la presencia de 5-metilcitosinas (5mC) en patrón CpG y no-CpG, 5-hidroximetilcitosinas (5hmC), la remodelación de nucleoides y el ARN de interferencia. De igual manera estos patrones de regulación se han correlacionado con niveles aberrantes en la transcripción de genes mitocondriales. Se ha encontrado asociación de estos cambios con entidades patológicas, como cáncer, NASH (esteatopatía no alcohólica), enfermedades neurológicas y retinopatía diabética, así como cambios ambientales e incluso envejecimiento. A partir de estos resultados y los avances recientes, se podría considerar la mitoepigenética como un campo de estudio fascinante y creciente, que puede contribuir al entendimiento de la fisiopatología de enfermedades complejas.

Palabras clave: Epigénesis Genética, Metilación de ADN, ADN mitocondrial, mitocondria.

INTRODUCCIÓN

La mitocondria es, indudablemente, uno de los organelos celulares más importantes para el correcto funcionamiento de la maquinaria celular eucariota. Es la encargada de la producción aeróbica de ATP por medio de un proceso denominado fosforilación oxidativa, el cual es mediado por complejos enzimáticos multiproteicos en su membrana, encargados de transporte en cadena de electrones (Saraste, 1999). De igual manera, la mitocondria juega un papel fundamental en la apoptosis, la generación de especies reactivas de oxígeno y cuenta con funciones tejidas específicas, como el metabolismo del colesterol, la biosíntesis del grupo hemo y la detoxificación de derivados del amonio en tejido hepático, entre otros (Skorecki *et al.*, 2012). Sin embargo, a pesar de su rol importante en el mantenimiento de la homeostasis y el metabolismo, se conoce poco de la relación entre la mitocondria y los mecanismos epigenéticos específicos que regulan su síntesis y funcionamiento en mamíferos, un campo novedoso conocido como “mitoepigenética”.

Dentro de la epigenética tradicional la mitocondria cumple un papel primordial (van der Wijst y Rots, 2015). La mitocondria participa en la regulación del epigenoma al formar parte del metabolismo de un carbono a través de la producción de ATP y folato (Taylor y Tumbull, 2005; Locasale, 2013). Las mitocondrias regulan el cambio entre S-AdenosilMetionina (SAM), un donante de grupos metilo, y la síntesis de nucleótidos por la acción de la enzima mitocondrial bifuncional, el cual es un paso fundamental en la metilación del ADN (Iacobazzi *et al.*, 2013).

ADN MITOCONDRIAL: MECANISMOS MOLECULARES

La mitocondria, junto al ribosoma, son los únicos organelos en la célula animal cuya composición incluye ácidos nucleicos. El ADN mitocondrial (ADNmt) fue descubierto en 1963 por Margit M.K. Nass- Edelson y Sylvan Nass mediante microscopía de electrones (Nass y Nass, 1963), sin embargo, su presencia se intuía aún antes debido a la similitud ultraestructural entre la mitocondria y las bacterias. Por este motivo se propuso la teoría de la endosimbiosis, en la cual se formula que durante el proceso evolutivo células bacterianas fueron engullidas y modificadas para convertirse en organelos eucariotas, como la mitocondria (Manev *et al.*, 2012).

El ADNmt es un ADN circular de 16,6 kb, el cual representa menos de 1% del ADN total de una célula y es transmitido casi exclusivamente por linaje materno. Su composición consta de 2 cadenas, 1 ligera y 1 pesada (Figura 1), no contiene intrones y codifica 13 proteínas en total, todas parte de la cadena de transporte de electrones. El ADNmt también contiene el bucle de desplazamiento (*D-Loop*), una región no codificante en el ADNmt que contiene elementos que regulan la replicación y la transcripción *in cis*. En el *D-Loop* también se ubican las regiones hipervariables 1 y 2 (HVR1 y HVR2), las cuales son los sitios de mayor acúmulo de mutaciones dentro de la secuencia (Gosh *et al.*, 2015). El genoma mitocondrial también transcribe 2 ARNr y 22 ARNt necesarios para la traducción del ADNmt (Minocherhomji *et al.*, 2012). Otros estudios han encontrado un 14vo. polipéptido de 21 a 24 aminoácidos, llamado humanina, factor protector para el desarrollo de la enfermedad de Alzheimer, el cual es codificado por el gen MT-RNR2 (Maximov *et al.*, 2002). El resto de elementos necesarios para el correcto funcionamiento de la mitocondria son codificados nuclearmente y son transportados por complejos proteicos dentro y fuera de la membrana mitocondrial.

El ADNmt no contiene histonas, sin embargo, se encuentra recubierto por proteínas y empaquetado en agregados llamados nucleoides o “mitocromosoma”. El factor de transcripción mitocondrial A (TFAM) es el componente más abundante del nucleoide, es análogo a las histonas y juega un rol importante en el empaquetamiento y compactación del ADNmt. Los datos indican que la mayoría de los nucleoides tienen sólo una copia de ADNmt (Manev *et al.*, 2012).

La replicación del ADNmt es mediada de manera exclusiva por la polimerasa γ , la cual cumple doble función: y reparar, por lo que existe una disminución en la fidelidad por una baja corrección de errores de la replicación en comparación con el genoma nuclear (Skorecki *et al.*, 2012). Sin embargo, debido al alto número de copias de ADNmt (hasta 1.000 copias por célula), estas mutaciones no conllevan necesariamente efectos fenotípicos, sino que aumentan el nivel de heteroplasmia de la célula, definida como la divergencia entre las copias de ADNmt dentro la misma; siempre que no se supere un umbral de heteroplasmia, diferente para cada tejido y tipo de célula, no va a existir enfermedad o alteración, lo que se conoce como “mitocheckpoint” (Minocherhomji *et al.*, 2012). Existe un mayor nivel de heteroplasmia en células post-mitóticas que en los tejidos que se dividen rápidamente (Manev *et al.*, 2012).

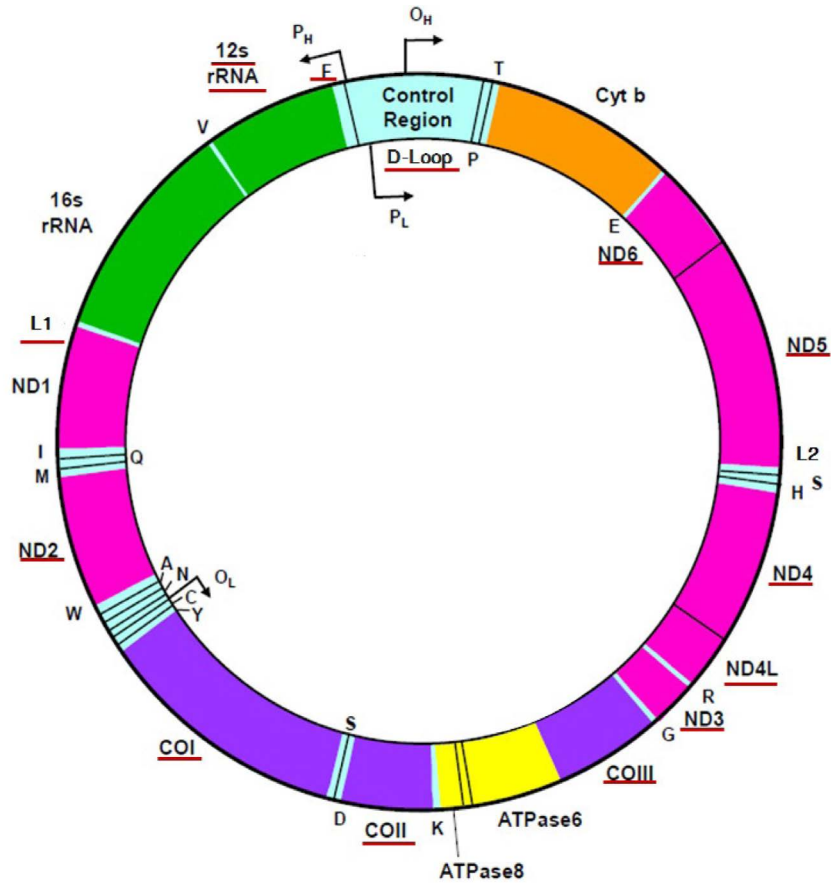


Figura 1. Diagrama del mtDNA. Se representan la posición de los 37 genes, la región de control (D-Loop) y los orígenes de replicación de las cadenas pesada (O_H , P_H) y ligera (O_L , P_L). Genes que han sido reportados con metilación o transcripción diferencial se encuentran subrayados. Modificado de www.mitomap.org (Lott *et al.*, 2013).

MITOPIGENÉTICA Y REGULACIÓN

En 1984, Pollack y su grupo publicaron reportes que indicaban la presencia de 5-Metilcitosinas (5mC) en ADNmt en tejido murino (Pollack *et al.*, 1984). El equipo de Pollack observó que, en comparación con el ADN no codificante (ADNnc), el ADNmt se encontraba generalmente submetilado, con sólo un 1-5 % de las citosinas metiladas. Sin embargo, la presencia de metilación en el ADN mitocondrial fue controvertida en la comunidad científica (Manev *et al.*, 2012; Chinnery *et al.*, 2012), hasta que los experimentos de Shock *et al.* (2011) demostraron la translocación de DNMT1 (ADN Metiltransferasa-1) nuclear a la matriz mitocondrial (mtDNMT1) en fibroblastos murinos. En condiciones normales la DNMT1 es la encargada del mantenimiento de los patrones de metilación en el ADN nuclear en mamíferos.

La mtDNMT1 es una isoforma de la enzima que contiene tanto el dominio catalítico de la DNMT1 como la secuencia diana mitocondrial. La migración de la DNMT1 hacia la mitocondria es un proceso regulado mediante la expresión de una secuencia diana conservada corriente arriba del sitio de inicio de transcripción dentro del gen codificado nuclearmente. Estudios posteriores han encontrado que la migración de mtDNMT1 está regulada positivamente por factores que responden al estrés oxidativo, como el PGC1 α (Coactivador α del receptor activado por proliferadores peroxisomales γ), el factor nuclear respiratorio 1 (NRF1), y la liberación de p53 a partir del promotor DNMT1 (Balinang, 2012). Bajo condiciones basales, la transcripción de mtDNMT1 constituye alrededor del 2 % del total de la DNMT1 (Shock *et al.*, 2011).

Adicional a la presencia de mtDNMT1, estudios posteriores en fracciones mitocondriales provenientes de cultivos murinos y sistema nervioso central humano detectaron la presencia de DNMT3a, pero bajos niveles de mtDNMT1, lo cual podría sugerir que la localización mitocondrial de DNMT3a es tejido-específica (Chestnut *et al.*, 2011; Wong *et al.*, 2013). Cabe recordar que la DNMT3, en el núcleo, es la encargada de regular los patrones de metilación de novo (Shock *et al.*, 2011).

En un estudio posterior, Bellizzi *et al.* (2013) reportaron citosinas metiladas e hidroximetiladas en el *D-Loop* de las muestras analizadas, encontrando que la metilación en el ADNmt se da principalmente en un patrón no-CpG en la región promotora de la cadena pesada y en bloques de secuencias conservadas, lo que sugiere su implicación en la regulación de la replicación y/o transcripción del ADNmt. El inactivar las DNMTs 1, 3a y 3b en células madre embrionarias de ratón resulta en una reducción de la metilación CpG, pero no la metilación no-CpG, lo cual sugiere que la modificación epigenética del *D-loop* sólo se establece parcialmente por estas enzimas (Bellizzi *et al.*, 2013).

Ghosh *et al.* (2014) analizaron 39 conjuntos de datos presentes en bases de datos públicas a fin de caracterizar el metiloma mitocondrial, encontrando que el patrón y la distribución de 5mC son relativamente constantes a través de la mitocondria, con la excepción de unos pocos loci metilados diferencialmente en diversos tejidos durante diferentes puntos en el tiempo; coincidentemente, estos loci han sido relacionados con enfermedades mitocondriales (Tabla 1). Estos resultados muestran que es plausible que la variabilidad en la metilación puede influir potencialmente en la regulación génica, con la consiguiente diversidad fenotípica.

OTROS MECANISMOS DE REGULACIÓN

La presencia de 5-hidroximetilcitosina (5hmC) ha sido confirmada en ADNmt de mamíferos (Shock *et al.*, 2011), en una densidad significativamente mayor en el ADNmt que en ADNnc (Sun *et al.*, 2013). Estudios realizados por Dzitoyeva *et al.* (2012) encontraron presencia de proteínas TET (TET1 y TET2) por inmunoreactividad en fracciones mitocondriales de cultivos neuronales, sin embargo, no se ha confirmado de manera concluyente la actividad de dichas proteínas en la mitocondria (Sadakierska-Chudy *et al.*, 2014; Gosh *et al.*, 2015). La expresión

coordinada de genes nucleares y mitocondriales parece ser particularmente importante durante la división celular y la biogénesis mitocondrial, los estudios sugieren que existe una correlación entre la regulación transcripcional de ADNmt mediada por 5hmC, 5mC y genes codificados en el núcleo (Minocherhomji *et al.*, 2012).

Sin embargo, no sólo la metilación tiene un componente importante en la regulación epigenética del ADNmt, las alteraciones en el contenido de TFAM determinan un cambio en el mitocromosoma y exponen al ADN a la acción de DNMTs. Se ha encontrado que las DNMTs tienen accesibilidad a diferentes sitios en ADNmt en función del nivel de ocupación de la proteína, y que existe correlación entre el nivel de empaquetamiento del nucleóide con el porcentaje de metilación e incluso con la presencia de ARN de interferencia en mitocondria (Rebelo *et al.*, 2009 ; Chinnery *et al.*, 2012). De manera curiosa, se ha descrito una regulación negativa del TFAM en amniocitos cultivados de pacientes con trisomía 21 en comparación con fetos normales (Infantino *et al.*, 2011).

En ratas, un pequeño grupo de micro ARN (ARNmi) ha sido identificado en derivados altamente purificados de mitocondrias hepáticas; estudios sugieren que estos regulan la expresión de genes relacionados con la apoptosis, la proliferación celular y la diferenciación celular (Kalani *et al.*, 2014). Evidencia de esta regulación por ARNmi se encontró al hallar la presencia de argonauta 2 en extractos mitocondriales (Sadakierska-Chudy *et al.*, 2014). Das *et al.* (2012) también hallaron evidencia de regulación mitocondrial del ADNmt a través de ARNmi en tejido cardíaco. Un diagrama resumiendo los diversos mecanismos de control epigenético en el ADNmt se puede encontrar en la Figura 2.

ROL DEL MITOEPÍGENOMA EN PATOLOGÍAS

Además de la diferencia en la presencia de TFAM en amniocitos con trisomía 21, otras modificaciones epigenéticas han sido correlacionadas o ligadas a entidades clínicas; por ejemplo, se encontró un aumento de niveles de DNMT1 y DNMT3A en mitocondrias de las motoneuronas en pacientes con esclerosis lateral amiotrófica respecto a controles sanos, lo cual sugiere que existen mecanismos que pueden contribuir a la muerte celular mediada por metilación del ADNmt (Chestnut *et al.*, 2011).

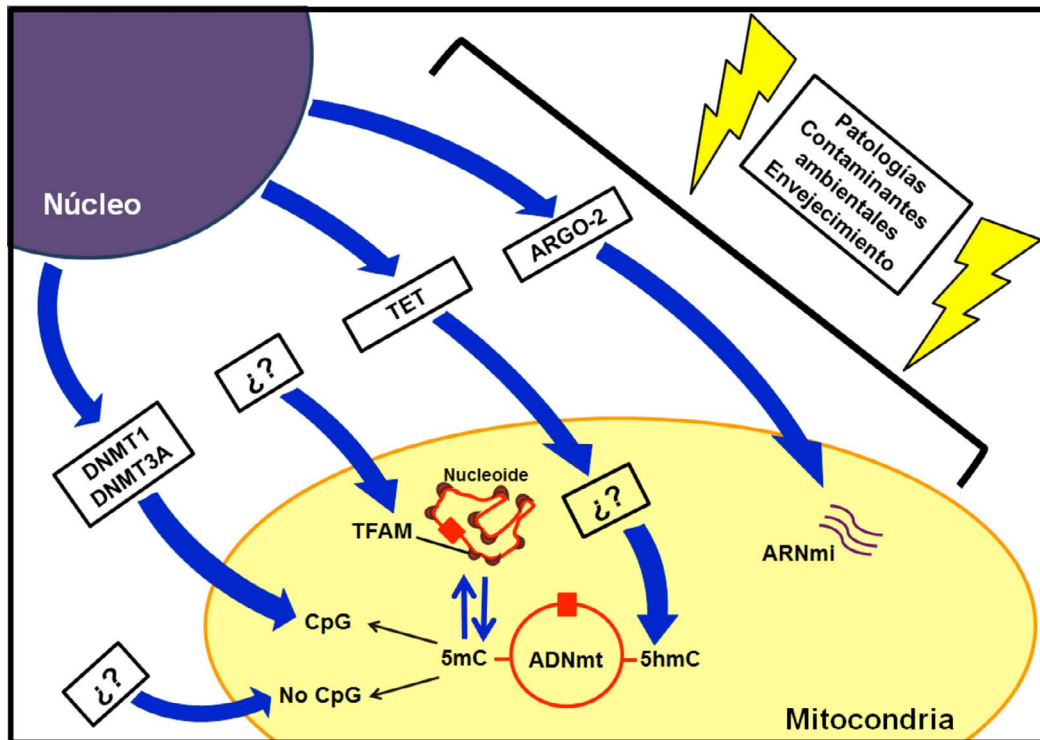


Figura 2. Un modelo que resume los principales mecanismos epigenéticos que regulan la expresión del ADN mitocondrial. El ADN nuclear codifica proteínas que involucran modificaciones epigenéticas dentro del núcleo y la mitocondria. Entre los mecanismos de regulación se encuentran en primer lugar, la metilación del ADNmt, la cual, en patrón CpG está mediado por las DNMTs 1 y 3a, las cuales son importadas desde el núcleo. Se desconocen los mecanismos involucrados en la metilación No CpG. Otros mecanismos incluyen la hidroximetilación, mediada por proteínas TET cuya actividad intramitocondrial no se ha podido confirmar de manera concluyente a pesar de encontrarse en fracciones mitocondriales, la remodelación de los nucleoides, regulada por la metilación de loci mitocondriales y cambios en TFAM y por último la regulación por medio de ARNmi, producidos por argonata 2 mitocondrial. Estudios han ligado estos cambios a desencadenantes como patologías, contaminantes ambientales y envejecimiento. TFAM: factor de transcripción mitocondrial A, ADNmt: ADN mitocondrial, 5hmC: 5-hidroximetilcitosina, 5mC: 5-metilcitosina, ARGO-2: argonata 2 mitocondrial, DNMT1: ADN Metiltransferasa-1, DNMT3A: ADN Metiltransferasa-3A.

Asimismo, se ha demostrado que la hipermetilación en sitios de unión de la polimerasa γ en el genoma mitocondrial durante estados hiperglucémicos podría potencialmente modular la memoria epigenética en retinopatía diabética (Ghosh *et al.*, 2014; Kowluru *et al.*, 2015). En estudios realizados por Feng y colaboradores (2012) también se ha observado una correlación directa entre la expresión de MT-ND2 mitocondrial y la demetilación de su secuencia de promotor en el ADNmt (*D-Loop*) en pacientes con cáncer colorrectal, encontrando que, en los fenotipos neoplásicos, la expresión de ND2 se aumentaba mientras que la metilación disminuía. Esta diferencia fue significativa incluso en estadios tempranos de cáncer (Feng *et al.*, 2012).

La evidencia sugiere que no sólo las enfermedades sino también factores como la exposición ambiental o el envejecimiento pueden afectar los niveles de metilación. El envejecimiento afecta los niveles de 5-hidroximetilcitosina en el ADNmt en la corteza frontal del cerebro en ratones y disminuye los niveles de transcripción de proteínas relacionadas con el complejo I de la fosforilación oxidativa (Dzitoyeva *et al.*, 2012). La exposición de cultivos celulares a estrés oxidativo derivado de la contaminación ambiental también ha mostrado una alteración en la metilación de los genes ND6 y 12SrRNA (Giordano *et al.*, 2012; Iacobazzi *et al.*, 2013) y el estrés oxidativo secundario a la exposición del etanol *in vitro* conlleva a un aumento en el nivel de

proteína mtDNMT1 seguido por el aumento del contenido de 5mC y 5hmC en el ADNmt (Balinang, 2012). *In vivo*, estudios en pacientes expuestos a material particulado rico en metales muestran una metilación selectiva en los genes MT-TF y MT-RNR1 con respecto a controles con baja exposición (Byun, 2013; Byun y Baccarelli, 2014).

De manera interesante, el estudio ENVIRONAGE (*Environmental influence on early ageing*) (Janssen *et al.*, 2015) evaluó una cohorte de 381 gestantes con diferentes grados de exposición a material particulado durante el curso de su embarazo. Al estudiar el nivel de metilación en el gen MT-RNR1 y el *D-Loop* en mitocondrias provenientes de la placenta de las mujeres, se encontró un aumento en la metilación en ambos loci. El porcentaje de metilación fue significativamente mayor cuando las gestantes se exponían al material particulado durante el primer trimestre de embarazo. La metilación aumentada también se relacionaba con disminución del número de copias del ADNmt en placenta, por lo cual los investigadores proponen que la exposición temprana a contaminantes ambientales (material particulado) podría llevar a mitofagia, disminución del ADNmt y muerte mitocondrial en el tejido placentario.

Uno de los casos más claros de la correlación entre la metilación diferencial de pacientes sanos y fenotipos patológicos en humanos vino a partir de los experimentos realizados por Pirola *et al.* (2013), quienes estudiaron los cambios moleculares que median la transición de un estado sin daño hepático (esteatosis simple) a una enfermedad más agresiva (esteatohepatitis no alcohólica- NASH). El equipo de Pirola (2013) encontró patrones de metilación diferencial en el gen que codifica para la deshidrogenasa NADH6 (MT-ND6), el cual se encuentra altamente metilado en el hígado de pacientes con NASH, pero no en pacientes con esteatosis simple. Este efecto mostró un impacto significativo en la regulación transcripcional del MTND6: en pacientes con NASH se encontró una menor actividad transcripcional y expresión de la proteína NADH6 a comparación con la esteatosis simple. Este patrón de metilación podría correlacionarse con el hallazgo de mayor expresión de la mtDNMT1 en pacientes con NASH (Pirola *et al.*, 2013).

Asimismo, Baccarelli y Byun (2015) estudiaron el grado de metilación del ADNmt en plaquetas como un marcador patológico en enfermedad cardiovascular (ECV). Las plaquetas, al carecer de ADN nuclear, constituyen un cultivo interesante para el estudio de la regulación del ADN mitocondrial. En este estudio los autores compararon el

porcentaje de metilación de 7 genes mitocondriales (MT-CO1, MT-CO2, MT-CO3, MT-TL1, MT-ATP6, MT-ATP8, MT-MD5) por medio de pirosecuenciación de ADN tratado con bisulfito en pacientes con ECV y controles sanos. Encontraron diferencias estadísticamente significativas en la metilación de los genes MT-CO1, MT-CO2, MT-CO3 y MT-TL1, en su mayoría encargados de la síntesis de ATP. Dichos genes se encuentran más metilados en pacientes que presentan ECV que en controles sanos, sin importar edad, índice de masa corporal o raza (Baccarelli y Byun, 2015). Una recopilación de los genes que se han encontrado con metilación selectiva, su transcrito y función puede encontrarse en la Tabla 1.

Estos hallazgos que sugieren que la expresión de genes mitocondriales está regulada por modificaciones epigenéticas y pueden desempeñar un papel importante en la patogénesis de la progresión de enfermedades, o como biomarcador temprano de patogénesis; asimismo los estudios muestran que dichos cambios epigenéticos podrían ser susceptibles a factores ambientales y como tal, podrían utilizarse como marcadores de polución o cambio ambiental, mostrando patrones diferenciales incluso en gestación.

CONCLUSIÓN

La mitoepigenética ha sido un campo altamente controvertido y pobremente estudiado desde el descubrimiento de 5mC en el ADN mitocondrial en 1984. Sin embargo, estudios recientes han probado la existencia de mecanismos epigenéticos tales como la metilación e hidroximetilación del ADNmt mediante la respectiva migración de enzimas nucleares como la DNMT a la mitocondria. También se ha comprobado la existencia de otros mecanismos epigenéticos como la remodelación de nucleoides y la presencia de miARN.

Estas alteraciones en los patrones de regulación se han correlacionado con niveles aberrantes en la transcripción y producción de ciertas proteínas mitocondriales y se ha encontrado asociación con entidades patológicas y cambios ambientales. A partir de estos resultados y los avances recientes en este campo, se podría considerar la epigenética mitocondrial como un campo de estudio fascinante, que podría dilucidar parte de los mecanismos moleculares detrás de enfermedades complejas, y que incluso podría ser usado como un biomarcador predictor en el futuro.

Tabla 1. Recopilación de los genes reportados con metilación selectiva, transcritos y funciones correspondientes. Se listan los genes que codifican para las 13 proteínas de la cadena de transporte de electrones, las dos subunidades del ribosoma mitocondrial (uno de los cuales codifica para la proteína Humanina, factor neuroprotector e inhibidor de apoptosis) y los dos ARN de transferencia que han mostrado metilación diferencial. Los principales genes reportados con cambios epigenéticos en diversas condiciones son MT-ND2, MT-ND6, que codifican subunidades del complejo NADH Deshidrogenasa y MT-RNR1, que codifica para la subunidad 12S del ribosoma mitocondrial. ECV (Enfermedad cardiovascular), NASH (Esteatopatía no alcohólica).

Función	GEN	OMIM	Transcripto	Regulación	Condición	Reporte	Ref.
Complejo I (NADH Deshidrogenasa)	MT-ND1	*516000	NADH-ubiquinona oxidoreductasa, Subunidad ND1	No reportado	-	-	-
	MT-ND2	*516001	NADH-ubiquinona oxidoreductasa, Subunidad ND2	Sí	Cáncer colorrectal Envejecimiento	Aumento de transcripción de ND2 en cáncer colorrectal desde estadios tempranos, correlacionada con demetilación del promotor ND2 en el <i>D-Loop</i> Aumento en la transcripción del gen en corteza cerebral de ratones con mayor edad (región rica en 5hmC).	Feng <i>et al.</i> , 2012 Dzitoyeva <i>et al.</i> , 2012
	MT-ND3	*516002	NADH-ubiquinona oxidoreductasa, Subunidad ND3	Sí			
	MT-ND4	*516003	NADH-ubiquinona oxidoreductasa, Subunidad ND4	Sí	Envejecimiento	Aumento en la transcripción del gen en corteza cerebral de ratones con mayor edad (región rica en 5hmC).	Dzitoyeva <i>et al.</i> , 2012
	MT-ND4L	*516004	NADH-ubiquinona oxidoreductasa, Subunidad ND4L	Sí			
	MT-ND5	*516005	NADH-ubiquinona oxidoreductasa, Subunidad ND5	Sí			
Complejo III (Coenzima Q - citocromo C reductasa)	MT-ND6	*516006	NADH-ubiquinona oxidoreductasa, Subunidad ND6	Sí	NASH Envejecimiento	Mayor metilación de ND6 en NASH vs. esteatosis simple. Menor actividad transcripcional y expresión de la proteína ND6 en NASH. Aumento en la transcripción del gen en corteza cerebral de ratones con mayor edad (región rica en 5hmC).	Pirola <i>et al.</i> , 2013 Dzitoyeva <i>et al.</i> , 2012
	MT-CYB	*516020	Citocromo B	No reportado	-	-	-
Complejo IV (Citocromo C Oxidasa)	MT-CO1	*516030	Citocromo c oxidasa I	Sí			
	MT-CO2	*516040	Citocromo c oxidasa II	Sí	ECV	Mayor metilación del gen en mitocondrias de plaquetas de pacientes con ECV vs. pacientes sanos.	Baccarelli y Byun, 2015
	MT-CO3	*516050	Citocromo c oxidasa III	Sí			
Complejo V (ATP sintasa)	MT-ATP6	+516060	ATP Sintasa, Subunidad ATPasa 6	No reportado	-	-	-
	MT-ATP8	+516070	ATP Sintasa, Subunidad ATPasa 8	No reportado	-	-	-
ARNr mitocondrial	MT-RNR1	*561000	Subunidad 12S ribosomal	Sí	Exposición a material particulado rico en metal Exposición a material particulado en gestación Envejecimiento	Aumento en metilación del gen en población expuesta a material particulado rico en metales pesados vs. población no expuesta Aumento de metilación del gen en placenta de mujeres expuestas a material particulado. Esta relación fue significativamente más fuerte en exposición en primer trimestre. Patrones diferenciales de metilación según edad. El porcentaje de metilación del gen disminuye significativamente con la edad en hombres	Byun <i>et al.</i> , 2013 Jannsen <i>et al.</i> , 2015 Giordano <i>et al.</i> , 2012
	MT-RNR2	*561010	Subunidad 16S ribosomal Humanina	No reportado No reportado	- -	- -	- -
ARNt mitocondrial	MT-TL1	*590050	tRNA-Leucina	Sí	ECV	Mayor metilación del gen en mitocondrias de plaquetas de pacientes con ECV vs. pacientes sanos.	Baccarelli y Byun, 2015
	MT-TF	*590070	tRNA-Fenilalanina	Sí	Exposición a material particulado rico en metal	Aumento en metilación del gen en población expuesta a material particulado rico en metales pesados vs. población no expuesta	Byun <i>et al.</i> , 2013!

BIBLIOGRAFÍA

- Baccarelli A.A., Byun H.M. (2015) Platelet mitochondrial DNA methylation: a potential new marker of cardiovascular disease. *Clinical epigenetics*, 7 (1): 44.
- Balinang J. (2012) The Regulation of Mitochondrial DNMT1 During Oxidative Stress. M.S. thesis, Virginia Commonwealth University, Richmond, Virginia.
- Bellizzi D., D'Aquila P., Scafone T., Giordano M., Riso V., Riccio A., Passarino G. (2013) The control region of mitochondrial DNA shows an unusual CpG and non-CpG methylation pattern. *DNA research*, 20 (6): 537-547.
- Byun H.M., Baccarelli A.A. (2014) Environmental exposure and mitochondrial epigenetics: study design and analytical challenges. *Human genetics* 133 (3): 247-257.
- Byun H.M., Panni T., Motta V., Hou L., Nordio F., Apostoli P., Baccarelli A.A. (2013) Effects of airborne pollutants on mitochondrial DNA methylation. *Part Fibre Toxicol.* 10 (1): 18.
- Chestnut B.A., Chang Q., Price A., Lesuisse C., Wong M., Martin L.J. (2011) Epigenetic regulation of motor neuron cell death through DNA methylation. *J. Neurosci.* Nov. 16; 31 (46): 16619-36.
- Chinnery P.F., Elliott H.R., Hudson G., Samuels D.C., Relton C.L. (2012) Epigenetics, epidemiology and mitochondrial DNA diseases. *Int. J. Epidemiol.* Feb; 41 (1): 177-87.
- Das S., Ferlito M., Kent O.A., Fox-Talbot K., Wang R., Liu D., Raghavachari N., Yang Y., Steenbergen C. (2012) Nuclear miRNA regulates the mitochondrial genome in the heart. *Circulation research* 110 (12): 1596-1603.
- Dzitoyeva S., Chen H., Manev H. (2012) Effect of aging on 5-hydroxymethylcytosine in brain mitochondria. *Neurobiology of aging* 33 (12): 2881-2891.
- Feng S., Xiong L., Ji Z., Cheng W., Yang H. (2012) Correlation between increased ND2 expression and demethylated displacement loop of mtDNA in colorectal cancer. *Mol. Med. Rep.* 6: 125-130.
- Ghosh S., Sengupta S., Scaria V. (2014) Comparative analysis of human mitochondrial methylomes shows distinct patterns of epigenetic regulation in mitochondria. *Mitochondrion.* Jul 22, pii: S1567-7249 (14).
- Ghosh S., Singh K.K., Sengupta S., Scaria V. (2015) Mitoeigenetics: The different shades of grey. *Mitochondrion* 25: 60-66.
- Giordano M., Cristiani C., Crocco P., D'Aquila P., De Rango F., Pisani F., Scafone T., Tallaro F., Rose G., Passarino G., Bellizzi D. (2012) Methylation of the human mitochondrial 12S rRNA gene is correlated with aging. In 12th International FISV Congress, Rome, Italy (Vol. 9).
- Iacobazzi V., Castegna A., Infantino V., Andria G. (2013) Mitochondrial DNA methylation as a next-generation biomarker and diagnostic tool. *Mol. Genet. Metab.* Sep-Oct; 110 (1-2): 25-34.
- Infantino V., Castegna A., Iacobazzi F., Spera I., Scala I., Andria G., Iacobazzi V. (2011) Impairment of methyl cycle affects mitochondrial methyl availability and glutathione level in Down's syndrome. *Molecular genetics and metabolism* 102 (3): 378-382.
- Janssen B.G., Byun H.M., Gyselaers W., Lefebvre W., Baccarelli A.A., Nawrot T.S. (2015) Placental mitochondrial methylation and exposure to airborne particulate matter in the early life environment: An ENVIRONAGE birth cohort study. *Epigenetics* 10 (6): 536-44.
- Kalani A., Kamat P.K., Voor M.J., Tyagi S.C., Tyagi N. (2014) Mitochondrial epigenetics in bone remodeling during hyperhomocysteinemia. *Mol. Cell Biochem.* Oct; 395 (1-2): 89-98.
- Kowluru R.A., Kowluru A., Mishra M., Kumar B. (2015) Oxidative stress and epigenetic modifications in the

- pathogenesis of diabetic retinopathy. *Progress in retinal and eye research* 48: 40-61.
- Locasale J.W. (2013) Serine, glycine and the one-carbon cycle: cancer metabolism in full circle. *Nature Reviews Cancer* 13 (8): 572-583.
- Lott M.T., Leipzig J.N., Derbeneva O., Xie H.M., Chalkia D., Sarmady M., Procaccio V., Wallace D.C. (2013) mtDNA variation and analysis using MITOMAP and MITOMASTER. *Current Protocols in Bioinformatics* 1 (123): 1.23.1-26. PMID: 25489354 URL: <http://www.mitomap.org>.
- Manev H., Dzitoyeva S., Chen H. (2012) Mitochondrial DNA: A Blind Spot in Neuroepigenetics. *Biomol. Concepts* Apr.; 3 (2): 107-115.
- Maximov V., Martynenko A., Hunsmann G., Tarantul V. (2002) Mitochondrial 16S rRNA gene encodes a functional peptide, a potential drug for Alzheimer's disease and target for cancer therapy. *Medical hypotheses* 59 (6): 670-673.
- Minocherhomji S., Tollefsbol T.O., Singh K.K. (2012) Mitochondrial regulation of epigenetics and its role in human diseases. *Epigenetics* Apr.; 7 (4): 326-34.
- Nass M.M., Nass S. (1963) Intramitochondrial fibers with DNA characteristics I. Fixation and electron staining reactions. *The Journal of cell biology* 19 (3): 593-611.
- Pirola C.J., Gianotti T.F., Burgueño A.L., Rey-Funes M., Loidl C.F., Mallardi P., Martino J.S., Castaño G.O., Sookoian S. (2013) Epigenetic modification of liver mitochondrial DNA is associated with histological severity of nonalcoholic fatty liver disease. *Gut*. Sep; 62 (9): 1356-63.
- Pollack Y., Kasir J., Shemer R., Metzger S., Szyf M. (1984) Methylation pattern of mouse mitochondrial DNA. *Nucleic acids research* 12 (12): 4811-4824.
- Rebelo A.P., Williams S.L., Moraes C.T. (2009) In vivo methylation of mtDNA reveals the dynamics of protein-mtDNA interactions. *Nucleic acids research* 37 (20): 6701-6715.
- Sadakierska-Chudy A., Frankowska M., Filip M. (2014) Mitoeipigenetics and drug addiction. *Pharmacol. Ther.* Nov; 144 (2): 226-233.
- Saraste M. (1999) Oxidative phosphorylation at the fin de siecle. *Science* 283 (5407): 1488-1493.
- Shock L.S., Thakkar P.V., Peterson E.J., Moran R.G., Taylor S.M. (2011) DNA methyltransferase 1, cytosine methylation, and cytosine hydroxymethylation in mammalian mitochondria. *Proc. Natl. Acad. Sci. USA* Mar 1; 108 (9): 3630-5.
- Skorecki K., Mandel H. (2012) DNA mitochondrial y enfermedades y rasgos hereditarios. En: Longo D.L. (Ed.) *Harrison Principios de Medicina Interna*, Capítulo 18, parte 3. McGraw-Hill Interamericana Editores 18a edición.
- Sun Z., Terragni J., Borgaro J.G., Liu Y., Yu L., Guan S., Zheng Y. (2013) High-resolution enzymatic mapping of genomic 5-hydroxymethylcytosine in mouse embryonic stem cells. *Cell reports* 3 (2): 567-576.
- Taylor R.W., Turnbull D.M. (2005) Mitochondrial DNA mutations in human disease. *Nat. Rev. Genet.* May; 6 (5): 389-402.
- Van der Wijst, M. G., & Rots, M. G. (2015). Mitochondrial epigenetics: an overlooked layer of regulation?. *Trends in Genetics*, 31(7), 353-356.
- Wong M., Gertz B., Chestnut B.A., Martin L.J. (2013) Mitochondrial DNMT3A and DNA methylation in skeletal muscle and CNS of transgenic mouse models of ALS. *Front Cell Neurosci.* Dec. 25 (7): 279.

AGRADECIMIENTOS

Agradecimiento especial al grupo de genética Humana de la Universidad de Los Andes y a la clase de epigenética del Doctor Valeriano López, donde nació el interés por el tema.

DIAGNÓSTICO GENÉTICO PREIMPLANTATORIO (DGP) EN PORTADORES DE REARREGLOS ESTRUCTURALES BALANCEADOS POR aCGH (ARRAY COMPARATIVE GENOMIC HYBRIDIZATION)

PREIMPLANTATIONAL GENETIC DIAGNOSIS (PGDs) IN CARRIERS OF BALANCED STRUCTURAL REARRANGEMENTS BY ARRAY COMPARATIVE GENOMIC HYBRIDIZATION (aCGH)

Ducatelli M.E.¹, A. Grazziotin Mondadori¹, F. Coco¹, N. Neuspiller¹, F.L. Gismondi¹, R. Coco^{1*}

¹Fecunditas, Instituto de Medicina Reproductiva, Afiliada a la Universidad de Medicina, UBA.

*Corresponding author: robertococo@fecunditas.com.ar

ABSTRACT

Carriers of chromosomal rearrangements have an increased risk of producing aneuploid gametes, which originate abnormal embryos, most of them lethal. Only a minority of them complete the pregnancy, delivering malformed children with unbalanced chromosomes. The majority of Preimplantational Genetic Diagnosis (PGDs) reported in these carriers were detected in blastomere using FISH. The objective of this work was to determine the risk of chromosome misbalance in blastocysts. Herein, we report 26 couples with chromosome rearrangements: five with pericentric inversions, four with Robertsonian translocations and 17 with reciprocal translocations. Trophectoderm biopsies were performed on days five or six post Intracytoplasmic Sperm Injection (ICSI). The array Comparative Genomic Hybridization (aCGH) was done with 24 Sure Plus of BlueGnome-Illumina® and the blastocyst transfer in a deferred cycle. Average of aspired oocytes was 16.2 ($r= 5-46$), of normal fertilized oocytes was 11.1 ($r= 2-26$) and of obtained blastocysts was 4.8 ($r= 1-11$). In the reciprocal translocation group, 63 blastocysts were biopsied, 44.4 % of them was normal and 55.6 % abnormal. In the Robertsonian translocations group, 23 blastocysts were studied, 56.5 % of them was normal and 43.5 % abnormal. In the chromosomal inversions group, 25 blastocysts were analyzed, 52 % was normal and 44 % abnormal. We could not obtain any information for one blastocyst due to a failure in DNA amplification. The results allow us to conclude that the theoretical expected risk for structural chromosome rearrangements in trophectoderm biopsy is lower than in blastomeres and/or gametes. Since not all fertilized oocytes reach the blastocyst stage, the quantity of biopsied blastocysts is much lower than in blastomere biopsy, reducing thus the costs of the PGDs.

Key words: Diagnóstico preimplantatorio, cariotipo molecular, biopsia de trofoblasto.

RESUMEN

Los portadores de rearrreglos cromosómicos tienen mayor riesgo para generar gametas aneuploides, que originan embriones anormales. Sólo una minoría llega a término y origina niños malformados con cromosomas desbalanceados. La mayoría de los reportes de DGPs (Diagnóstico Genético Preimplantatorio) en portadores de rearrreglos corresponden al estudio en blastómeras por FISH. El objetivo de nuestro trabajo fue determinar el riesgo de desbalance cromosómico en el estadio de blastocisto. La serie estuvo compuesta por 26 pacientes portadores de rearrreglos: cinco inversiones pericéntricas, cuatro fusiones céntricas y 17 translocaciones recíprocas. La biopsia de trofotodermo fue realizada en el día cinco o seis post Inyección Intracitoplasmática del Espermatozoide (ICSI). El aCGH (array Comparative Genomic Hybridization) fue realizado con el kit 24 Sure Plus BlueGnome-Illumina® y la transferencia del blastocisto en un ciclo diferido. El promedio de ovocitos aspirados fue 16,2 ($r= 5-46$), el de ovocitos fecundados normales fue 11,1 ($r= 2-26$) y el promedio de blastocistos obtenidos fue 4,8 ($r= 1-11$). De los 63 blastocistos biopsiados en el grupo de translocaciones recíprocas, 44,4 % resultaron normal y 55,6 % anormal. De los 23 blastocistos de las fusiones céntricas, 56,5 % resultó normal y 43,5 % anormal. De los 25 blastocistos provenientes de las inversiones pericéntricas, 52 % resultó normal y 44% anormal. Los resultados hallados evidencian que la llegada a blastocisto disminuye sustancialmente el riesgo teórico esperado en la fecundación. Por otro lado, como no todos los ovocitos fecundados llegan a blastocisto, se reduce el costo del DGP.

Palabras clave: molecular karyotype, preimplantation diagnosis, throphectoderm biopsy.

Fecha de recepción: 9/05/2016
Fecha de aceptación de versión final: 14/07/2016

INTRODUCTION

It is well recognized the increased frequency of balanced structural chromosome rearrangements in human infertility. One in every 500 live newborns has a structural balanced rearrangement, such as a reciprocal translocation, a Robertsonian translocation or a pericentric or paracentric inversion (Nielsen, 1991). The carriers of these rearrangements do not express abnormal phenotypes, but the rearrangements may determine sterility and/or infertility in adulthood, due to meiotic arrest or to abnormal chromosome segregation in Anaphase I from the meiotic multivalents or loops.

In meiosis, there are five chromosome segregation alternatives from quadrivalents resulting from reciprocal translocations: alternate, adjacent 1, adjacent 2, 3:1 segregation and 4:0. Only the alternate segregation originates normal gametes, with either the two normal or the two derivative chromosomes. The other segregation types originate abnormal gametes with various chromosomal unbalances. Therefore, in carriers of reciprocal translocations, the theoretical risk of producing abnormal gametes is 80%. The empiric risk observed in such carriers is almost equal or even higher than the theoretical 80% risk. In trivalents resulting from Robertsonian translocations, there are four segregation alternatives: alternate, adjacent 1, adjacent 2 and 3:0. Only the alternate segregation generates normal gametes, with either the two normal or the two derivative chromosomes. The other segregation types generate abnormal gametes with various chromosomal unbalances. Therefore, in carriers of Robertsonian translocations, the theoretical risk of the producing abnormal gametes is 75%. In contrast to reciprocal translocations, the empiric risk observed in such carriers is lower than the theoretical one.

On the other hand, there are three chromosome segregation alternatives from the meiotic loop of heterozygous for pericentric or paracentric inversions with symmetrical and asymmetrical crossing over: normal, non-disjunction and recombined chromosomes. Therefore, in heterozygous carriers for chromosomal inversions, the theoretical risk of the producing abnormal gametes is 66%. The observed empiric risk is almost lower than the theoretical one (Durban, 2001; Coco, 2005).

It is important to remark the differences in the percentage of empirical chromosomal risks during gamete production, preimplantational development at cleavage or at the blastocyst, embryonic, fetal stage or

newborn. The risks of abnormal gamete formation and preimplantational embryos stage are higher than the observed during pregnancy or in the newborn because most of chromosomal abnormalities are lethal. The lethality of a complete chromosome embryo aneuploidy is 99%, most embryo aborting during preimplantation or in the postimplantation stage during the first pregnancy trimester. Instead, partial aneuploidies are less lethal than the complete ones and can result in malformed newborns. Thus, it is important that carriers of the balanced chromosomal rearrangements receive advice for the preimplantational diagnoses, even more if they need the *In Vitro* Fertilization/Intracytoplasmic Sperm Injection (IVF/ICSI) procedure to conceive.

There are three types of biopsies: (a) in polar body; (b) in cleaved embryos on day three; and (c) in trophoctoderm on day five. According to the last data of the PGD ESHRE Consortium, 16.3% of the PGDs had been performed in polar bodies, 79.8% in D3 blastomeres and 2.3% in D5 blastocysts. According to the last register of the PGD Consortium, from 1,213 PGDs performed in carriers of Robertsonian translocations, 72% of the cycles achieved an euploid embryo transfer with a pregnancy rate per aspiration of 21% and 30% per transfer. From 2,413 PGDs performed in reciprocal translocations carriers, 58% of the cycles achieved an euploid embryo transfer, with 14% of pregnancy rate per aspiration and 24% per transfer (Goossens, 2012).

The blastocyst is the maximum stage reached of *in vitro* preimplantational embryo development. As the cells of the trophoctoderm and the majority of the inner cell mass conform the placenta, the villus and the extraembryonic mesoderm and ectoderm, the trophoctoderm biopsy is considered equivalent to the chorionic villus puncture. Therefore, both are indirect procedures whose results are not necessarily in agreement with the constitution of the future embryo-fetal-newborn, being 98% the estimated diagnostic efficiency.

When the choice is the trophoctoderm biopsy, it is preferable to perform the blastocyst transfer in a deferred cycle, due to the fact that not all the blastocysts are obtained on day five, the genetic laboratory has more time for performing the studies and, principally, because the pregnancy rate is equal or better than with the fresh embryo transfer. Nowadays, with better laboratory conditions and vitrification success, the blastocyst biopsy is the PGD best option, due to the robust results, the lower cost and the fact

that the genetic study can be programmed for working days.

Herein, we report our experience with blastocyst biopsy and deferred transfer in 26 PGD cycles in carriers of 17 reciprocal translocation, four Robertsonian translocations and five pericentric inversions, all of them evaluated by aCGH with the 24 Sure Plus BlueGnome-Illumina® platform.

PATIENTS AND METHODS

The couples accessed to PGD because of their recurrent abortion experiences and failures in IVF/ICSI procedures, the malformed stillbirths or the birth of a malformed child. The balanced rearrangements observed were:

(a) reciprocal translocations: 46,XX,t(7;11)(q11.2;q12); 45,XX,der(14)t(14;22)(q32;q11.1),-22; 46,XY,t(13;17)(q14;q23); 46,XX,t(6;10)(q13;q24); 46,XY,t(3;6)(q26;q24); 46,XX,t(9;13)(q34.3;q14.3); 46,XX,t(1;8)(q41-q42;q12); 46,XX,t(6;7)(q23;q34); 46,XY,t(3;8)(p21;p11.2); 46,XX,t(9;13)(q21;q21.2); 46,XX,t(2;22)(q35;q13); 46,XY,t(6;14)(q13;q31); 46,XY,t(6;7)(q25;p15); 46,XY,t(9;11)(q32;q21); 46,XY,t(2;9)(q37;p12); 46,XX,t(4;9)(q21;p22); 46,XY,t(10;19)(q10;q10).

(b) Robertsonian translocations or centric fusions: 45,XY,rob(13;21)(q10;q10); 45,XY,rob(13;14)(q10;q10),9ph; 45,XY,rob(14;21)(q10;q10); 45,XX,rob(13;14)(q10;q10).

(c) Pericentric inversions: 46,XY,inv(5)(p12q22); 46,XX,inv(5)(p14q21); 46,XY,inv(9)(p21q22); 46,XY,inv(8)(p23q11.2); 46,XY,inv(10)(p15q25.2).

The women's average age was 36.1 years old (range: 25-46). Previous to the PGD procedure, all women were subjected to the following hormonal studies: FSH, LH, E2 y AMH, antral follicles count and the hysterosalpingography. All men underwent a complete semen analysis with bacteriological sperm culture.

The ovarian stimulation was done with recombinant gonadotrophines and GnRH antagonist or agonist. Thirty six hours post-HCG application, oocytes were retrieved by trans-vaginal ultrasonography. ICSI procedure was performed in all cases with ejaculated semen. The normal fertilized oocytes were cultured until reaching the blastocyst stage on day five or six. On day four post-ICSI

the *zona pellucida* was perforated with laser shots to allow the aspiration of the hatched trophectoderm cells. After biopsy, the blastocysts were vitrified and the removed and washed trophectoderm cells were placed in Eppendorf microtubes for the molecular karyotype study. This study was performed with the BlueGnome-Illumina® 24 Sure Plus platform. Only the normal blastocysts were devitrified and transferred in a deferred cycle. The endometrium was prepared with estrogens and progesterone.

RESULTS

The average of aspired oocytes was 16.2 (5-46), the average of normal fertilized oocytes was 11.1 (2-26) and the average of reached blastocysts was 4.8 (1-11). Every couple achieved at least one blastocyst for biopsy.

In the reciprocal translocations group, 63 blastocysts were biopsied. Twenty eight (44.4 %) had normal karyotypes and 35 (55.6 %) had abnormal chromosome complements, being 15 of them (42.9%) product of a homogeneous missegregation of the meiotic quadrivalent, nine of them (25.7%) associated with aneuploidies, and 11 (31.4 %) of them with *de novo* aneuploidies unrelated to reciprocal translocations.

In the Robertsonian translocations group, 23 blastocysts were biopsied. Thirteen of them (56.5%) had normal karyotypes, and 10 (43.5 %) had abnormal karyotypes, with one of the abnormal karyotype (10%) being product of missegregation of the meiotic trivalent, one (10 %) associated with aneuploidy and eighty (80%) with *de novo* aneuploidies unrelated to the Robertsonian translocations.

In the pericentric inversion group, 25 blastocysts were biopsied. Thirteen (52%) had normal karyotypes, 11 (44%) had abnormal karyotypes and one (4%) was an amplification failure. Two of abnormal karyotypes (18.2 %) were products of a recombinant aneusomy of the meiotic loop, seven (63.6 %) with *de novo* aneuploidies unrelated with the rearrangement and two (18.2 %) were unbalanced segregation of a balanced complex rearrangement between a reciprocal translocation and a pericentric inversion. This case was a finding because the couple consulted for a chromosome 10 inversion but they really had a reciprocal translocation between chromosome 8 and an inverted chromosome 10. The abnormal karyotypes obtained in each group of balanced structural chromosome rearrangements (BSR) are presented in Table 1.

Twenty-six PGD cycles were performed; 20 of them were transferred with a normal devitrified blastocyst. Eight couples achieved an ongoing pregnancy, but one of them ended in a spontaneous abortion at eight weeks of gestation. In six couples, their blastocysts have not transferred yet.

Table 1. Blastocyst chromosome abnormalities of balanced structural rearrangement carriers.

Balanced Structural rearrangements (BSR)	Abnormal blastocyst karyotypes	Types of abnormal segregations	Chromosome anomalies unrelated to the BSR
	Reciprocal Translocations		
46,XY,t(3;8)(p21;p11.2)	46,XY,der(3)t(3;8)(p21;p11.2)	Adjacent 1	
	46,XY,der(3)t(3;8)(p21;p11.2)	Adjacent 1	
	46;XX,+3,-22	Seg 3:1	Monosomy 22
46,XX,t(6;10)(q13;q24)	45,X		Monosomy X
	46,XX,der(6)t(6;10)(q13;q24)	Adjacent 2	
46,XX,t(6;7)(q23;q34)	45,XY,-22		Monosomy 22
	46,XX,-7,+13	Seg 3:1	Trisomy 13
	46,XY,der(6)t(6;7)(q23;q34)	Adjacent 2	
	46,XX,-7,+22	Seg 3:1	Trisomy 22
46,XX,t(1;8)(q41-42;q12)	46,XY,der(1)t(1;8)(q41-42;q12)	Adjacent 1	
	45,XY,-8	Seg 3:1	
	46,XY,der(1)t(1;8)(q41-42;q12)	Adjacent 1	
	45,X,der(1)t(1;8)(q41;q12),del(3q10)	Adjacent 1	Partial Monosomy 3
	46,XX,der(8)t(1;8)(q41;q12),del(6)	Adjacent 1	
	45,XY,-15		Partial Monosomy 6
	46,XX,-3,+21		Monosomy 15 Monosomy 3 Trisomy 21
46,XY,t(3;6)(q26;q24)	45,XY,-15		Monosomy 15
46,XX,t(9;13)(q21;q21.2)	46,XX,+9,-21	Seg 3:1	Monosomy 21
	47,XX,+2,-9,+13,+17,-18,+20,-21		Caotic
	46,XY,+14,-15		Trisomy 14 Monosomy 15
	48,XXY,+19		Disomy X Trisomy 19

Balanced Structural rearrangements (BSR)	Abnormal blastocyst karyotypes	Types of abnormal segregations	Chromosome anomalies unrelated to the BSR
46,XY,t(6;14)(q13;q31)	46,XX,der(6)t(6;14)(q13;q31)	Adjacent 1	
	46,XX,der(6)t(6;14)(q13;q31)	Adjacent 1	
	46,XX,+19		Trisomy 19
46,XY,t(6;7)(q25;p15)	45,XX,der(7)t(6;7)(q25;p15),-22	Adjacent 1	Monosomy 22
	46,XX,der(6)t(6;7)(q25;p15)	Adjacent 1	
	46,XY,+14,-15		Trisomy 14 Monosomy 15
	46,XX,-7,+22	Seg 3:1	Trisomy 22
46,XY,t(9;11)(q32;q21)	46,XY,der(11)t(9;11)(q32;q21)	Adjacent 1	
46,XY,t(2;9)(q37;p12)	46,XX,der(2)t(2;9)(q37;p12)	Adjacent 1	
46,XX,t(4;9)(q21;p22)	45,XX,-9	Seg 3:1	
	46,XX,+4,-22	Seg 3:1	Monosomy 22
	47,XX,+19		Trisomy 19
46,XX,t(7;11)(q11.2;q12)	46,XX,der(11)t(7;11)(q11.2;q12)	Adjacent 1	
46,XY,t(10;19)(q10;q10)	46,XX,der(10)t(10;19)(q11.21;q12)	Adjacent 1	
Robertsonian Translocations			
45,XY,rob(13;21)(q10;q10)	45,XY,-13	Adjacent 2	
	47,XY,+3		Trisomy 3
	45,XX,-16		Monosomy 16
	47,XXY		Disomy X
45,XY,rob(13;14)(q10;q10),9ph	45,XX,-21		Monosomy 21
45,XY,rob(14;21)(q10;q10)	48,XX,+2,+9		Trisomy 2 y 9
	44,XY,-13,-22		Monosomy 13 y 22
	47,XX,+16		Trisomy 16
	46,X,+21	Adjacent 2	Monosomy X
	47,XX,-7,+15,+16		Monosomy 7
			Trisomies 15, 16

Balanced Structural rearrangements (BSR)	Abnormal blastocyst karyotypes	Types of abnormal segregations	Chromosome anomalies unrelated to the BSR
Pericentric inversions			
46,XY,inv(5)	45,XX,-16		Monosomy 16
	47,XXY		Disomy X
	46,XY,+6,-10		Monosomy 10 Trisomy 6
46,XY,inv(9)	47,XXY		Disomy X
46,XX,inv(5)	46,XX,rec(5)del5p,dup5q	Recombined	
	46,XX,rec(5)del5q	Recombined	
	48,XX,+17,+20		Trisomies 17, 20
46,XY,inv(8)	47,XX,+15		Trisomy 15
	46,XX,+16,-18		Trisomy 16 Monosomy 18
46,XY,inv(10)	46,XX,der(8)t(8;10)(q24.23;q25.1) ^a	Adjacent 1	
	46,XX,der(8)t(8;10)(q24.23;q25.1) ^a	Adjacent 1	

^a: These abnormal unbalances allow us to infer that the true karyotype of the patient is a reciprocal translocation between a chromosome #8 and an inverted chromosome #10

DISCUSSION

Before the development of the arrays technology, the PGDs in balanced rearrangement carriers were approached by FISH or by quantitative fluorescent Polymerase Chain Reaction (qfPCR) using probes or polymorphic markers linked to the involved chromosomes.

Actually, we have the possibility of molecular karyotyping, by using the qf-PCR, aCGH, simple nucleotide polymorphism arrays (aSNPs) and, ultimately, NGS or new generation sequencing (Bono, 2015), all of them better than the FISH technique because information is obtained for all chromosomes. Because of the high cost of the molecular studies it is preferable to perform it in embryos at the maximum stage of *in vitro* preimplantation development, because they have higher chances of becoming implanted in the endometrium. However, in the last report of the ESRHE PGD Consortium, it can be

appreciated that only 2.3 % of the PGDs are performed in the trophectoderm with deferred cycle transfer. We decided to end with the blastomere biopsies for several reasons: the best results published with devitrified blastocyst transfer, the less invasiveness of the trophectoderm biopsy *vs.* cleavage embryos, the largest number of aspirated cells for the genetic studies, the robust results and the possibility to organize the studies in workable days. The genetic study in a single cell is difficult and becomes a real challenge for the patients and the medical team.

In the present series of 26 cycles in carriers of balanced structural rearrangements, 111 blastocysts were biopsied, 54 had normal karyotypes (48.6%), 56 were abnormal (50.4%) and for one (0.09 %) no results were obtained due to a DNA amplification failure.

In Table 1, we report the details of the abnormal karyotypes product of a homogeneous missegregation

of the meiotic multivalent and loops, an unbalanced rearrangement associated with aneuploidies and with *de novo* simple or more complex aneuploidies. The aneuploidies associated or not with the unbalanced karyotypes related to the balanced structural rearrangements could be due to an interchromosomal effect during meiosis in the carrier, or to other causes, such as the employed ICSI procedure, the suboptimal conditions of *in vitro* embryo development, or an inherent aneuploidy risk related to the maternal age; however, with the aCGH performed in this series, we could not know the origin of the aneuploides to infer a likely explanation.

In the present series we found 44.4% of transferrable blastocyst in reciprocal translocations carriers, 56.5% in Robertsonian translocations carriers and 54.2% in pericentric inversion carriers. These results agree with those reported by Xiong *et al.* (2014) and Idowu *et al.* (2015), who performed the PGDs with aSNPs.

We consider that eight pregnancies out of 20 with only one euploid blastocyst transfer is a great stimulus to continue with the trophectoderm biopsy and deferred cycle transfer. In all cases we achieved at least one blastocyst for biopsy and one euploid for transfer.

Another important fact from the economic point of view is that the genetic studies are less expensive because it is performed only in those which reach the blastocyst stage, which are transferrable. Molecular karyotyping performed in few cells is expensive. For that reason, the PGD study in potentially transferrable embryos becomes relevant, especially in countries with few economic resources. The change of studying only one cell in cleavage embryos versus several cell at the blastocyst stage is significant because results are obtained in most of the assays performed. In fact, in our series of 111 blastocyst biopsied with five to 10 removed trophectoderm cells, we had a result in almost all of them, except in one due to an amplification failure. It is an important difference with the biopsy of a single blastomere in day three, in which the assays without results varied between 10 and 15 %.

A PGD program with fresh transfer requires that the molecular geneticist team works all the day, every day, without free days. Instead, a PGD program with deferred transfer is more organized for the genetic laboratory and, if there were doubts about some of the results, the studies can be either completed or repeated. Actually, there are no doubts that the pregnancy rate in cases of deferred cycle transfer is better than in the fresh transfer, not only due to

the pregnancy rate but also because of the better obstetric and perinatal results that can be obtained. These better results probably are due to the embryo-endometrium synchronization and the better endometrial receptivity in a non-stimulated cycle. In fact, it is known that the high concentration of estrogens in a stimulated cycle disturbs and damages the endometrial receptivity (Haouzi, 2010).

BIBLIOGRAPHY

- Bono S., A. Biricik, L. Spizzichino, A. Nuccitelli, M.G. Minasi, E. Greco, F. Spinella, F. Fiorentino (2015) Validation of a semiconductor next-generation sequencing-based protocol for preimplantation genetic diagnosis of reciprocal translocations. *Prenatal Diagnosis* 35: 938-944.
- Coco R., F. Coco Ludueña, M. Urquiza, J. Mincman, A. Gallo, N. Neuspiller (2005) Riesgo genético reproductivo en portadores de rearreglos cromosómicos. *Reproducción* 20 (1): 25-36.
- Coco R., A. Mondadori, M.E. Ducatelli, J. Mincman, A. Gallo, F. Coco, S. Neuspiller, F.L. Gismondi, N. Neuspiller (2012) Preimplantation diagnosis in blastocyst biopsy and deferred cycle transfer. *JBRA Assist Reprod.* 16 (5): 268-270.
- Durban M., J. Benet, M. Boada, E. Fernandez, J.M. Calafell, J.M. Laila (2001) PGD in female carriers of balanced Robertsonian and reciprocal translocations by first polar body analysis. *Hum. Reprod. Update* 7 (6): 591-602.
- Goossens V., J. Traeger-Synodinos, E. Coonen, M. De Rycke, C. Moutou, T. Pehlivan, I.A.P. Derks-Smeets, G. Harton (2012) ESHRE PGD Consortium data collection XI: cycles from January to December 2008 with pregnancy follow-up to October 2009. *Hum. Reprod.* 27 (7): 1887-1911.
- Haouzi D., S. Assou, C. Dechanet, T. Anahory, H. Dechaud, J. De Vos, S. Hamamah (2010) Controlled ovarian hyperstimulation for *In Vitro* Fertilization alters endometrial receptivity in Humans: Protocol effects. *Biol. Reprod.* 82 (4): 679-686.

Idowu D., K. Merrion, N. Wemmer, J. Gessner Mash, B. Pettersen, D. Kijacic, R. Lathi (2015) Pregnancy outcome following 24-Chromosome preimplantation genetic diagnosis in couple with balanced reciprocal or Robertsonian translocation. *Fertility and Sterility* 103: 1037-42.

Nielsen J., M. Wohlert (1991) Chromosome abnormalities found among 34,910 newborn children: results from a 13-year incidence study in Arhus, Denmark. *Hum. Genet.* 87 (1): 81-3.

Xiong B., K. Tan, Y.Q. Tan, F. Gong, S.P. Zhang, C.F. Lu, K.L. Luo, G.X. Lu, G. Lin (2014) Using SNP array to identify aneuploidy and segmental imbalance in translocation carriers. *Genomics Data* 2: 92-95.

PERSISTENT GENETIC DAMAGE INDUCED BY TOPOISOMERASE II POISONS IN NORMAL HUMAN FIBROBLASTS: GENERATION OF CHROMOSOME INSTABILITY

DAÑO GENÉTICO PERSISTENTE INDUCIDO POR VENENOS DE TOPOISOMERASA II EN FIBROBLASTOS HUMANOS NORMALES: GENERACIÓN DE INESTABILIDAD CROMOSÓMICA

de Campos-Nebel M.¹, Palmitelli M.¹, Acevedo S.², González-Cid M.^{1*}

¹Laboratorio de Mutagénesis, Instituto de Medicina Experimental, IMEX-CONICET, Academia Nacional de Medicina, Buenos Aires, Argentina.

²División Genética, Departamento de Patología y Diagnóstico, Instituto de Investigaciones Hematológicas (IIHema), Academia Nacional de Medicina, Buenos Aires, Argentina.

*Corresponding author: margoncid@hematologia.anm.edu.ar

ABSTRACT

Topoisomerase II (Top2) poisons idarubicin (IDA) and etoposide (ETO) are anticancer drugs that target Top2 stabilizing DNA-enzyme complexes and generating double-strand breaks (DSB). These DNA lesions are dangerous because they lead to genomic instability and promote tumorigenesis. These drugs are associated with the development of leukemias characterized by translocations of the *MLL* gene in treated patients. Our aim was to analyze the residual genetic damage induced by IDA and ETO in normal human fibroblasts at different times. Cells were treated with sublethal concentrations of IDA and ETO for 2 h and persistent DSB were evaluated in interphase nuclei and chromosomal aberrations (CA) in metaphase at 26 h post-treatment. In addition, micronuclei and *MLL* gene rearrangements were determined in interphase nuclei at 30 h post-treatment. Unrepaired persistent DSB induced by IDA and ETO turned into chromatid and chromosome breaks and improper repair in chromatid and chromosome exchanges. Simultaneously with increased CA, there was a marked reduction of the mitotic index, principally in cultures treated with IDA, due to the accumulation of cells in G2/M phase of cell cycle. This chromosomal damage progressed to the following interphase causing an increase in the micronucleated cells and in the rearrangements of *MLL* gene. The persistent DNA damage produced by IDA and ETO in normal human cells plays an important role in the possible induction of Top2 poisons-mediated secondary malignancies.

Key words: idarubicin, etoposide, DNA damage, *MLL* gene, normal human cells.

RESUMEN

Los venenos de topoisomerasa II (Top2) idarubicina (IDA) y etopósido (ETO) son drogas antitumorales que actúan sobre Top2 estabilizando complejos ADN-enzima y generando rupturas de doble cadena (RDC). Estas lesiones en el ADN son peligrosas debido a que llevan a inestabilidad genómica y a promover tumorigénesis. Estas drogas están asociadas al desarrollo de leucemias caracterizadas por translocaciones del gen *MLL* en los pacientes tratados. Nuestro objetivo fue analizar el daño genético residual inducido por IDA y ETO en fibroblastos humanos normales a diferentes tiempos. Las células se trataron con concentraciones subletales de IDA y ETO durante 2 hs y se evaluaron RDC persistentes en núcleos interfásicos y aberraciones cromosómicas (AC) en metafases a las 26 hs postratamiento. Además, se determinaron los micronúcleos y los rearrreglos del gen *MLL* en núcleos interfásicos a las 30 hs. Las lesiones persistentes inducidas por IDA y ETO se transformaron en rupturas cromatídicas y cromosómicas por falta de reparación y en figuras de intercambio cromatídico y cromosómico por reparación incorrecta. Junto al aumento de AC existió una marcada reducción del índice mitótico, en particular en los cultivos tratados con IDA, debido a la acumulación de células en la fase G2/M del ciclo celular. Este daño cromosómico progresó hacia la siguiente interfase causando un incremento de micronúcleos y de rearrreglos del gen *MLL*. El daño persistente en el ADN producido por IDA y ETO en células humanas normales juega un rol importante en la posible inducción de neoplasias secundarias debidas al tratamiento con venenos de Top2.

Palabras clave: idarubicina, etopósido, daño al ADN, gen *MLL*, células humanas normales.

INTRODUCCIÓN

Las topoisomerasas de tipo II (Top2) resuelven problemas topológicos del ADN durante la replicación, transcripción y condensación cromosómica. Esta enzima cataliza la liberación de torsiones en el ADN mediante la introducción de rupturas transitorias en la doble cadena del mismo (RDC) (McClendon y Osheroff, 2007). En los seres humanos existen dos isoformas, α y β , cuyas expresiones están reguladas durante el ciclo celular. Top2 α es fundamental para la supervivencia de células proliferantes y su expresión está relacionada con el crecimiento celular. En cambio, la concentración de Top2 β es independiente del estado proliferativo de las células (Heisig, 2009).

Un gran número de los regímenes antitumorales están basados en la utilización de drogas dirigidas contra Top2 debido a las numerosas funciones celulares desempeñadas por la enzima (Nitiss, 2009). Estas drogas son llamadas venenos de Top2 pues convierten a esta enzima esencial en una potente toxina que fragmenta al genoma en su presencia (Deweese y Osheroff, 2009). Los venenos de Top2 interfieren con el paso de religación de los extremos rotos durante el ciclo catalítico estabilizando los complejos ADN-enzima, conocidos como complejos de clivaje. La acumulación de los complejos de clivaje genera RDC en el ADN, que pueden producir rearrreglos cromosómicos y promover el desarrollo tumoral o conducir a la muerte celular (Elliott y Jasin, 2002). Entre los venenos de Top2 se encuentran idarubicina (IDA) y etopósido (ETO).

La antraciclina IDA es utilizada en el tratamiento de leucemia mieloide crónica, leucemia linfocítica aguda y cáncer de mama (Hollingshead y Faulds, 1991), mientras que la epipodofilotoxina ETO es empleada en el cáncer de pulmón a células pequeñas, sarcomas, tumores de la línea germinal, ovario, gástrico, mama, leucemias y linfomas (Hande, 1998).

Nuestros resultados previos demostraron que Top2 α es responsable de la inducción de RDC en el ADN y de la citotoxicidad de los venenos de Top2 IDA y ETO en células humanas normales (de Campos-Nebel *et al.*, 2010). Asimismo, se ha establecido que la isoforma α involucrada en la proliferación celular, sirve como blanco terapéutico para estas drogas antineoplásicas (Azarova *et al.*, 2007; Chen *et al.*, 2012).

Si bien los venenos de Top2 son ampliamente utilizados en la clínica médica, su empleo está asociado con severos efectos adversos como el desarrollo de leucemias agudas relacionadas a la terapia, especialmente leucemia mieloide aguda (t-LMA). La incidencia de leucemias relacionadas con la terapia ha aumentado debido a una mayor tasa de supervivencia de los pacientes y al uso de regímenes quimioterapéuticos más intensivos para tratar cánceres primarios (Cowell y Austin, 2012; Pendleton *et al.*, 2014). La presencia de t-LMA reflejaría un daño genético no letal inducido por venenos de Top2 en células precursoras hematopoyéticas.

Las translocaciones cromosómicas balanceadas representan las anomalías cariotípicas características de la t-LMA asociada con venenos de Top2. En este sentido, el locus de leucemia de linaje mixto (MLL) ubicado en 11q23 es el más frecuentemente involucrado en estas translocaciones, manifestándose en un período de latencia corto luego del tratamiento (Felix *et al.*, 2006; Leone *et al.*, 2010).

Debido a la elevada frecuencia con la cual los regímenes antineoplásicos que utilizan venenos de Top2 inducen tumores secundarios relacionados a la terapia de diversas patologías, el objetivo del presente trabajo fue evaluar la contribución de las RDC persistentes inducidas por IDA y ETO en generar inestabilidad cromosómica en células humanas normales.

MATERIALES Y MÉTODOS

Cultivo celular y tratamientos

La línea de fibroblastos humanos normales PTP (pasajes 25–35) fue proporcionada por la Dra. M.I. Tous, Servicio de Cultivo de Tejidos, Depto. de Virología, INEI-ANLIS “CG Malbrán”, Buenos Aires, Argentina. Las células crecieron en medio esencial mínimo MEM con 10 % suero fetal bovino, 2 mM L-glutamina, 100 U/ml penicilina y 100 μ g/ml estreptomina (medio completo) a 37° C en una atmósfera húmeda con 5 % CO₂. Las células en crecimiento exponencial se trataron con IDA (CAS N° 58957-92-9, Pharmacia y Upjohn) disuelta en agua destilada o ETO (CAS N° 33419-42-0, Sigma) disuelto en DMSO durante 2 hs, se lavaron dos veces con una solución de buffer fosfato (PBS) y se mantuvieron en medio completo durante diferentes tiempos. Los cultivos controles se trataron con los diluyentes correspondientes a cada agente empleado.

Evaluación de RDC (γ H2AX) persistentes

Por citometría de flujo

Las células se trataron con IDA, se lavaron con PBS y permanecieron en cultivo con medio completo por 4 o 26 hs. Las células tratadas con ETO se cosecharon inmediatamente, o se lavaron y dejaron en cultivo con medio completo por 26 hs. Luego de los tratamientos, las células se recolectaron, centrifugaron y fijaron en 1 % paraformaldehído (PFA) en PBS por 10 min a 37° C y se resuspendieron en 90 % metanol durante por lo menos 30 min a 4° C. Luego se incubaron en solución bloqueante (1 % seroalbúmina bovina en PBS) durante 1 h, en solución bloqueante conteniendo el anticuerpo anti- γ H2AX (1:200, *Cell Signaling*) por 1 h a temperatura ambiente y se resuspendieron en solución bloqueante conteniendo el anticuerpo secundario marcado con isotiocianato de fluoresceína (FITC, 1:100, Vector Laboratories) por 30 min en oscuridad a temperatura ambiente. Las células se lavaron en PBS y se adicionaron 200 μ g/ml RNAsa A (libre de DNAsas) y 10 μ g/ml ioduro de propidio (IP) incubando durante 15 min en oscuridad. La adquisición de 30.000 células por muestra se realizó mediante un citómetro de flujo FACSCalibur utilizando el software Cell Quest. Se agregaron controles de isotipo apropiados y se determinó la media de la intensidad de fluorescencia (IF) de γ H2AX en 4 experimentos independientes.

Por inmunofluorescencia en núcleos interfásicos

Las células crecidas sobre cubreobjetos se trataron con IDA o ETO y se mantuvieron en medio completo durante 26 hs. Las células se fijaron en 2 % PFA y permeabilizaron con 0,25 % Tritón X-100 en PBS por 10 min a temperatura ambiente. Se incubaron, sucesivamente, en solución bloqueante (5 % suero fetal bovino y 0,25 % Tritón X-100 en PBS) durante 1 h, en presencia del anticuerpo primario anti- γ H2AX (1:300) en solución bloqueante por 2 hs a 37° C en cámara húmeda y del anticuerpo secundario conjugado con FITC (1:200) en solución bloqueante durante 1 h a 37° C en cámara húmeda. Se colocaron los cubreobjetos sobre portaobjetos con medio de montaje conteniendo 4',6-diamidino-2-fenilindol diclorhidrato (DAPI, Vector Laboratories) como contra-colorante. El porcentaje de células con focos nucleares γ H2AX se analizó en 200 células por tratamiento en 2 experimentos independientes. En este recuento no se consideraron los focos γ H2AX presentes en los núcleos en fase S por estar relacionados con la replicación del ADN.

Aberraciones cromosómicas estructurales (AC) en metafase e índice mitótico (IM)

Las células se trataron y mantuvieron en medio completo por 26 hs. Se agregó 0,1 μ g/ml colchicina durante 2-4 hs, las células se recolectaron, se expusieron a una solución hipotónica 0,075 M KCl, se fijaron en 3 metanol: 1 ácido acético y se colorearon con 10 % Giemsa por 2,5 min. Se analizaron las AC (rupturas e intercambios cromatídicos y cromosómicos, figuras de intercambio cromatídico complejos y células altamente dañadas) en 100 metafases por tratamiento. El IM se calculó como el número de metafases en 2.000 núcleos y se expresó como porcentaje. Los valores de las AC y del IM se obtuvieron de 3 experimentos independientes.

Ciclo celular

Las células se trataron con IDA o ETO y se mantuvieron en medio completo por 26 hs. Se fijaron con 90 % metanol durante 1 h a -20° C, se lavaron en PBS y resuspendieron en PBS conteniendo 200 μ g/ml RNAsa A (libre de DNAsas) y 10 μ g/ml IP incubando 30 min a 37° C en oscuridad y se mantuvieron a 4° C hasta el momento de la adquisición mediante citómetro de flujo. Se adquirieron 30.000 células por muestra en 3 experimentos independientes utilizando el software Cell Quest. El análisis del ciclo celular se realizó utilizando el software ModFit.

Ensayo de micronúcleos (MN)

Las células crecidas sobre cubreobjetos se trataron y mantuvieron en medio completo por 30 hs. Luego se fijaron y permeabilizaron (1 % PFA y 0,25 % Tritón X-100 en PBS) y se procedió a la inmunomarcación para detectar señales γ H2AX (como se detalló previamente) en los MN y la presencia de MN en 2.000 células interfásicas por tratamiento en 4 experimentos independientes.

Estudio del gen MLL mediante hibridación in situ fluorescente (FISH)

Las células crecidas sobre cubreobjetos se trataron y cultivaron con medio completo durante 30 hs. Luego, se fijaron con 3 metanol: 1 ácido acético durante 20 min y se efectuó la técnica de FISH para determinar la presencia o ausencia del rearrreglo en el gen *MLL*. Para ello se utilizó la sonda de ADN *break apart* (Cytocell) siguiendo el protocolo establecido por el fabricante. Los preparados pretratados con pepsina y formaldehído, se lavaron con PBS y deshidrataron en alcoholes crecientes (70, 90 y 100 %). Cada preparado se desnaturalizó con formamida 70 % en

2XSSC a $72 \pm 1^\circ \text{C}$ durante 3 min y se pasó por alcoholes crecientes (70, 90 y 100 %). Por su parte, la sonda de ADN desnaturalizada a $72 \pm 1^\circ \text{C}$ se colocó sobre la zona elegida del preparado, se cubrió con un cubreobjeto y se incubó en cámara húmeda y oscura a 37°C durante 16 hs. Luego, se realizaron lavados en 2X SSC 0,4 % NP40 a 37°C por 2 min y 2X SSC 0,1 % NP40 a temperatura ambiente por 1 min. Se dejó secar en oscuridad y se contra-coloreó con DAPI. Como control positivo, se utilizó un cultivo de 24 hs de la médula ósea de un paciente (con los requisitos de consentimiento informado debidamente cumplimentados) con LMA confirmada mediante biología molecular (RT-PCR). El paciente era portador de una translocación entre el cromosoma 4 y 11.

Evaluación de células apoptóticas

Las células se trataron con IDA o ETO y se mantuvieron en medio completo por 30 hs. La caracterización morfológica de apoptosis se determinó usando naranja de acridina (NA)/bromuro de etidio (BE). Suspensiones celulares de 25 μl se mezclaron con 1 μl de la solución NA 100 $\mu\text{g}/\text{ml}$ + BE 100 $\mu\text{g}/\text{ml}$ en PBS y se extendieron sobre portaobjetos. Se contaron 200 células por tratamiento en 4 experimentos independientes.

Análisis estadístico

Las diferencias entre las medias de IF y de focos γH2AX , del IM, de MN y apoptosis se analizaron usando el test t de Student. Las AC se evaluaron mediante la prueba de Chi cuadrado (Primer of Biostatistics, version 3.0 by S.A. Glantz, McGraw-Hill, 1992).

RESULTADOS

RDC persistentes inducidas por venenos de Top2

Para analizar la cinética de reparación de las RDC inducidas por IDA o ETO en la línea de fibroblastos humanos normales PTP, se evaluó la IF media de γH2AX por citometría de flujo (Tabla 1 y Figura 1). Los valores máximos se obtuvieron a las 4 hs para IDA e inmediatamente para ETO luego de 2 hs de tratamiento. Una disminución de la IF media de γH2AX de $\sim 58\%$ se observó a las 26 hs postratamiento para ambas drogas, permaneciendo una fracción de RDC sin reparar. La determinación del

número de focos γH2AX en los núcleos interfásicos por inmunofluorescencia corroboró estos resultados (Figura 2). A las 26 hs, el porcentaje de células sin focos γH2AX fue de $63,4 \pm 5,9\%$, $16,8 \pm 2,4\%$ y $23,5 \pm 2,1\%$ en el cultivo control, en el tratado con IDA y en el tratado con ETO, respectivamente. El porcentaje de células con ≥ 20 focos γH2AX por célula fue de $11,0 \pm 1,4\%$ para IDA y de $17,0 \pm 3,8\%$ para ETO en relación a $1,5 \pm 0,7\%$ en los controles. Por otro lado, el análisis del número de focos γH2AX por núcleo mostró para IDA un valor de $4,87 \pm 0,62$ y para ETO de $4,22 \pm 0,06$, siendo de $1,07 \pm 0,04$ focos γH2AX por núcleo en los controles ($p < 0,013$).

Los venenos de Top2 causan AC en metafase

Las RDC persistentes en los núcleos interfásicos pueden dar origen a AC en la metafase inmediata por falta o incorrecta reparación de las mismas. La Tabla 2 muestra un aumento significativo en el porcentaje de metafases anormales en los cultivos tratados con IDA o ETO en comparación a sus respectivos controles. Las principales AC inducidas fueron rupturas cromatídicas y cromosómicas (resultado de RDC sin reparar) y figuras de intercambio cromatídico (debido a una incorrecta reparación de RDC). El efecto clastogénico de IDA y ETO se determinó convirtiendo las AC obtenidas como eventos de rupturas por célula (Figura 3a). Para obtener este cálculo se consideraron a las rupturas cromatídicas y cromosómicas como una ruptura y a los intercambios cromatídicos (cuadrirradiales y trirradiales) y cromosomas dicéntricos o en anillo como dos rupturas. En esta evaluación no se tuvieron en cuenta las figuras de intercambio cromatídico complejo y las células altamente dañadas debido a la dificultad de establecer el número de rupturas involucradas. La Figura 3b muestra los diferentes tipos de AC estructurales inducidas por los venenos de Top2. El aumento de las AC fue acompañado de una marcada reducción del IM (Figura 3c), en particular en los cultivos tratados con IDA, lo cual condicionó el recuento de metafases. El análisis del ciclo celular por citometría de flujo reveló que el IM descendió debido a la acumulación de células en la fase G2/M luego del tratamiento con los venenos de Top2 (Figura 3d). El porcentaje de células en G2/M se incrementó 1,7 veces en el tratamiento con IDA y 1,4 veces con ETO en comparación a sus respectivos controles (considerados como 1).

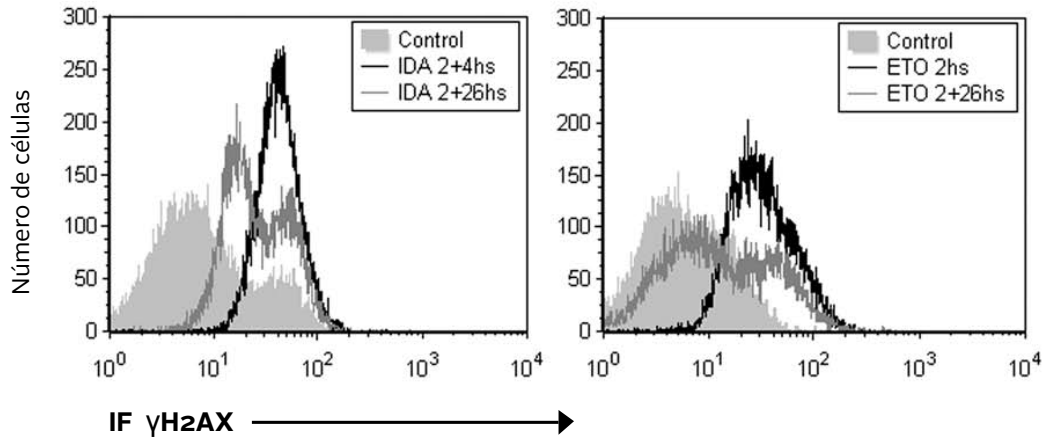


Figura 1. Histogramas representativos de la intensidad de fluorescencia (IF) de γ H2AX inducidas por IDA 0,1 μ g/ml y ETO 10 μ g/ml a diferentes tiempos en células humanas normales.

Tabla 1. Rupturas de doble cadena (γ H2AX) inducidas por venenos de Top2 en células humanas normales.

Tratamiento	IF media de γ H2AX \pm DE
Isotipo	0,43 \pm 0,11
Control	2,01 \pm 0,90
IDA 0,1 μ g/ml (2+4 hs)	76,45 \pm 3,40*
IDA 0,1 μ g/ml (2+26 hs)	31,53 \pm 11,89*
ETO 10 μ g/ml (2 hs)	42,88 \pm 4,39*
ETO 10 μ g/ml (2+26 hs)	18,04 \pm 6,45*

IF: intensidad de fluorescencia, DE: desvío estándar
*p= 0,0001 vs. control, test t de Student.

Tabla 2. Aberraciones cromosómicas estructurales inducidas por venenos de Top2 en células humanas normales a las 26 hs postratamiento.

Tratamiento	Células contadas	Células anormales (%)	rc	RC	FIC	Dicéntricos	Anillos	FIC complejos (%)	Células altamente dañadas (%)
CONTROL	300	4 (1,3)	4	0	0	0	0	0	0
DMSO 0,5 %	300	6 (2)	6	0	0	0	0	0	0
IDA 0,01 μ g/ml	150	38 (25,3)*	160	11	55	1	0	5 (3,3)	9 (6,0)
ETO 5 μ g/ml	300	37 (12,3)*	48	27	52	2	1	4 (1,3)	3 (1,0)

rc: rupturas cromatídicas, RC: rupturas cromosómicas, FIC: figuras de intercambio cromatídico (cuadrirradiales y trirradiales).

Células altamente dañadas: > 15 rupturas por metafase.

*p= 0,0001 vs. respectivos controles, prueba de Chi cuadrado.

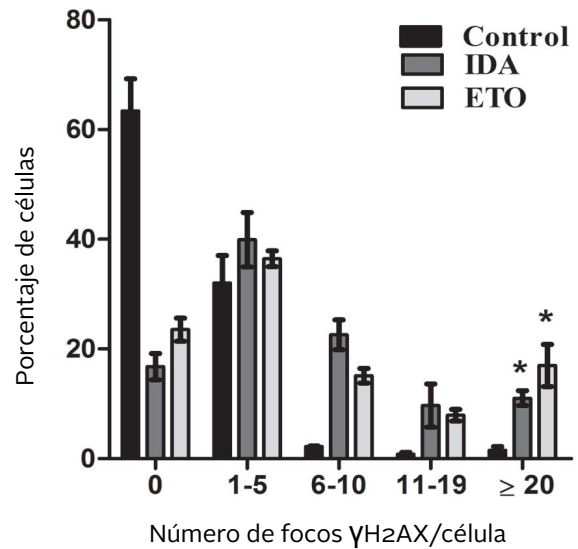


Figura 2. Distribución de rupturas de doble cadena (focos γ H2AX, media \pm DE) en núcleos interfásicos de células humanas normales tratadas con IDA 0,01 μ g/ml y ETO 5 μ g/ml durante 2 hs y analizadas por inmunofluorescencia a las 26 hs postratamiento. Dado que los controles sin tratar y tratados con DMSO 0,5% no mostraron diferencias estadísticamente significativas, los datos fueron agrupados.
* p<0,03 vs. control, test t de Student, n=2

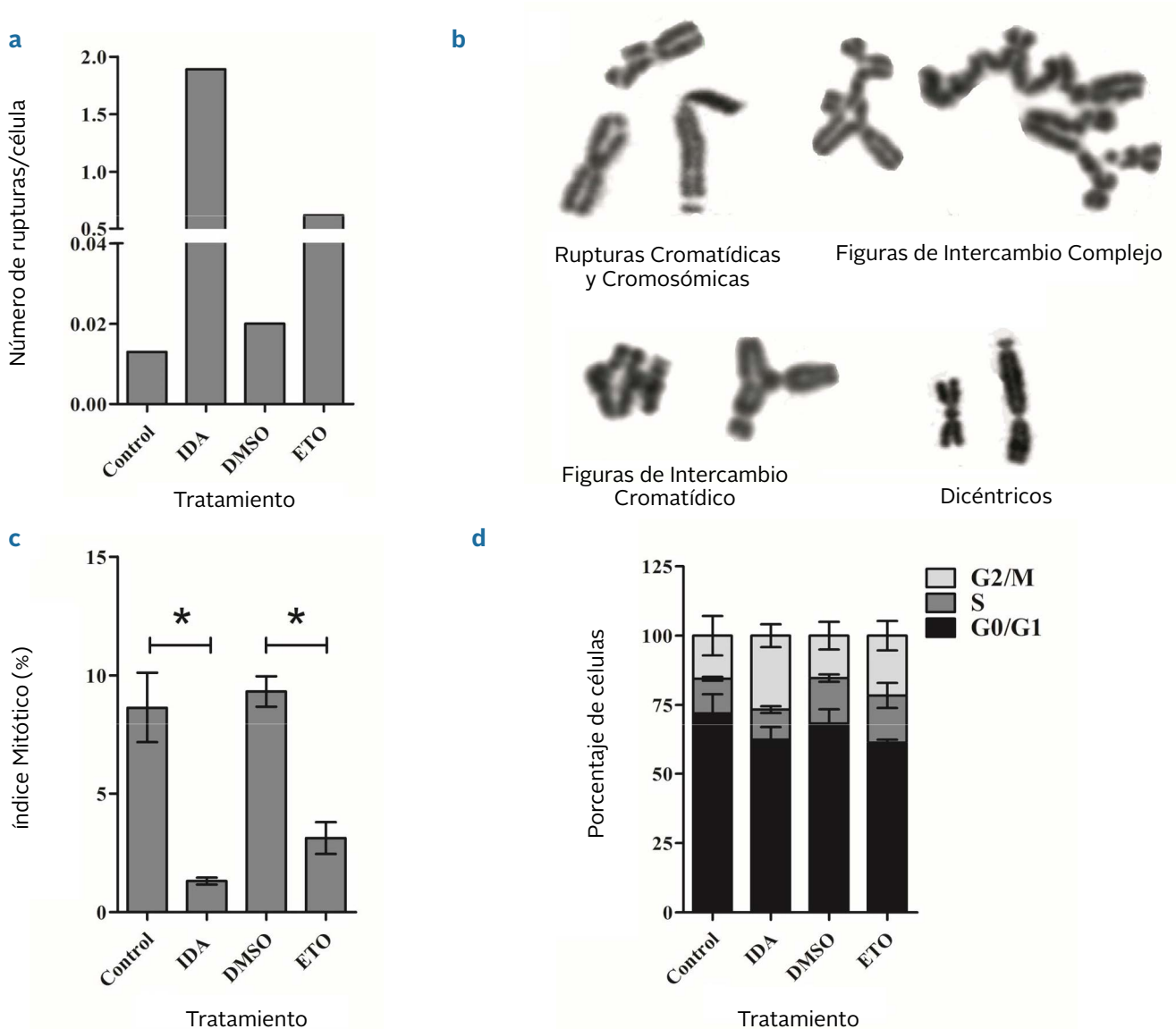


Figura 3. a. Rupturas cromosómicas totales por célula, b. Diferentes tipos de aberraciones cromosómicas estructurales, c. Índice mitótico (media \pm DE) y d. Distribución de las poblaciones celulares en el ciclo celular inducidos por IDA 0,01 μ g/ml y ETO 5 μ g/ml durante 2 hs y analizados a las 26 hs postratamiento. Rupturas cromosómicas totales: rupturas cromatídicas y cromosómicas se contaron como 1 ruptura y figuras de intercambio cromatídico, cromosomas dicéntricos y en anillo se contaron como 2 rupturas.

* p=0,0001 vs. controles respectivos, test t de Student.

MN inducidos por venenos de Top2 en interfase posmitótica

La progresión del daño cromosómico se determinó analizando los MN y los focos γ H2AX en los MN de células interfásicas posmitóticas a las 30 hs de efectuados los tratamientos (Figura 4). Los MN se forman a partir de cromosomas enteros (efecto aneugénico, MN sin focos γ H2AX) o fragmentos acéntricos (efecto clastogénico, MN con focos γ H2AX) excluidos de los núcleos de las nuevas células durante la anafase mitótica. La Figura 4a muestra un incremento del porcentaje de células micronucleadas entre 3-4 veces luego del tratamiento con IDA o ETO comparado con sus controles. Además, ambas drogas aumentaron el porcentaje de células micronucleadas con

focos γ H2AX (Figura 4b). De los MN inducidos por IDA, el 70,7 \pm 4,8 % presentó focos γ H2AX, mientras que en el control sin tratar el 52,7 \pm 4,3 % de los MN tenían señal γ H2AX. Por otro lado, en el tratamiento con ETO, el 61,9 \pm 3,1 % de los MN mostraron focos γ H2AX, siendo este valor de 54,7 \pm 2,9 % en el control tratado con DMSO. Estos resultados indicaron que los venenos de Top2 inducen MN cuyo origen es, en su mayor parte, clastogénico. En la Figura 4d se muestran MN con y sin focos γ H2AX. Por otro lado, a las concentraciones usadas de IDA y ETO se indujo un ligero aumento del porcentaje de células apoptóticas en relación a los controles, donde no se observó este efecto (Figura 4c).

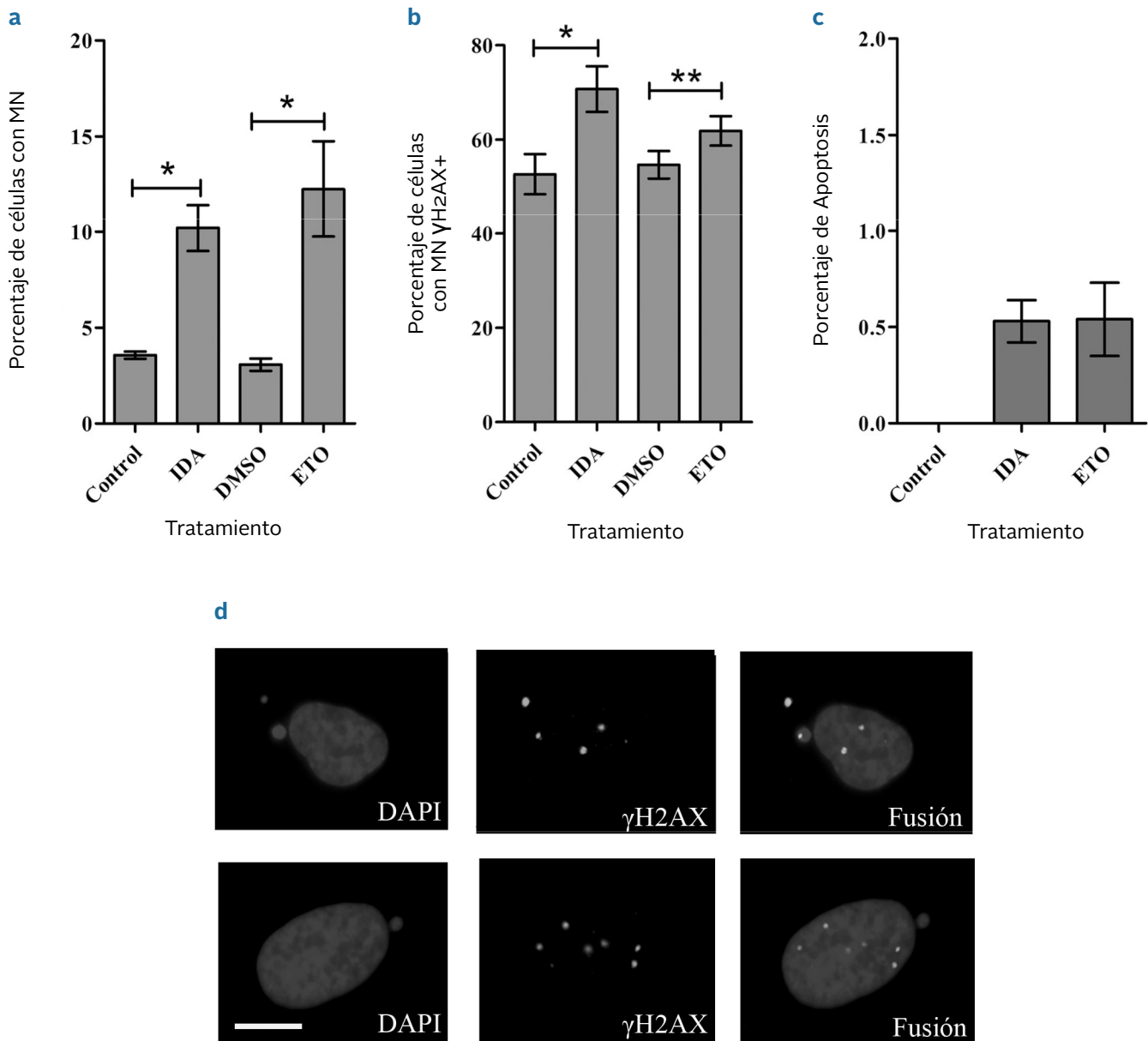


Figura 4. a. Células con micronúcleos (MN) (media \pm DE), b. Células con MN conteniendo focos γ H2AX (media \pm DE), c. Células apoptóticas (media \pm DE) inducidas por IDA 0,01 μ g/ml y ETO 5 μ g/ml durante 2 hs y analizadas a las 30 hs postratamiento y d. Núcleo con MN conteniendo focos γ H2AX (arriba) y núcleo con MN sin focos γ H2AX (abajo). Barra de escala: 10 μ m.

* p=0,001 vs. controles respectivos. ** p=0,013 vs. control tratado con DMSO 0,5%, test t de Student.

Rearreglos del gen *MLL* inducido por venenos de *Top2*

Los rearreglos cromosómicos recíprocos involucrando al gen *MLL* están asociados al desarrollo de LMA debido al uso de quimioterapias con venenos de *Top2*. Estos rearreglos son el resultado de errores en la reparación de las RDC persistentes. Los núcleos interfásicos de fibroblastos humanos normales tratados con IDA o ETO se analizaron a las 30 hs mediante la técnica de FISH. La sonda de ADN

dual color está diseñada de forma tal que flanquea al gen *MLL* y cuando se separa se considera que ha ocurrido un rearreglo del mismo. Ambos tratamientos indujeron rearreglos de este gen *MLL* revelados por la separación de las señales, mientras que en los cultivos controles se observó una estructura normal del mismo. El control positivo, médula ósea de un paciente con LMA, presentó el rearreglo en todas las células analizadas (Tabla 3).

Tabla 3. Núcleos con rearrreglos (señal *split*) del gen *MLL* inducidos por venenos de Top2 en células humanas normales a las 30 hs postratamiento.

Tratamiento	Células contadas	Núcleos con señal <i>MLL split</i>
Control	400	0
IDA 0,01 $\mu\text{g/ml}$	400	3
ETO 5 $\mu\text{g/ml}$	400	8
Control positivo (t-LMA)	100	100

DISCUSIÓN

Los venenos de Top2 actúan estabilizando los complejos de clivaje ADN-enzima, impidiendo la reunión de los extremos de ADN rotos y generando numerosas RDC para eliminar las células con activa proliferación (McClendon y Osheroff, 2007). Su uso está asociado a la aparición de neoplasias secundarias en los pacientes tratados (Cowell y Austin, 2012; Pendleton *et al.*, 2014). Además, estas drogas dañan al ADN de células normales, lo cual conduce a la formación de rearrreglos cromosómicos y al posible desarrollo de un nuevo tumor.

Las células humanas cuentan con dos mecanismos de reparación de las RDC para prevenir que estas lesiones se perpetúen: la vía de reunión de extremos no homólogos (*non-homologous end joining*, NHEJ) y la vía de recombinación homóloga (*homologous recombination*, HR). La principal diferencia entre ambas está en que NHEJ opera en forma rápida, aunque está sujeta a cometer errores y HR actúa en forma lenta pero es una vía precisa. Mientras que HR utiliza la cromátida hermana para restaurar la información original en la molécula dañada; NHEJ reúne los extremos rotos sin necesidad de homología originando mutaciones como sustitución de bases, inserciones o deleciones, en los sitios de la lesión (Shrivastav *et al.*, 2008).

Nuestros resultados previos establecieron la contribución de ambas vías en la reparación de las RDC inducidas por los venenos de Top2 IDA y ETO en fibroblastos humanos normales (de Campos-Nebel *et al.*, 2010). En este sentido y con el fin de profundizar los resultados obtenidos se evaluó el daño residual presente en los fibroblastos humanos por citometría de flujo. Se observó que la máxima inducción de RDC se alcanzó a las 4 hs o inmediatamente luego del tratamiento durante 2 hs para IDA y ETO, respectivamente,

y que existió una disminución por reparación de las mismas hacia las 26 hs, permaneciendo una parte de las lesiones sin reparar. Asimismo, la cuantificación de focos γH2AX en núcleos interfásicos por inmunofluorescencia confirmó la presencia de RDC persistentes en el ADN luego de 26 hs.

La fosforilación de la histona H2AX en serina 139, γH2AX , es un evento temprano en la reparación de las RDC y el recuento de los focos γH2AX , es usado como un marcador biológico de estas rupturas (Scully y Xie, 2013).

La falta o la incorrecta reparación de las RDC llevan a la formación de AC estructurales. Durante *et al.* (2013) establecieron que las RDC sin reparar aparecen en la mitosis como rupturas en el cromosoma y las RDC erróneamente reunidas producirán aberraciones de tipo intercambio intra- o entre cromosomas. Los venenos de Top2 causaron fundamentalmente un aumento de rupturas cromatídicas y cromosómicas y de figuras de intercambio cromatídico en las metafases evaluadas a las 26 hs postratamiento. El daño cromosómico inducido afectó la dinámica del ciclo celular y las células se acumularon en la fase G2/M, con la consiguiente reducción del número de metafases analizables.

Löbrich *et al.* (2010) reportaron que la duración del arresto en G2 depende de la dosis de las radiaciones y de la capacidad de reparación de las células. Sin embargo, este arresto no se mantiene hasta completar la reparación de las RDC. Se ha observado que fibroblastos humanos normales expuestos a rayos X son liberados del punto de control G2/M con ~ 20 focos γH2AX por núcleo. En este sentido, se estimó que ~ 1 de 10 RDC inducidas puede visualizarse como 1 ruptura cromosómica mitótica (Deckbar *et al.*, 2007).

La persistencia posmitótica de daño en el ADN llevará a inestabilidad genómica. Para evaluar la progresión del daño persistente se analizaron los MN presentes en interfase luego de 30 hs de realizados los tratamientos. Ambas drogas, IDA y ETO, indujeron un aumento en el porcentaje de células micronucleadas. Resultados similares se obtuvieron con venenos de Top2 en células de hámster chino CHO9 y V79 (Wang *et al.*, 2007; Elguero *et al.*, 2012).

El tipo de aberración cromosómica contenida en los MN depende del agente inductor (Norppa y Falck, 2003). Los venenos de Top2 promueven tanto aberraciones cromosómicas estructurales como numéricas (Degrassi *et al.*, 2004). En base a ello, se determinó el origen de los MN mediante inmunofluorescencia para detectar la presencia o no de focos γH2AX en los mismos. Los resultados

mostraron que la mayoría de los MN inducidos por IDA y ETO poseían señales γ H2AX en relación a sus respectivos controles, donde el porcentaje de MN con señales γ H2AX fue cercano al 50 %. Esto concuerda con que los MN inducidos por los venenos de Top2 contienen fragmentos acéntricos provenientes de rupturas en los cromosomas o en los puentes nucleoplásmicos originados a partir de figuras de intercambio cromatídico o cromosómico. Además, el tratamiento con IDA provocó más rupturas cromosómicas por metafase que ETO y esto se manifestó en un porcentaje más alto de MN con focos γ H2AX.

Así, las células con daño cromosómico alcanzaron la siguiente interfase. A su vez, en estas células interfásicas se pueden formar rearrreglos cromosómicos. Las translocaciones entre cromosomas representan el producto final de una serie de eventos celulares erróneos. Entre ellos, la presencia concomitante de RDC sin reparar, la reunión inadecuada de estas lesiones y la proximidad espacial y/o la relocalización de los extremos rotos de cromosomas diferentes dentro del espacio nuclear (Iarovaia *et al.*, 2014; Roukos y Misteli, 2014).

Numerosa evidencia clínica muestra una clara relación entre exposición previa a venenos de Top2 y el desarrollo subsecuente de t-LMA caracterizada por la translocación del gen *MLL* ubicado en la banda cromosómica 11q23 (Felix *et al.*, 2006; Le *et al.*, 2009; Leone *et al.*, 2010; Cowell y Austin, 2012; Pendleton *et al.*, 2014). Además, muchos de los rearrreglos aislados de los pacientes con t-LMA se localizan dentro de una región de 8,3 kb (*breakpoint cluster region*) que contiene secuencias de reconocimiento de clivaje de la enzima Top2 y elementos repetitivos (Nakada *et al.*, 2006; Sung *et al.*, 2006).

Teniendo en cuenta estos hallazgos, se evaluó la presencia de rearrreglos del gen *MLL* en núcleos interfásicos posmitóticos a las 30 hs. Los resultados obtenidos indicaron que los venenos de Top2 causaron rearrreglos cromosómicos en los fibroblastos humanos normales.

Las lesiones inducidas por IDA o ETO no generaron un daño genético masivo capaz de desencadenar procesos de muerte celular por apoptosis en forma substancial. Estos resultados son similares a los obtenidos en un trabajo previo de nuestro grupo (de Campos-Nebel *et al.*, 2010). Las concentraciones empleadas en la presente investigación son bajas (subletales) en relación a la hallada en el plasma de los pacientes tratados con ETO (Soubeyrand *et al.*, 2010).

La inestabilidad genómica es una característica fundamental de las células tumorales. La acumulación progresiva de mutaciones somáticas o translocaciones durante sucesivos ciclos celulares es un rasgo propio de la transformación maligna (Helleday *et al.*, 2014). A diferencia de este proceso gradual, se ha descrito recientemente el fenómeno de cromotripsis. Este fenómeno ocurre cuando un cromosoma se fragmenta en varias partes en un solo paso (*all at once*) y se vuelve a reunir mediante reparación por NHEJ en un orden aleatorio durante un único ciclo celular (Zhang *et al.*, 2015). Este evento está presente en muchos tipos de cánceres evidenciado al secuenciar el genoma de las células tumorales (Stephens *et al.*, 2011; Kloosterman y Cuppen, 2013).

Hatch y Hetzer (2015) mostraron la existencia de un nexo entre cromotripsis y MN. Los MN pueden ser descartados de la célula con la consiguiente pérdida de cromatina. Sin embargo, en otros estudios se observó su reincorporación al núcleo permaneciendo durante varias generaciones celulares (Terradas *et al.*, 2010; Crasta *et al.*, 2012). La cromatina de estos MN puede fragmentarse e integrarse al genoma celular. Crasta *et al.* (2012) reportaron que existe la posibilidad que el daño adquirido en el MN pueda causar translocaciones, aunque a una frecuencia baja (~1/100 núcleos por generación).

En resumen, ambas alteraciones cromosómicas, los MN y los rearrreglos del gen *MLL* observadas en células humanas normales generadas por IDA y ETO pueden contribuir al proceso tumorigénico como es la aparición de neoplasias secundarias en pacientes tratados con regímenes que involucran a los venenos de Top2.

BIBLIOGRAFÍA

- Azarova A.M., Lyu Y.L., Lin C.P., Tsai Y.C., Lau J.Y., Wang J.C., Liu L.F. (2007) Roles of DNA topoisomerase II isozymes in downregulated from chemotherapy and secondary malignancies. *Proc. Natl. Acad. Sci. U.S.A.*, 104: 11014-11019.
- Chen W., Qiu J., Shen Y.M. (2012) Topoisomerase II α , rather than II β , is a promising target in development of anti-cancer drugs. *Drug Discov. Ther.* 6: 230-237.
- Cowell I.G., Austin C.A. (2012) Mechanism of generation of therapy related leukemia in response to anti-

- topoisomerase II agents. *Int. J. Environ. Res. Public Health* 9: 2075–2091.
- Crasta K., Ganem N.J., Dagher R., Lantermann A.B., Ivanova E.V., Pan Y., Nezi L., Protopopov A., Chowdhury D., Pellman D. (2012) DNA breaks and chromosome pulverization from errors in mitosis. *Nature* 482: 53–58.
- de Campos-Nebel M., Larripa I., González-Cid M. (2010) Topoisomerase II-mediated DNA damage is differently repaired during the cell cycle by non-homologous end joining and homologous recombination. *PLoS One* 5: e12541.
- Deckbar D., Birraux J., Krempler A., Tchouandong L., Beucher A., Walker S., Stiff T., Jeggo P., Löbrich M. (2007) Chromosome breakage after G2 checkpoint release. *J. Cell. Biol.* 176: 749–755.
- Degrassi F., Fiore M., Palitti F. (2004) Chromosomal aberrations and genomic instability induced by topoisomerase-targeted antitumour drugs. *Curr. Med. Chem. Anticancer Agents* 4: 317–325.
- Deweese J.E., Osheroff N. (2009) The DNA cleavage reaction of topoisomerase II: wolf in sheep's clothing. *Nucleic Acids Res.* 37: 738–748.
- Durante M., Bedford J.S., Chen D.J., Conrad S., Cornforth M.N., Natarajan A.T., van Gent D.C., Obe G. (2013) From DNA damage to chromosome aberrations: Joining the break. *Mutat. Res.* 756: 5–13.
- Elguero M.E., de Campos-Nebel M., González-Cid M. (2012) DNA-PKcs-dependent NHEJ pathway supports the progression of topoisomerase II poison-induced chromosome aberrant cells. *Environ. Mol. Mutagen.* 53: 608–618.
- Elliott B., Jasin M. (2002) Double-strand breaks and translocations in cancer. *Review Cell. Mol. Life Sci.* 59: 373–385.
- Felix C.A., Kolaris C.P., Osheroff N. (2006) Topoisomerase II and the etiology of chromosomal translocations. *Review DNA Repair* 5: 1093–1108.
- Hande K.R. (1998) Etoposide: four decades of development of a topoisomerase II inhibitor. *Eur. J. Cancer* 34: 1514–1521.
- Hatch E.M., Hetzer M.W. (2015) Linking Micronuclei to Chromosome Fragmentation. *Cell* 161: 1502–1504.
- Heisig P. (2009) Type II topoisomerases-inhibitors, repair mechanisms and mutations. *Mutagenesis* 24: 465–469.
- Helleday T., Eshtad S., Nik-Zainal S. (2014) Mechanisms underlying mutational signatures in human cancers. *Nature Rev. Genet.* 15: 585–598.
- Hollingshead L.M., Faulds D. (1991) Idarubicin. A review of its pharmacodynamic and pharmacokinetic properties, and therapeutic potential in the chemotherapy of cancer. *Review Drugs* 42: 690–719.
- Iarovaia O.V., Rubtsov M., Ioudinkova E., Tsfasman T., Razin S.V., Vassetzky Y.S. (2014) Dynamics of double strand breaks and chromosomal translocations. *Review Mol. Cancer* 13: 249.
- Kloosterman W.P., Cuppen E. (2013) Chromothripsis in congenital disorders and cancer: similarities and differences. *Review Curr. Opin. Cell. Biol.* 25: 341–348.
- Le H., Singh S., Shih S.J., Du N., Schnyder S., Loredó G.A., Bien C., Michaelis L., Toor A., Diaz M.O., Vaughan A.T. (2009) Rearrangements of the MLL gene are influenced by DNA secondary structure, potentially mediated by topoisomerase II binding. *Genes Chromosomes Cancer* 48: 806–815.
- Leone G., Fianchi L., Pagano L., Voso M.T. (2010) Incidence and susceptibility to therapy-related myeloid neoplasms. *Chem. Biol. Interact.* 184: 39–45.
- Löbrich M., Shibata A., Beucher A., Fisher A., Ensminger M., Goodarzi A.A., Barton O., Jeggo P.A. (2010) γ H2AX foci analysis for monitoring DNA double-strand break repair: strengths, limitations and optimization. *Cell Cycle* 9: 662–669.
- McClendon A.K., Osheroff N. (2007) DNA topoisomerase II, genotoxicity, and cancer. *Review Mutat. Res.* 623: 83–97.

- Nakada S., Katsuki Y., Imoto I., Yokoyama T., Nagasawa M., Inazawa J., Mizutani S. (2006) Early G2/M checkpoint failure as a molecular mechanism underlying etoposide-induced chromosomal aberrations. *J. Clin. Invest.* 116: 80-89.
- Nitiss J.L. (2009) Targeting DNA topoisomerase II in cancer chemotherapy. *Nat. Rev. Cancer* 9: 338-350.
- Norppa H., Falck G.C. (2003) What do human micronuclei contain? *Review Mutagenesis* 18: 221-233.
- Pendleton M., Lindsey R.H.Jr., Felix C.A., Grimwade D., Osheroff N. (2014) Topoisomerase II and leukemia. *Review Ann. N.Y. Acad. Sci.* 1310: 98-110.
- Roukos V., Misteli T. (2014) The biogenesis of chromosome translocations. *Nat. Cell. Biol.* 16: 293-300.
- Scully R., Xie A. (2013) Double strand break repair functions of histone H2AX. *Review Mutat. Res.* 750: 5-14.
- Shrivastav M., De Haro L.P., Nickoloff J.A. (2008) Regulation of DNA double-strand break repair pathway choice. *Cell Res.* 18: 134-147.
- Soubeyrand S., Pope L., Haché R.J. (2010) Topoisomerase II α -dependent induction of a persistent DNA damage response in response to transient etoposide exposure. *Mol. Oncol.* 4: 38-51.
- Stephens P.J., Greenman C.D., Fu B., Yang F., Bignell G.R., Mudie L.J., Pleasance E.D., Lau K.W., Beare D., Stebbings L.A., McLaren S., Lin M.L., McBride D.J., Varela I., Nik-Zainal S., Leroy C., Jia M., Menzies A., Butler A.P., Teague J.W., Quail M.A., Burton J., Swerdlow H., Carter N.P., Morsberger L.A., Iacobuzio-Donahue C., Follows G.A., Green A.R., Flanagan A.M., Stratton M.R., Futreal P.A., Campbell P.J. (2011) Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* 144: 27-40.
- Sung P.A., Libura J., Richardson C. (2006) Etoposide and illegitimate DNA double-strand break repair in the generation of MLL translocations: new insights and new questions. *Review DNA Repair* 5: 1109-1118.
- Terradas M., Martín M., Tusell L., Genescà A. (2010) Genetic activities in micronuclei: is the DNA entrapped in micronuclei lost for the cell? *Mutat. Res.* 705: 60-67.
- Wang L., Roy S.K., Eastmond D.A. (2007) Differential cell cycle-specificity for chromosomal damage induced by merbarone and etoposide in V79 cells. *Mutat. Res.* 616: 70-82.
- Zhang C.Z., Spektor A., Cornils H., Francis J.M., Jackson E.K., Liu S., Meyerson M., Pellman D. (2015) Chromothripsis from DNA damage in micronuclei. *Nature* 522: 179-184.

AGRADECIMIENTOS

Este trabajo ha sido financiado por subsidios de la Fundación A.J. Roemmers y por el PIP-0114/14 de CONICET.



ANCESTRAL GENETIC COMPOSITION OF A HUMAN POPULATION FROM THE COLOMBIAN SOUTHWEST USING AUTOSOMAL AIM-InDels

COMPOSICIÓN GENÉTICA ANCESTRAL DE UNA POBLACIÓN HUMANA DEL SUDOESTE DE COLOMBIA USANDO AIM-InDels AUTOSÓMICOS

Urbano L.^{1,2}, Portilla E.C.^{1,2}, Builes J.J.³, Gusmão L.^{4,5}, Sierra-Torres C.H.^{1,2,*}

¹ Human Genetics Laboratory, Department of Physiological Sciences, Faculty of Health Sciences, University of Cauca, Popayán, Colombia.

² Molecular Diagnostics Unit, Innovagen Foundation, Popayán, Colombia.

³ Genes Ltda., Medellín Colombia.

⁴ IPATIMUP-Institute of Molecular Pathology and Immunology of the University of Porto, Porto, Portugal.

⁵ DNA Diagnostic Laboratory (LDD), State University of Rio de Janeiro (UERJ), Rio de Janeiro, Brazil.

*Corresponding author: hsierra@unicauca.edu.co

ABSTRACT

Genetic markers with allele distribution differences in human populations provide human ancestry information that can be used to evaluate admixture proportions in different ethnic groups, including stratification in genetic association studies. The aim of this study was to determine the African, European and Native American admixture proportion by making use of 46 insertion-deletion informative ancestry record markers set in 500 Latin American people in the Department of Cauca, Colombia. Allelic frequencies were established, Hardy Weinberg test was performed, comparisons in genetic distances among population groups were estimated using Arlequin v 3.5, and STRUCTURE v 2.3.4 software was used to determine the population's ancestry admixture. The results revealed a diverse genetic base, resulting from the three continental populations tested. The average ancestry observed in the Cauca population was: 48 % Native American, 39 % European, and 14 % African. This study reveals the complex population stratification of the Department of Cauca, which is the result of the genetic admixture of three continental populations.

Key words: genetic ancestry, InDels, AIMS, Latin America, human populations.

RESUMEN

Los marcadores genéticos con diferencias en la distribución de alelos en las poblaciones humanas proporcionan información de ancestría que se puede utilizar para evaluar las proporciones de mezcla de diferentes grupos étnicos, incluyendo la estratificación en los estudios de asociación genética. El objetivo de este estudio fue determinar las proporciones de mezcla Africana, Europea y Nativo Americana utilizando un conjunto de 46 marcadores informativos de ancestría de inserción-delección en una población de 500 sujetos Latino-Americanos del Departamento del Cauca, Colombia. Se estimaron las frecuencias alélicas, la prueba de Hardy Weinberg, se realizaron comparaciones de distancias genéticas entre poblaciones utilizando Arlequin v 3.5. Se utilizó El programa STRUCTURE v 2.3.4 para determinar la mezcla ancestral de la población. Los resultados mostraron una base genética muy diversa producto de las tres poblaciones continentales evaluadas. La ancestría promedio observada en la población del Cauca fue del 48 % de Nativos Americanos, el 39 % Europeos y el 14 % de Africanos. Este estudio revela la compleja estratificación poblacional del Departamento del Cauca, la cual es el resultado de la mezcla genética de tres poblaciones continentales.

Palabras clave: ancestría genética, InDels, AIMS, América Latina, poblaciones humanas.

Fecha de recepción: 26/04/2015
Fecha de aceptación de versión final: 11/09/2016

INTRODUCTION

The genetic composition of the Colombian population, like most Latin American populations, is the result of complex admixture processes among European, Native American and African ancestors that took place during the Spanish conquest of America (Chacón-Duque *et al.*, 2014). In addition to the Spaniards, an important number of Afro-descendants were brought to the continent as slaves. Throughout the last 500 years of history, the Colombian population has undergone genetic mixture and, as a consequence, the Colombian Native American population has been notably reduced (Acosta *et al.*, 2009). The population with mixed ancestry is mainly concentrated in the urban areas of the country, particularly in the Colombian Andes; the Afro-Colombian population lives mainly in the Caribbean and Pacific coasts and the islands, whereas the Native American populations are located predominantly in the East and in rural areas of the Southwest and North of the country (Rojas *et al.*, 2010). The population from the Department of Cauca is well known for its highly multiethnic background as a result of the admixture of parental populations (Córdoba *et al.*, 2012). According to data from the National Administrative Department of Statistics, Cauca has 1,182,022 inhabitants, of whom 56.25 % are white and “Mestizo” and are located mainly in the central area of the Department; 21.55 % are Native Americans located in the East; and 22.2 % are Afro-descendants who reside in the Pacific Coast (DANE, 2005). Genetic ancestry studies of human populations from highly admixed regions such as the Department of Cauca is currently considered a powerful tool to determine their admixture patterns (Neves Manta *et al.*, 2013). The implementation and use of polymorphic markers such as Ancestry Informative Markers (AIMs)-InDels, which present high intercontinental allelic differentiation across populations, allows the identification of allele frequency differences among populations and to infer their ancestral origin (Bastos-Rodrigues *et al.*, 2006; *et al.*, 2009; Pena *et al.*, 2009; Tishkoff *et al.*, 2009; Pena *et al.*, 2011; Manta *et al.*, 2013; Neves Manta *et al.*, 2013; Pepinski *et al.*, 2013; Zaumseger *et al.*, 2013; LaRue *et al.*, 2014; Shi *et al.*, 2015).

The aim of this study was to determine the admixture proportions of a population sample from the Department of Cauca based on the evaluation of ancestral contributions of three different continental populations, Native Americans, Europeans and Africans, using autosomal AIMs-InDels.

MATERIALS AND METHODS

Ethics statement

The questionnaires, procedures, and protocols used were reviewed and approved by the Ethics Committee for Scientific Research at the University of Cauca; following guidelines and bioethical principles established in the Helsinki declaration of the 206 World Medical Association and the parameters outlined in Resolution 8430 of 1993 by the Colombian Ministry of Health.

Sample collection and DNA extraction

During this study, a total of 500 samples were collected from unrelated individuals from 23 municipalities of the Department of Cauca, Colombia, as follows: 67.2 % were collected in the central region and 1.2 %, 26.6 % and 5.0 % were collected from the north, south and east regions, respectively. According to 2005 census by the National Administrative Department of Statistics, the Department of Cauca had a 68.3 % of inhabitants in the central zone, 3.5 % in the north zone, 23 % in the south zone, and 5.2 % in the east zone. Therefore, the population sample included in this study represents the general population (DANE, 2005). The genetic markers information was obtained from Marshfield Diallelic Insertion/Deletion Polymorphisms database website (Weber *et al.*, 2002) and GENES Ltda. Blood samples were collected in tubes containing EDTA and coded with internal control codes to protect the identity of each participant. Then, DNA was extracted using the standard salting-out procedure (Miller *et al.*, 1988). A DNA concentration of 30 ng/ μ L, measured with a NanoDrop 2000c Spectrophotometer, was used for the analysis (Desjardins *et al.*, 2010).

Genotyping

We used a set of 46 autosomal AIMs-InDels (Table 1), previously reported and validated for human population groups of Africans, Europeans and Native Americans (Shriver *et al.*, 1997; Pereira *et al.*, 2009). Amplification and genotyping of the 46 AIM-InDels was performed using the QIAGEN Multiplex PCR kit (Qiagen) combined with a mixture of 0.1 μ M of all primers and 5 ng of genomic DNA to complete a 6 μ L of final reaction solution. The amplified products were separated by electrophoresis in 5 % polyacrylamide gels and visualized with a FMBIO IIe Genetic Analyzer (HITACHI GENETIC Systems, Alameda, CA, USA).

Polymerase Chain Reaction

PCR reaction admixture calculus were done multiplying a sample volume (μL) by the total number of individuals; following, the 46 InDels panel markers were prepared according to the Genes laboratory protocol; the final volume per sample was 6 μL . For each assembly, a negative control under the same condition was considered (Tables 2 and 3).

Statistical analysis

Estimation of allele frequencies, Hardy Weinberg Equilibrium testing and population comparisons by pairwise genetic distances (F_{ST}) were carried out using Arlequin v 3.5 (Excoffier *et al.*, 2010) (Schneider *et al.*, 2009). For an easier visualization of the observed genetic distances, a multidimensional scaling (MSD) plot of the pairwise F_{ST} was performed using STATISTICA v7.0.

The software STRUCTURE v 2.3.4 was used to determine the ancestral admixture of the population under study and to infer the population structure (Pritchard *et al.*, 2000). The evaluation of the ancestral proportions was performed by using published information on the geographic origin of the reference samples. PDGH-CEPH were used as reference samples (Pereira *et al.*, 2012), and a tri-hybrid contribution from Native Americans, Europeans and Africans ($K= 3$) to the current genetic makeup of Cauca population was assumed. Furthermore, 100,000 burning steps followed by 100,000 Markov Chain Monte Carlo iterations, using both “Admixture Model” and only the reference individual with POPFLAG=1, were considered in the STRUCTURE analysis (Pritchard *et al.*, 2000). The option “Use population Information to test for migrants” was used in the Admixture Model.

Table 1. InDels allele frequency, in HGDP-CEPH diversity panel population samples from Africa (AFR), Europe (EUR) and Native America (NAM).

IndeL	Rs Number	Chromosome	Frequency of allele 1			Frequency of allele 2		
			AFR	EUR	NAM	AFR	EUR	NAM
MID 1470	rs2307666	11	0.162	0.563	0.031	0.838	0.437	0.969
MID-777	rs1610863	16	0.390	0.288	0.844	0.610	0.712	0.156
MID-196	rs16635	6	0.538	0.472	0.063	0.462	0.528	0.938
MID-881	rs1610965	5	0.324	0.880	0.992	0.676	0.120	0.008
MID-3122	rs35451359	18	0.552	0.991	1.000	0.448	0.009	0.000
MID-548	rs140837	6	0.338	0.196	0.172	0.662	0.804	0.828
MID-659	rs1160893	2	0.276	0.104	0.258	0.724	0.896	0.742
MID-2011	rs2308203	2	0.162	0.791	0.906	0.838	0.209	0.094
MID-2929	rs33974167	8	0.790	0.655	0.977	0.210	0.345	0.023
MID-593	rs1160852	6	0.695	0.006	0.008	0.305	0.994	0.992
MID-798	rs1610884	5	0.295	0.642	0.117	0.705	0.358	0.883
MID-1193	rs2067280	5	0.178	0.098	0.141	0.822	0.902	0.859
MID-1871	rs2308067	7	0.081	0.326	0.289	0.919	0.674	0.711
MID-17	rs4183	3	0.795	0.338	0.898	0.205	0.662	0.102
MID-2538	rs3054057	15	0.000	0.494	0.063	1.000	0.506	0.938
MID-1644	rs2307840	1	0.229	0.953	0.195	0.771	0.047	0.805

Indel	Rs Number	Chromosome	Frequency of allele			Frequency of allele		
			1			2		
			AFR	EUR	NAM	AFR	EUR	NAM
MID-3854	rs60612424	6	0.748	0.013	0.008	0.252	0.987	0.992
MID-2275	rs3033053	14	0.486	0.095	0.523	0.514	0.905	0.477
MID-94	rs16384	22	0.071	0.155	0.648	0.929	0.845	0.352
MID-3072	rs34611875	18	0.100	0.940	0.977	0.900	0.060	0.023
MID-772	rs1610859	5	0.943	0.968	0.781	0.057	0.032	0.219
MID-2313	rs3045215	1	0.110	0.247	0.730	0.890	0.753	0.270
MID-397	rs25621	6	0.771	0.766	0.344	0.229	0.234	0.656
MID-1636	rs2307832	1	0.148	0.785	0.992	0.852	0.215	0.008
MID-51	rs16343	4	0.057	0.642	0.867	0.943	0.358	0.133
MID-2431	rs3031979	8	0.052	0.11	0.234	0.948	0.889	0.763
MID-2264	rs34122827	13	0.162	0.465	0.133	0.838	0.449	0.867
MID-2256	rs133052	22	0.000	0.222	0.836	1.000	0.778	0.164
MID-128	rs6490	12	0.024	0.481	0.047	0.976	0.519	0.953
MID-15	rs4181	2	0.195	0.453	0.836	0.805	0.547	0.164
MID-2241	rs3030826	6	0.814	0.293	0.070	0.186	0.707	0.930
MID-419	rs140708	6	0.919	0.810	0.563	0.081	0.190	0.438
MID-943	rs1611026	5	0.176	0.791	0.398	0.824	0.209	0.602
MID-159	rs16438	20	0.752	0.570	0.734	0.248	0.430	0.266
MID-2005	rs2308161	10	0.029	0.671	0.325	0.971	0.329	0.675
MID-250	rs16687	7	0.824	0.715	0.328	0.176	0.285	0.672
MID-1802	rs2307998	5	0.543	0.003	0.000	0.457	0.997	1.000
MID-1607	rs2307803	3	0.424	0.161	0.555	0.576	0.839	0.445
MID-1734	rs2307930	6	0.852	0.835	0.375	0.148	0.165	0.625
MID-406	rs25630	6	0.048	0.823	0.703	0.952	0.177	0.297
MID-1386	rs2307582	1	0.138	0.218	0.766	0.862	0.782	0.234
MID-1726	rs2307922	1	0.176	0.700	0.758	0.824	0.300	0.242
MID-3626	rs11267926	15	0.124	0.699	0.039	0.876	0.301	0.961
MID-360	rs25584	12	0.591	0.864	0.523	0.317	0.136	0.477
MID-1603	rs2307799	5	0.081	0.389	0.820	0.919	0.611	0.180
MID-2719	rs34541393	20	0.448	0.345	0.469	0.552	0.655	0.531

Modified from Pereira *et al.* (2012).

Table 2. Reactive mixture for PCR (polymerase chain reaction).

Reactive	µL / Sample
Water	1.9
Mix Quiagen Multiplex PCR	3.0
Primer's	0.6
DNA	0.5
Final volume	6

Table 3. Multiplex PCR profile.

First denaturing/ Denaturalization	Anneal 10 cycles	Anneal 20 cycles	Final Extension	
95° C 15 minutes	94° C 30 Secs	94° C 30 Secs	60° C 10 Minutes	15° C 10 Infinite
	60° C 1.5 Minutes	58° C 1.5 Minutes		
	70° C 1 Minutes	72° C 1 Minutes		

RESULTS

Allele frequencies

The allele frequencies and expected heterozygosities estimated for the 46 AIM-InDels are presented in Table 4. The generated data from all samples were used to test Hardy Weinberg Equilibrium and p values were corrected by Bonferroni's correction for multiple tests (Bland *et al.*, 1995). The InDels frequency distribution showed no statistically significant deviations from HWE within Cauca population.

Genetic distance analysis

A two-dimensional MDS plot was drawn to improve the visualization of genetic distances (Figure 1). It can be observed that Cauca and Manaus populations are neighbors and that these two populations are the result of the Native Americans and Europeans admixture, as other Brazilian populations.

Analysis of admixture

The ancestry proportions from the admixture analysis are shown in Figures 2A and 2B. The average ancestry of the studied population is 48 % Native American, 14 % African and 39 % European (Figure 3).

Principal component analysis

Principal component analysis was done to reduce the data dimension (Figure 4). It can be observed that the studied population is composed of European and Amerindian populations (biplot's right side); on the other hand, a minority arrangement towards the African population is also observed (biplot's left side). These findings set the Department of Cauca as a mixed race population territory (Figure 4).

Table 4. Frequencies of 46 insertion/deletion polymorphisms, obtained in individuals from the Department of Cauca.

Population Samples	n (%)	Frec_AM	Frec_AF	Frec_EU
North	20 (3.9)	0.31	0.35	0.34
South	103 (20.6)	0.52	0.12	0.36
Center	355 (71.0)	0.46	0.13	0.41
Eastern	22 (4.5)	0.49	0.10	0.41
Average Department of Cauca	500 (100)	0.45	0.17	0.38

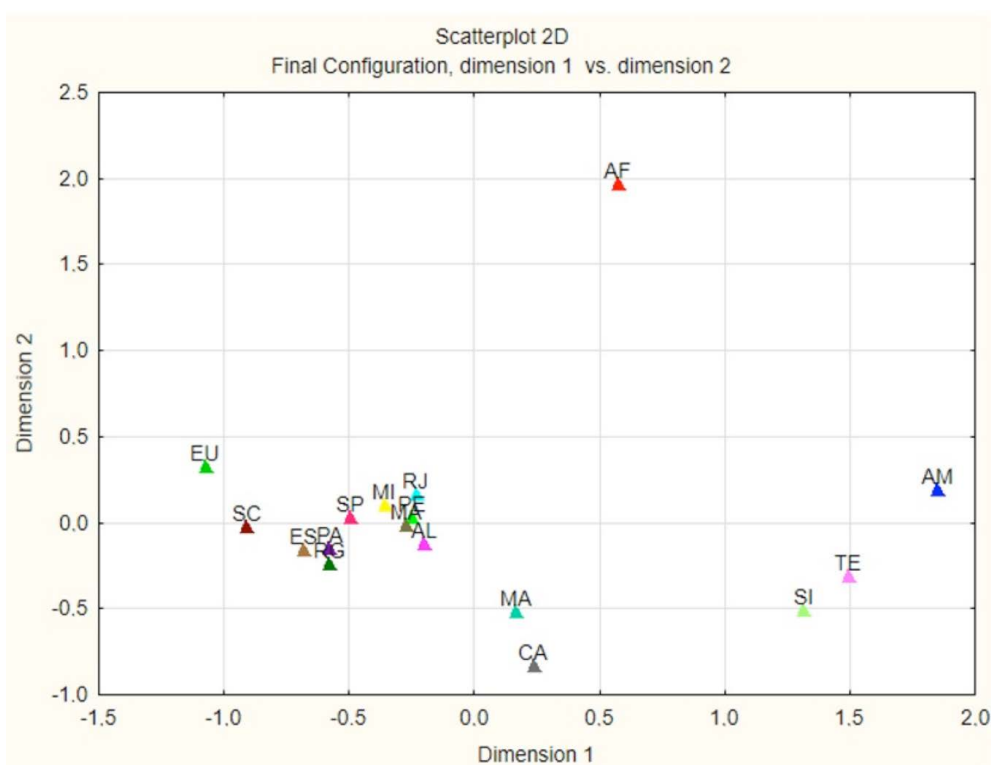
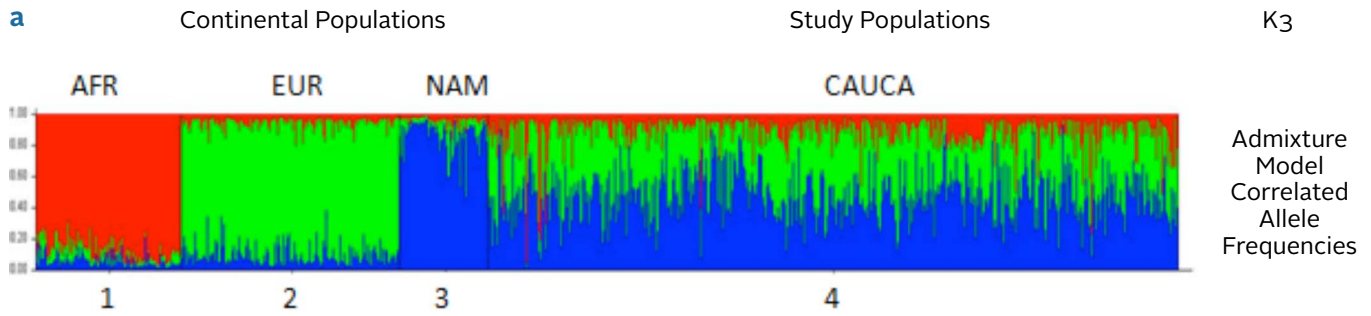


Figure 1. Multidimensional scaling plot of the pairwise F_{st} genetic distances calculated for samples from Native Americans, European, African and Cauca populations, including samples of Brazil; SI: Santa Isabel do Rio Negro; MA: Manaus; PE: Pernambuco; AL: Alagoas; MT: Mato Grosso do Sul; TE: Terena; MI: Minas Gerais; ES: Espírito Santo; RJ: Rio de Janeiro; SP: São Paulo; PR: Paraná; SC: Santa Catarina; RG: Rio Grande do Sul).



Bart Plots- direct output from STRUCTURE (Modified Colors)

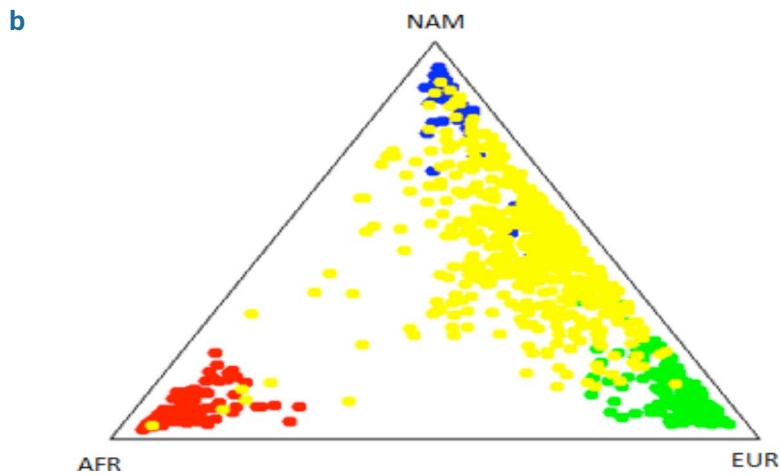


Figure 2A. Bar plot showing individual ancestry estimates obtained from Cauca population and parental populations. Based on STRUCTURE analysis with 100,000 burning steps followed by 100,000 MCMC iterations and $k=3$ with POPFLAG=1, treated in CLUMPP and plotted with destruct (AFR: Africa; EUR: Europe; NAM: Native America). Panels 1, 2 and 3 present the distribution of the African, European and Native American populations, respectively. On the other hand, panel 4 shows the distribution of individuals from the Department of Cauca evidencing a homogeneous contribution from both the Native American and European populations; notice that the largest proportion corresponds to “mestizo”. The African contribution is present in a low proportion.

Figure 2B. Triangle plot representation to distinguish individuals from different parental populations. Based on STRUCTURE results from the run with highest $-\ln P(D)$ (left: admixture model; right: using population information; red: Africa; green: European; blue: Native American; yellow: Cauca). For each individual, the values of the three ancestry coefficients are indicated by the distance to the edges of the triangle. Each individual is represented by a colored point and a homogenous distribution of the population from Cauca in both clusters of European and Native American population is shown. Furthermore, a lower proportion of individuals from Cauca in the African population cluster are observed.

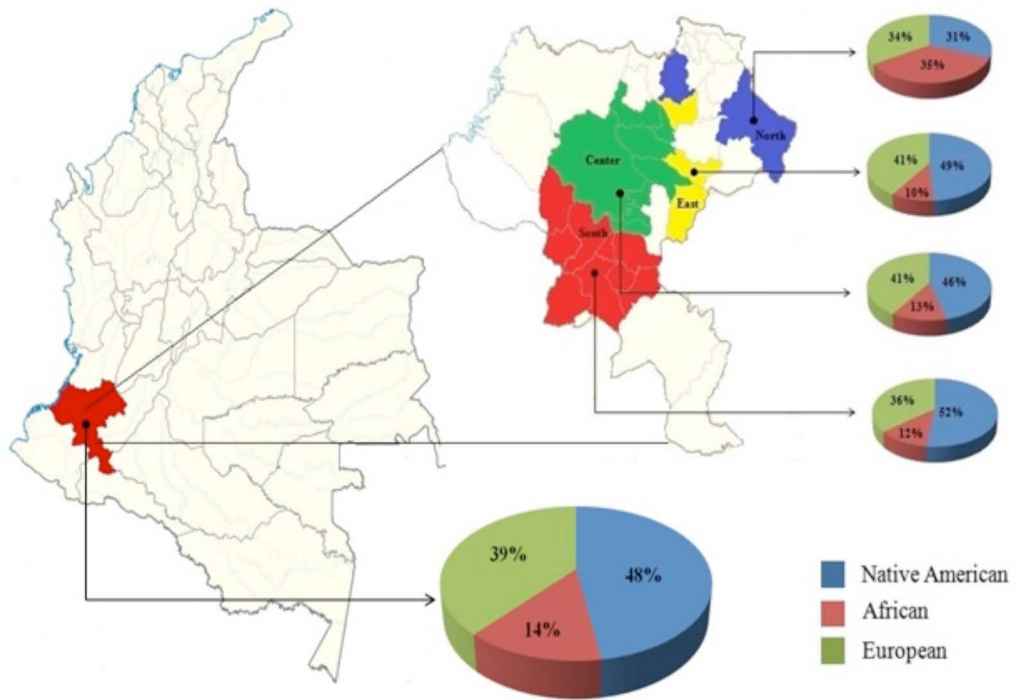


Figure 3. Map of regions studied from the Department of Cauca showing African, European and Native American ancestry contribution.

Principal Component Analysis Plot

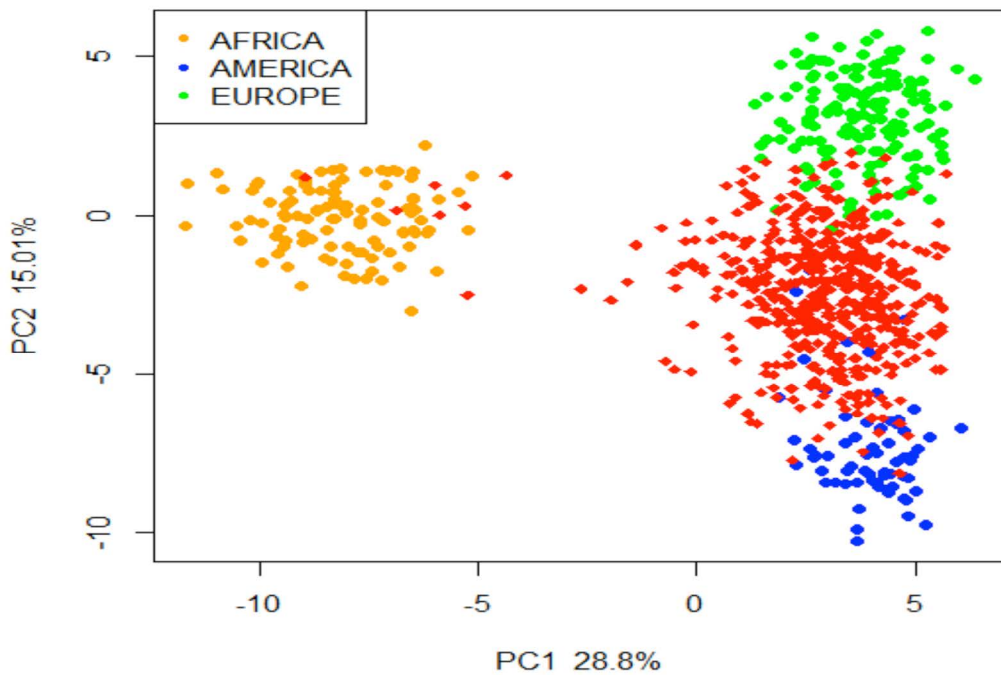


Figure 4. Principal Component Analysis, Biplot right side shows that the studied population is spread and compounded by the European and Amerindian origin, in contrast, in the biplot left side there is a minority population of African origin.

DISCUSSION

The estimation of the genetic composition of populations resulting from admixtures has been important in genetic analyses of many human populations (Lai *et al.*, 2009; Lee *et al.*, 2010; Lins *et al.*, 2010; Wang *et al.*, 2010; Avena *et al.*, 2012; Crouch *et al.*, 2012; Campos-Sánchez *et al.*, 2013; Heinz *et al.*, 2013; Shahabi *et al.*, 2013; Parolin *et al.*, 2014). A set of 46 AIMs-InDels reported by Pereira *et al.* (2012) was used to characterize the composition of genetic ancestry proportions of a population sample of 500 individuals from 23 out of 42 municipalities representing the central, northern, southern and eastern regions of the Department of Cauca (DANE, 2005). Individuals from western municipalities were not included in this study because jungle and maritime barriers limited the access to this part of the State. Thus, a random assessment of the genetic composition of the most representative regions from Cauca was accomplished.

The pairwise genetic distances analysis of the studied population showed a higher number of individuals closer to both Native American and European parental populations and a lower frequency closer to African population. The lower frequency of African population could be explained by the occurrence of cultural and geographic barriers which have resulted in the isolation of this ethnic group. These results are consistent with the admixture composition analysis conducted by STRUCTURE, which indicates that the ancestry genetic composition of the Cauca population is predominantly “Mestizo”, resulting from the genetic contributions of both Native American and European populations and, to a lesser extent, the African population. The observed differences in ancestry proportions are the result of the admixture by both, European and African populations with the existent Native American population of Cauca. Our findings are consistent with previous genetic composition studies carried out in individuals from several municipalities of Cauca using different AIMs. The study by Acosta *et al.* showed a high percentage of European contribution in the Caucasian-Mestizo population and a high influence of African ancestry in Afro-Colombian individuals (Acosta *et al.*, 2009). Furthermore, it has been indicated that the Native American and European populations have contributed to a greater proportion to the current gene pool of Cauca population than the African (Rojas *et al.*, 2010; Córdoba *et al.*, 2012).

A previous study conducted in Cauca suggest that the DNAm_t lineage and Y chromosome ancestry did not contribute in the same way to the admixture process, as only 72.8 % of Y chromosomes comes from an European origin, but just a 0.65 % of DNAm_t comes from such origin (Córdoba *et al.*, 2012). In this study, a 17 % paternal contribution against an 89 % maternal contribution in the Amerindian component was also observed. In the African people an opposite phenomenon is observed, as there is a greater paternal contribution. These results suggest that there is a sexual bias in the studied population admixture process due to interethnic reproduction mainly between European conquerors and native women, which could also explain the results found in the present study.

The Department of Cauca harbors the highest percentage of Native American population in Colombia. About 20 % of the Cauca population belongs to this ethnic group; moreover, there are eight officially recognized indigenous groups which are established in 26 of the 42 municipalities (Zaumsegel *et al.*, 2013). The admixture analysis in our study revealed greater proportions of Native Americans in both eastern and southern regions of Cauca. This distribution is explained by the fact that numerous Native Americans tribes have inhabited predominantly in the municipalities of the eastern region (Caldono, Puracé, Totoró) since the Pre-Colombian period. Furthermore, several Native American tribes (e.g. Yanacona) are located predominantly in the South of the Department. On the other hand, the admixture proportions of Africans were significantly higher in the northern Region compared to the others. Despite the fact that the North of Cauca is inhabited mainly by Afro-Colombians, Native American and “Mestizo” populations exhibit a wide ethnic diversity. The various hard-labor activities carried out in this region, such as mining and agriculture, contribute to the large concentration of male Afro-descendants linked to them (DANE, 2005).

The population of the Department of Cauca according to the estimates of the National Administrative Department of Statistics –based on the census carried out at the national level in 2005– is constituted by 1,330,756 inhabitants who are spread over a territory of 30,169 km². According to this census, 20.50 % of the population is classified as indigenous. However, our study reveals that according to InDels markers the Amerindian ancestry is 48 %, being twice of that reported by the government. On the other hand, our study reveals a 14 % of the population having an

African ancestry, while the government reports 21 %; and the European ancestry is not established by the government. Data discrepancies are obviously the result of self-reported ethnicity that people provide during census interviewing, while genetic studies provide a more precise approach to ancestry origin.

CONCLUSIONS

All populations presented identical European and Native American contributions, which were slightly higher than the African contribution. Therefore it was determined that the Department of Cauca has a mixed population.

BIBLIOGRAPHY

- Acosta M.A., Blanco-Verea A., Lareu M.V., Brion M., Carracedo A. (2009) The genetic male component of two South-Western Colombian populations. *Forensic Sci. Int. Genet.* 3: 59-61.
- Avena S., Via M., Ziv E., Pérez-Stable E.J., Gignoux C.R., Dejean C., Huntsman S., Torres-Mejía G., Dutil J., Matta J.L. (2012) Heterogeneity in genetic admixture across different regions of Argentina. *PloS one*: e34695.
- Bastos-Rodrigues L., Pimenta J.R., Pena S.D. (2006) The genetic structure of human populations studied through short insertion-deletion polymorphisms. *Ann. Hum. Genet.* 70: 658-665.
- Bland J.M., Altman D.G. (1995) Multiple significance tests: the Bonferroni method. *BMJ* 310: 170.
- Campos-Sánchez R., Raventós H., Barrantes R. (2013) Ancestry Informative Markers Clarify the Regional Admixture Variation in the Costa Rican Population. *Hum. Biol.* 85: 721-740.
- Chacón-Duque J.C., Adhikari K., Avendaño E., Campo O., Ramirez R., Rojas W., Ruiz-Linares A., Restrepo B.N., Bedoya G. (2014) African genetic ancestry is associated with a protective effect on Dengue severity in Colombian populations. *Infect. Genet. Evol.* 27: 89-95.
- Córdoba L., García J.J., Hoyos L.S., Duque C., Rojas W., Carvajal S., Escobar L.F., Reyes I., Cajas N., Sanchez A. (2012) Genetic Structure of a Southwestern Colombian Population. *RCA* 48: 21-48.
- Crouch D.J., Weale M.E. (2012) Inferring separate parental admixture components in unknown DNA samples using autosomal SNPs. *Eur. J. Hum. Genet.* 20: 1283-1289.
- DANE (2005) http://www.dane.gov.co/files/censo2005/etnia/sys/Afro_indicadores_sociodemograficos_censo2005.pdf (Accessed April 2014).
- DANE (2005) <http://www.dane.gov.co/index.php/poblacion-y-demografia/censos> (Accessed April 2014).
- Desjardins P., Conklin D. (2010) NanoDrop microvolume quantitation of nucleic acids. *JoVE*: e2565-e2565.
- Excoffier L., Lischer H.E. (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10: 564-567.
- Heinz T., Álvarez-Iglesias V., Pardo-Seco J., Taboada-Echalar P., Gómez-Carballa A., Torres-Balanza A., Rocabado O., Carracedo Á., Vullo C., Salas A. (2013) Ancestry analysis reveals a predominant Native American component with moderate European admixture in Bolivians. *Forensic Sci. Int. Genet.* 7: 537-542.
- Kosoy R., Nassir R., Tian C., White P.A., Butler L.M., Silva G., Kittles R., Alarcon-Riquelme M.E., Gregersen P.K., Belmont J.W. (2009) Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in America. *Hum. Mutat.* 30: 69-78.
- Lai C.Q., Tucker K.L., Choudhry S., Parnell L.D., Mattei J., García-Bailo B., Beckman K., Burchard E.G., Ordovás J.M. (2009) Population admixture associated with disease prevalence in the Boston Puerto Rican health study. *Hum. Genet.* 125: 199-209.
- LaRue B.L., Lagacé R., Chang C.W., Holt A., Hennessy L., Ge J., King J.L., Chakraborty R., Budowle B. (2014)

- Characterization of 114 insertion/deletion (INDEL) polymorphisms, and selection for a global INDEL panel for human identification. *Leg. Med.* 16: 26-32.
- Lee Y.L., Teitelbaum S., Wolff M.S., Wetmur J.G., Chen J. (2010) Comparing genetic ancestry and self-reported race/ethnicity in a multiethnic population in New York City. *J. Genet.* 89: 417-423.
- Lins T.C., Vieira R.G., Abreu B.S., Grattapaglia D., Pereira R.W. (2010) Genetic composition of Brazilian population samples based on a set of twenty eight ancestry informative SNPs. *Am. J. Hum. Biol.* 22: 187-192.
- Manta F.S.N., Pereira R., Caiafa A., Silva D.A., Gusmão L., Carvalho E.F. (2013) Analysis of genetic ancestry in the admixed Brazilian population from Rio de Janeiro using 46 autosomal ancestry-informative indel markers. *Ann. Hum. Biol.* 40: 94-98.
- Miller S., Dykes D., Polesky H. (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 16: 1215.
- Neves Manta F.S., Pereira R., Vianna R., Araújo A.R.B., Gitaí D.L.G., da Silva D.A., de Vargas Wolfgramm E., da Mota Pontes I., Aguiar J.I., Moraes M.O. (2013) Revisiting the genetic ancestry of Brazilians using autosomal AIM-InDels. *PloS one* 8: e75145.
- Parolin M.L., Carreras-Torres R., Sambuco L.A., Jaureguiberry S.M., Iudica C.E. (2014) Analysis of 15 autosomal STR loci from Mar del Plata and Bahía Blanca (Central Region of Argentina). *Int. J. Legal Med.* 128: 457-459.
- Pena S.D., Bastos-Rodrigues L., Pimenta J., Bydlowski S. (2009) DNA tests probe the genomic ancestry of Brazilians. *Braz. J. Med. Biol. Res.* 42: 870-876.
- Pena S.D., Di Pietro G., Fuchshuber-Moraes M., Genro J.P., Hutz M.H., Kehdy F.d.S.G., Kohlrausch F., Magno L.A.V., Montenegro R.C., Moraes M.O. (2011) The genomic ancestry of individuals from different geographical regions of Brazil is more uniform than expected. *PloS one* 6: e17063.
- Pepinski W., Abreu-Glowacka M., Koralewska-Kordel M., Michalak E., Kordel K., Niemcunowicz-Janica A., Szeremeta M., Konarzewska M. (2013) Population genetics of 30 InDels in populations of Poland and Taiwan. *Mol. Biol. Rep.* 40: 4333-4338.
- Pereira R., Phillips C., Alves C., Amorim A., Carracedo Á., Gusmão L. (2009) A new multiplex for human identification using insertion/deletion polymorphisms. *Electrophoresis* 30: 3682-3690.
- Pereira R., Phillips C., Pinto N., Santos C., dos Santos S.E.B., Amorim A., Carracedo Á., Gusmão L. (2012) Straightforward inference of ancestry and admixture proportions through ancestry-informative insertion deletion multiplexing. *PloS one* 7: e29684.
- Pritchard J.K., Stephens M., Donnelly P. (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.
- Rojas W., Parra M.V., Campo O., Caro M.A., Lopera J.G., Arias W., Duque C., Naranjo A., García J., Vergara C. (2010) Genetic make up and structure of Colombian populations by means of uniparental and biparental DNA markers. *Am. J. Phys. Anthropol.* 143: 13-20.
- Schneider S., Lischer H.E. (2009) Arlequin, version 3.5: An integrated software package for population genetics data analysis. University of Berne: 1-145.
- Shahabi A., Wilson M.L., Lewinger J.P., Goodwin T.M., Stern M.C., Ingles S.A. (2013) Genetic admixture and risk of hypertensive disorders of pregnancy among Latinas in Los Angeles County. *Epidemiology* 24: 285-294.
- Shi M., Liu Y., Bai R., Jiang L., Lv X., Ma S. (2015) Population data of 30 insertion-deletion markers in four Chinese populations. *Int. J. Legal Med.* 129: 53-56.
- Shriver M.D., Smith M.W., Jin L., Marcini A., Akey J.M., Deka R., Ferrell R.E. (1997) Ethnic-affiliation estimation by use of population-specific DNA markers. *Am. J. Hum. Genet.* 60: 957.
- Tishkoff S.A., Reed F.A., Friedlaender F.R., Ehret C., Ranciaro A., Froment A., Hirbo J.B., Awomoyi A.A.,

- Bodo J.M., Doumbo O. (2009) The genetic structure and history of Africans and African Americans. *Science* 324: 1035-1044.
- Wang H., Haiman C.A., Kolonel L.N., Henderson B.E., Wilkens L.R., Le Marchand L., Stram D.O. (2010) Self-reported ethnicity, genetic structure and the impact of population stratification in a multiethnic study. *Hum. Genet.* 128: 165-177.
- Weber J.L., David D., Heil J., Fan Y., Zhao C., Marth G. (2002) Human diallelic insertion/deletion polymorphisms. *Am. J. Hum. Genet* 71: 854-862.
- Zaumsegel D., Rothschild M.A., Schneider P.M. (2013) A 21 marker insertion deletion polymorphism panel to study biogeographic ancestry. *Forensic Sci. Int. Genet.* 7: 305-312.

ACKNOWLEDGEMENTS

We are indebted to all the volunteers who participated in the study from the municipalities of El Tambo, Puracé, Silvia, Mercaderes, Rosas, Timbio, Bolivar, La Sierra, La Vega, Cajibío, Argelia, Tunia, Sotará, Piendamó, Miranda, El Bordo, Santander de Quilichao, Totoró, Almaguer, Inzá, Puerto Tejada, Balboa, Morales, San Sebastián, Belalcazar, Coconuco, Sucre, and Popayán. Finally, we acknowledge the collaboration of the administrative personnel of the Human Genetics Laboratory and the Vice-presidency for Research of the University of Cauca.

STATISTICAL MODELS FOR PHENOTYPE-GENOTYPE ASSOCIATION STUDIES IN GENETICALLY STRUCTURED POPULATIONS

MODELOS ESTADÍSTICOS PARA ESTUDIOS DE ASOCIACIÓN FENOTIPO-GENOTIPO EN POBLACIONES GENÉTICAMENTE ESTRUCTURADAS

Peña Malavera A.¹, Gutierrez L.², Balzarini M.^{1*}

¹Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). Estadística y Biometría, Facultad de Ciencias Agropecuarias, Universidad Nacional de Córdoba. Av. Valparaíso s/n, Ciudad Universitaria, CP: 5000 (509) Córdoba, Argentina.

²Department of Agronomy, University of Wisconsin at Madison, 1575 Linden Dr. Madison, WI 53706.

*Corresponding author: mbalzari@agro.unc.edu.ar

ABSTRACT

Association mapping is used to find specific regions in the genome related to changes in a phenotypic trait. However, it has been found that in genetically structured populations, the number of false positives increases. The aim of this study was to compare the performance of several association mapping statistical models that take into account the underlying population genetic structure. Different statistical strategies developed under the mixed model theory were evaluated. The compared association models included the following matrices to model genetic structure: Q-matrix (probability of membership of each individual to each subpopulation), P-matrix (principal components of marker data capturing the structure variance) and K-matrix (containing genetic relationships between the individuals of the mapping population). The columns of Q-matrix and P-matrix were used in the associative mapping model as fixed effect covariates as well as random effect covariates. We also evaluated models including simultaneously Q-matrix and K-matrix, or either as P-matrix and K-matrix. The reference model (naïve model) was a regression model that did not account for genetic structure. Model comparison criteria were the empirical distributions of p-values, the FDR (False Discovery Rate) and the statistical power. The results suggest that the use of the K-matrix, alone or together with the Q-matrix reduced the false positive rate regardless of the level of genetic divergence among underlying subpopulations.

Key words: linear mixed models, population genetic structure, false discovery rate.

RESUMEN

El mapeo asociativo (MA) es usado para encontrar regiones específicas del genoma relacionadas con la variación de un carácter fenotípico. Sin embargo, se ha detectado que en poblaciones con estructura genética poblacional (EGP), la cantidad de falsos positivos en la asociación fenotipo-genotipo aumenta. El objetivo de este trabajo fue evaluar el desempeño de modelos de MA que consideran la EGP mediante distintas estrategias desarrolladas bajo la teoría de los modelos mixtos. Se evaluaron modelos de regresión fenotipo-genotipo incluyendo las siguientes matrices para modelar EGP: matriz Q (probabilidad de pertenencia de cada individuo a cada subpoblación), matriz P (componentes principales de los datos de marcadores), matriz K o de parentesco genético entre las líneas de la población de mapeo. Las columnas de las matrices Q y P fueron usadas en el modelo de MA como covariables de efecto fijo y alternativamente, como efectos aleatorios. También se evaluaron modelos incluyendo simultáneamente las matrices Q y K, así como P y K. El modelo de referencia (“naïve”) fue el modelo de regresión que no contempló EGP. Los criterios de comparación de modelos fueron la función de distribución empírica de valores-p, la tasa FDR (*False Discovery Rate*) y la potencia estadística. Los resultados sugieren que el uso de la matriz K, sola o junto con la matriz Q, fue la estrategia de mayor impacto para disminuir la tasa de detección de falsas asociaciones. Esto se observó independientemente del nivel de divergencia genética, entre las subpoblaciones que constituían la población de mapeo.

Palabras clave: modelos lineales mixtos, estructura genética poblacional, tasa de falsos positivos.

Fecha de recepción: 22/03/2015
Fecha de aceptación de versión final: 18/05/2016

INTRODUCCIÓN

En los últimos años se ha incrementado el uso del mapeo asociativo (MA) o mapeo por desequilibrio de ligamiento (LD, del inglés *Linkage disequilibrium*) para la identificación de regiones del genoma responsables de características complejas de interés agronómico. La técnica ha sido ampliamente adoptada en el mejoramiento de especies vegetales para el análisis de los *loci* de caracteres cuantitativos o QTL (Aranzana *et al.*, 2005; Breseghello y Sorrells, 2006; D'hoop *et al.*, 2008; Kraakman *et al.*, 2006; Remington *et al.*, 2001; Stich *et al.*, 2008; Thornsberry *et al.*, 2001; Zhu *et al.*, 2008). Cuando la población de individuos empleada en el análisis de mapeo por LD está estructurada genéticamente, aumenta la cantidad de falsos positivos en la detección de las asociaciones de interés (Malosetti *et al.*, 2007). Esto ocurre porque en una población genética con sub-poblaciones, cualquier carácter presente con mayor frecuencia en una de ellas mostrará asociación positiva con alelos que son más comunes en esta sub-población (Zhang *et al.*, 2010). Consecuentemente, es posible que se detecten marcadores asociados con la composición de la población más que con la característica de interés (Yu *et al.*, 2006). Por ello, se han propuesto distintas estrategias de modelado para los estudios de asociación fenotipo-genotipo, todas tendientes a controlar el aumento en la detección de asociaciones espurias. El modelo de MA básico es un modelo de regresión lineal, donde el fenotipo se asocia al genotipo (marcadores moleculares) mediante coeficientes de regresión. Este modelo básico es luego extendido con el fin de incorporar factores o covariables que representan la estructura genética subyacente en la población de mapeo (Cappa *et al.*, 2013; Muñoz-Amatriáin *et al.*, 2014; Wang *et al.*, 2012). Cuando se usa el programa *Structure* (Pritchard *et al.*, 2000) como herramienta para detectar estructura genética poblacional (EGP), se obtienen las probabilidades de pertenencia de cada individuo a las sub-poblaciones que componen la población y esta información puede ser incorporada al modelo de MA (Gutiérrez *et al.*, 2011; Yu *et al.*, 2006). Las columnas de la matriz Q resultante del análisis con el programa *Structure* (probabilidades de pertenencia a los distintos grupos o conglomerados) suelen ser usadas como covariables en el modelo de regresión fenotipo-genotipo. Estas covariables proveen información que dimensiona la EGP. Alternativamente se suele usar otra matriz, conocida como matriz P, compuesta por los componentes principales resultantes del análisis de

componentes principales (ACP) (Hotelling, 1936) realizado con la matriz de datos de marcadores genéticos (Price *et al.*, 2006). Los componentes principales (CP) significativos, según la prueba de Tracy-Widom (1994), o los primeros CP (los que explican mayor porcentaje de la variabilidad total de los datos moleculares, pueden también ser usados como covariables en el modelo de MA (Peña Malavera *et al.*, 2014). Las covariables que resumen la estructuración genética de los genotipos de la población de mapeo pueden contemplarse en el modelo como efectos fijos o aleatorios. Malosetti *et al.* (2007) recomiendan incluir la estructura genética subyacente en las poblaciones de mapeo como efecto aleatorio, independientemente del procedimiento utilizado para detectar dicha estructura. Otra estrategia ya difundida para contemplar relaciones genéticas durante el MA es la corrección de errores estándares asociados a los coeficientes de regresión mediante la incorporación de la matriz de parentesco genético (matriz K) en la modelación de la estructura de covarianza residual. La matriz K puede ser obtenida con la librería EMMA del programa R (www.Rproject.org) (Kang *et al.*, 2008) como la matriz de covarianza asociada a un vector aleatorio de efectos poligénicos adicionados a cada genotipo en estudio. Usualmente, la matriz K es una matriz de similitudes entre los perfiles moleculares individuales de distintos genotipos. Cualquiera sea la estrategia seleccionada para contemplar la EGP, matriz Q, P o K, la estimación del modelo puede realizarse bajo el marco teórico de los modelos lineales mixtos (MLM) (Peña Malavera, 2015). Sin embargo, poco se ha investigado sobre el impacto de estos modelos alternativos respecto a la detección de QTL en situaciones donde existan bajos o casi nulos niveles de ligamiento como sucede en colecciones de germoplasma. El objetivo de este trabajo fue comparar el desempeño, a nivel de las tasas de falsos positivos y también a nivel de potencia estadística para la detección de QTL, de modelos de MA alternativos bajo escenarios de baja y alta estructuración genética. También, se comparan resultados obtenidos en contextos con distinta cantidad de información con dos niveles de densidad de marcadores (baja y alta). Los modelos comparados fueron: modelos Q y QA (usan matriz Q como covariables de efecto fijo o aleatorio, respectivamente), modelos P y PA (usan matriz P como covariables de efecto fijo o aleatorio, respectivamente), modelo K, QK y PK; estos dos últimos incluyen simultáneamente dos matrices de control de EGP. En cada escenario también se ajustó como modelo de referencia, uno que no contempla corrección por EGP

(Modelo *naive*). La comparación de modelos se realizó usando bases de datos de marcadores moleculares simulados bajo dos escenarios, uno con 300 marcadores moleculares y tamaño poblacional de 150 genotipos los cuales contenían distintos niveles de EGP, y otro con una base de datos experimentales con 511 marcadores moleculares y 504 genotipos de maíz genéticamente estructurados (Hansey *et al.*, 2011).

MATERIALES Y MÉTODOS

Simulaciones

Los datos de marcadores moleculares fueron simulados usando QMSim (Sargolzaei y Schenkel, 2009) dando origen a dos escenarios con cantidad de genotipos y de marcadores moleculares que imitan datos usuales en mejoramiento genético vegetal. Se simuló un genoma con 300 marcadores multilocus-bialélicos, con diseño de cruzamientos y selección aleatorios para una EGP conformada por cinco poblaciones. Se crearon dos escenarios correspondientes a dos niveles de divergencia genética entre poblaciones (bajo y alto F_{ST}), con un tamaño de población de mapeo ($n=150$). Los datos simulados fueron creados a partir de una población de 200 individuos y el sistema de cruzamiento fue basado en la unión aleatoria de gametas por más de 10 generaciones. La coancestría promedio fue baja como sucede en numerosas poblaciones que se usan para MA en vegetales. Al variar el número de generaciones desde la población fundadora, se crearon diferentes niveles de divergencia genética poblacional. Los datos simulados para cada marcador fueron codificados como 0 y 1. El promedio del estadístico F_{ST} (Wright, 1951), provisto por el análisis molecular de la varianza (AMOVA) (Excoffier *et al.*, 2009), fue usado para cuantificar el grado de diferenciación genética entre poblaciones en cada escenario con bajo y alto F_{ST} ($F_{ST} = 0,03$ y $0,20$, respectivamente).

Dada la matriz de marcadores moleculares resultante de QMSim, para cada uno de los dos escenarios, se escogieron aleatoriamente 20 marcadores y con ellos se realizó una combinación lineal con efectos que siguen una distribución $\Gamma(2,5)$ para simular el efecto de los *loci* ligados a un QTL, es decir aquellos con información para determinar el fenotipo. Adicionalmente, se anexó a cada perfil molecular la realización de una variable aleatoria con distribución normal de media 100 (para representar la media del carácter, la cual depende del efecto poligénico de

background) y varianza 25 (para representar la variabilidad experimental, *i.e.* desvío estándar de 5 es decir no superior al 5 % de la media del carácter fenotípico). A esta variable $\sim N(100, 25)$ se le sumaron los efectos de los marcadores ligados extraídos de la distribución gamma.

Datos experimentales

Los modelos de MA también se evaluaron en un conjunto de datos públicos conformados por $n=504$ líneas de maíz genotipificadas con $p=511$ marcadores del tipo SNPs (Hansey *et al.*, 2011). Hansey *et al.* (2011) identificaron ocho conglomerados o sub-poblaciones en ese conjunto de líneas de maíz, datos que fueron luego verificados por investigaciones genéticas en esa especie. Consecuentemente, los análisis sobre los datos genéticos reales se hicieron asumiendo la existencia de ocho conglomerados. Para implementar el MA se simuló para cada uno de estos genotipos un valor fenotípico adicionando una variable aleatoria $\sim N(100, 25)$ y una variable aleatoria gamma $\Gamma(4, 2)$, asociada a cada uno de 22 marcadores moleculares elegidos aleatoriamente.

Modelos estadísticos ajustados

Se estimaron ocho modelos de mapeo asociativo para evaluar el efecto del marcador sobre el fenotipo cuya denotación se presentan en la Tabla 1. Los modelos de MA comparados surgen de usar algunos o todos los términos de la siguiente ecuación:

$$y = X\beta + EGPv + Zu + e$$

donde y es el vector de valores fenotípicos (con un dato fenotípico por genotipo), X es el vector de datos para el marcador molecular, β es el efecto del marcador, EGP es la matriz de estructura genética (construida alternativamente como la matriz Q de la salida del programa *Structure* o la matriz P de los componentes principales estadísticamente significativos, ambos realizados previamente sobre los datos moleculares), v es el vector de efectos de la estructura poblacional (en algunas aproximaciones considerado como vector de efectos fijos y en otras como vector de efectos aleatorios), Z es la matriz de incidencia que conecta el vector aleatorio u de efectos poligénicos con los datos fenotípicos (matriz identidad de dimensión igual al número de genotipos que componen la población de mapeo) y e es un vector de términos de error aleatorio, normalmente distribuido con media cero y varianza constante σ_e^2

El vector u se distribuye independientemente del vector e y con matriz de varianzas y covarianzas dada por $\sigma_e^2 \times K$, siendo K la matriz de similitud entre los pares de perfiles moleculares derivadas de la librería EMMA (Kang *et al.*, 2008).

Tabla 1. Ocho modelos comparados en datos reales y simulados.

Matriz de Parentesco	Covariable de Estructura				
	No	Q fijo	ACP fijo	Q aleatorio	ACP aleatorio
No	naive	Q	P	QA	PA
Si	K	QK	PK	--	--

Q es la matriz de probabilidades de pertenencia a los grupos de la EGP, probabilidad calculada con el software *Structure*, P es la matriz de componentes principales retenida mediante el estadístico de Tracy-Widom (1994) y K es la matriz de parentesco propuesta por Kang *et al.* (2008).

Ajuste de modelos y criterios de comparación

Todos los modelos fueron ajustados usando *Info-Gen* (Balzarini y Di Rienzo, 2004) y su interfaz con R (Core Team, 2013). El desempeño de los modelos se evaluó usando como criterio las curvas de distribución acumulada de valores-p (Gutierrez *et al.*, 2011; Cappa *et al.*, 2013; Peña Malavera *et al.*, 2014; Peña Malavera, 2015), las tasas de falsos descubrimientos o FDR (del inglés, *False Discovery Rate*) (Benjamini y Hochberg, 1995) y los cálculos de potencia estadística.

Para construir las curvas de distribución de valores-p, se usó la función de distribución empírica de la variable “valor-p asociado a cada una de las pruebas de hipótesis realizadas en un escenario”. En cada escenario hay tantas pruebas de hipótesis de asociación como marcadores. Es importante resaltar que en una distribución acumulada de valores-p se espera que si la modelación ha sido buena, la distribución se aproxime a una línea recta de 45 grados, ya que la distribución de los valores-p debe ser simétrica. Una distribución asimétrica hacia valores-p pequeños indica mayor significancia de la esperada, lo que sugiere un posible incremento de falsos positivos, es decir presencia de asociaciones espurias.

La tasa FDR se calculó en base a las proporciones de falsos positivos (FP) y verdaderos positivos (VP). Los FP son todos aquellos valores-p significativos vinculados a marcadores que no están asociados al fenotipo (no ligados a un QTL) y los VP son todos aquellos marcadores positivos que efectivamente están asociados al fenotipo (ligados a un QTL), de esta forma tenemos que:

$$FDR = \frac{FP}{VP + FP}$$

La potencia estadística en la detección de marcadores asociados con el fenotipo está referida a una medida de eficacia de los modelos y es la probabilidad de que la hipótesis nula H_0 sea rechazada cuando esta es falsa, o dicho de otra manera cuando la hipótesis alternativa H_a es verdadera. La potencia estadística (Φ) puede interpretarse como la probabilidad de no cometer error del tipo II (error que producen los eventos conocidos como falsos negativos, FN). La potencia en este trabajo fue calculada de la siguiente manera:

$$\Phi = \frac{VP}{VP + FN}$$

RESULTADOS

Datos genéticos simulados

En la Figura 1 se muestran los gráficos de dispersión de los dos primeros ejes resultantes del escalamiento multidimensional métrico (Gower, 1967) obtenido desde los datos moleculares observados bajo cada uno de los dos escenarios con diferentes niveles de F_{ST} , con 300 marcadores moleculares. La figura proporciona información con respecto a la distancia genética entre los genotipos y el grupo al que estos genotipos fueron asignados. En el escenario I se observa baja divergencia genética mientras que en el escenario II, los grupos o subpoblaciones que estructuran la población se presentan más distanciados por la mayor divergencia genética con la que fueron simulados.

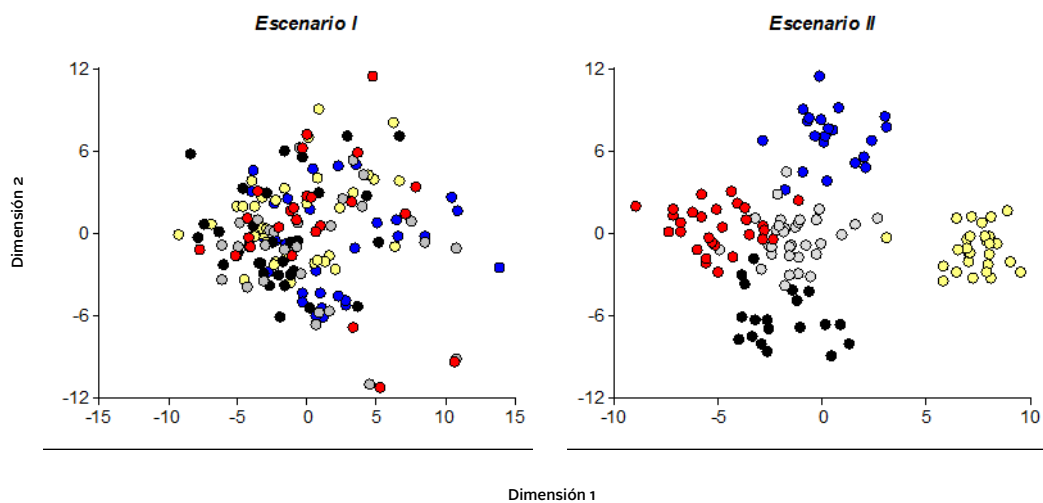


Figura 1. Gráficos de dispersión de los dos primeros ejes resultantes de un análisis de coordenadas principales (escalamiento multidimensional) de los datos moleculares (300 MM). En la columna de la izquierda bajo F_{ST} , y en la columna derecha alto F_{ST} . Los colores identifican los cinco grupos que definen la EGP.

Evaluación de modelos de mapeo asociativo

En la Figura 2 se muestran las funciones de distribución acumulada para los dos escenarios. Se puede ver que consistentemente en los dos escenarios con diferentes niveles de F_{ST} , los modelos con mejor ajuste son el modelo K y el modelo QK.

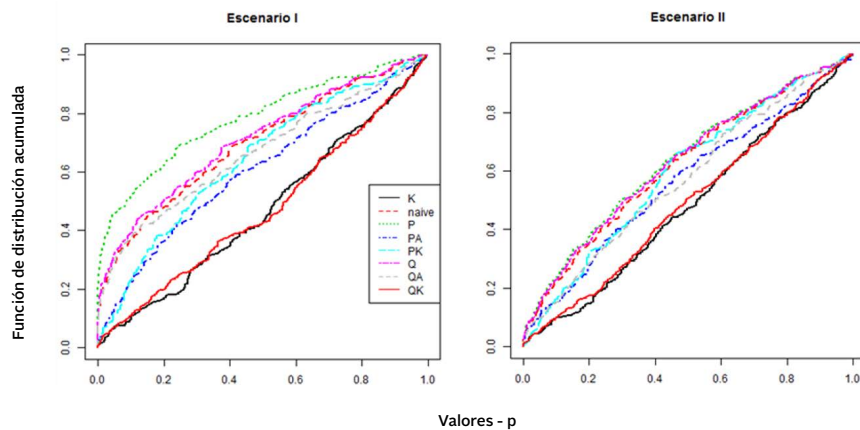


Figura 2. Gráfico de distribución acumulada de los valores-p de los ocho modelos evaluados en los escenarios simulados. En la columna de la izquierda escenarios con F_{ST} , bajo y en la columna derecha F_{ST} , alto.

Con bajo F_{ST} (izquierda, escenario I) se observa que el modelo de menor desempeño fue el modelo P (Figura 2). Los modelos se comportan de manera parecida cuando se trabaja en el contexto de alto F_{ST} (escenario II), situación en la que se observan menores diferencias entre los ajustes, principalmente para el caso de mayor estructuración relativa en la población, correspondiente a un valor de F_{ST} de 0,20 y a una población de 150 individuos.

Existen para el modelo P una cantidad mayor a la esperada de valores-p pequeños, es decir mayor probabilidad de detecciones falsas. Los modelos *naive* y Q también mostraron alta asimetría hacia pequeños valores-p. Con bajo nivel de EGP no se desempeñaron bien los modelos que incorporaron más parámetros para modelar su estructura de media (efectos fijos). Los modelos que incluyen la matriz K y los que incluyen covariables de efectos aleatorios para incorporar la estructura, fueron los de mejor desempeño. Con alto nivel de divergencia genética, QA se desempeñó mejor que PA mientras que con bajo nivel de divergencia la relación entre ambos fue inversa. Entre los modelos que usan la matriz K y que también incorporan covariables de efectos fijos, el modelo QK mostró menor FDR y una función de distribución acumulada de valores-p más cercana a la esperada bajo la

hipótesis nula. Es importante notar que en ninguno de los escenarios simulados la estrategia de modelado conjunto mejoró el desempeño del modelo que sólo usa la matriz K en términos de valores-p.

En la Figura 3 se muestran las funciones de distribución acumulada de los valores-p correspondiente a las pruebas de hipótesis implementadas para el conjunto de datos experimentales de maíz (Hansey *et al.*, 2011). Se observa que los modelos, excepto por PK, tienen un comportamiento relativo similar. Todos muestran menor cantidad de la esperada de valores-p, comportándose similar al modelo que no implementa ninguna corrección por estructura (modelo *naive*). Aún cuando se ha reconocido previamente la existencia de ocho grupos en los datos moleculares, los distintos modelos de MA ajustados sobre estos datos no provocaron significativos cambios en los valores-p (a excepción de PK que sobredimensionó la cantidad de pequeños valores-p). La estrategia de modelación más diferente del modelo *naive*, en este dominio de valores-p, fue el modelo QK. Estos resultados podrían estar asociados al bajo nivel de divergencia genética existente entre las ocho subpoblaciones ($F_{ST} = 0,02$), o a una variabilidad relativamente alta de los datos fenotípicos que enmascara el efecto de los QTL.

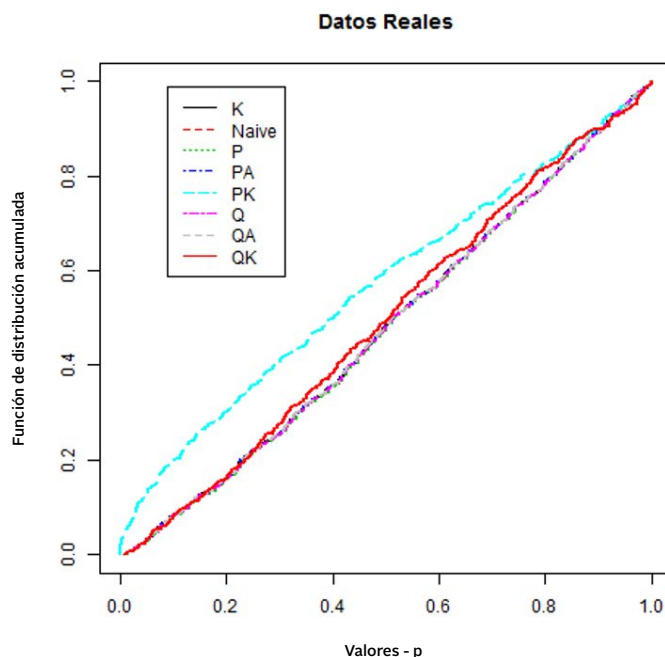


Figura 2. Gráfico de distribución acumulada de los valores-p de los ocho modelos evaluados en datos de 511 marcadores SNP sobre 504 genotipos de maíz estructurados genéticamente con un nivel de F_{ST} relativamente bajo ($F_{ST} = 0,02$).

FDR Y POTENCIA

Datos simulados

En la Tabla 2 se observa que los modelos que involucran la matriz K en el análisis, tienen los menores valores de FDR para los dos niveles de F_{ST} , mientras que la corrección por estructura con la matriz P incrementó el porcentaje de falsos descubrimientos de asociaciones con respecto al modelo *naive* (34 % vs. 30 %).

Como es de esperar por las relaciones teóricas existente entre los errores tipo I y tipo II de las pruebas de hipótesis estadísticas (Balzarini *et al.*, 2012), los métodos que mejor controlan FDR son los que más potencia pierden o los que tienen menor probabilidad de detectar mayor cantidad de QTL verdaderos. La introducción de la matriz K en

el modelo de MA, redujo en aproximadamente un tercio para el caso de bajo F_{ST} y en un medio para el caso de mayor F_{ST} , la probabilidad de detecciones falsas.

En escenarios de alta divergencia genética, todas las aproximaciones metodológicas representaron una mejora a nivel de FDR respecto al modelo *naive*, algunos sin pérdida de potencia y otros (los que introducen la matriz K) con pérdida de potencia. No obstante, el modelo QK fue el de menor pérdida relativa de potencia si se considera la disminución de FDR que provocó.

Para los escenarios de bajo F_{ST} , el modelo K fue el que mejor se desempeñó, reduciendo significativamente los valores de FDR respecto al modelo *naive* con la consecuente pérdida de potencia, aunque ésta no fue mayor que para otros modelos con menor impacto sobre la FDR.

Tabla 2. Tasas de falsos positivos y potencia de ocho modelos de mapeo asociativo para dos niveles de estructura genética poblacional (Bajo y Alto F_{ST}), 300 marcadores moleculares y $n=150$.

Modelo	FDR		Potencia	
	Bajo F_{ST}	Alto F_{ST}	Bajo F_{ST}	Alto F_{ST}
Naive	0,30	0,24	0,60	0,55
Q	0,31	0,20	0,65	0,50
P	0,34	0,22	0,85	0,50
K	0,12	0,11	0,35	0,25
QK	0,13	0,11	0,35	0,30
PK	0,22	0,18	0,35	0,30
QA	0,32	0,18	0,55	0,35
PA	0,25	0,20	0,35	0,35

Naive: sin corrección por estructura; Q: con corrección mediante la matriz de probabilidades *a posteriori* obtenida con el software *Structure*; P: con corrección por CP como covariables de efectos fijos; K: con corrección por matriz de parentesco; QK: modelo mixto con Q como factor de efectos fijos y K factor de efectos aleatorios; PK: modelo mixto con P como factor de efectos fijos y K factor de efectos aleatorios; QA: modelo con la matriz Q como efectos aleatorios; PA: con la matriz P como covariables de efectos aleatorios.

Estructura genética real

La cantidad de componentes retenidos por Tracy-Widom para estos datos fue de 36; este alto número es necesario para explicar una EGP con bajo nivel de divergencia, pero a su vez produce una sobre-parametrización del modelo PK y quita grados de libertad a la estimación de la varianza residual con las consecuencias negativas observadas a nivel de la detección de QTL. Para los otros modelos se observa una sub-detección de QTL, siendo los modelos K, QK y los que incluyen la EGP en la porción aleatoria del modelo de MA, los de mejor desempeño. La baja divergencia genética de los datos de marcadores ($F_{ST} = 0,02$) podrían explicar estos resultados.

DISCUSIÓN

Se observó que todos los métodos, excepto Q y P, contribuyen a bajar la FDR. El uso de la matriz K fue la estrategia estadística de mayor impacto para bajar la FDR, tanto usada sola como con las correcciones por estructura genética Q y P. No hubo diferencia entre las tasas de FDR para alto y bajo F_{ST} , FDR es una característica más asociada al modelo de MA (parametrización de las estructuras de medias y varianzas que realiza cada modelo) que al nivel de estructura subyacente. Sin embargo se puede ver que el nivel de estructura impacta en la potencia del modelo. Con un nivel alto de estructura todos los modelos excepto P y Q, si bien controlaron la tasa de falsos positivos (FP) respecto al modelo *naive*, mostraron pérdida de potencia. El método PK que combina la matriz P proveniente del ACP y la matriz K de parentesco, fue el de menor pérdida de potencia en los escenarios simulados. No obstante, PK mostró en todos los escenarios menor capacidad de controlar FDR que K y QK. Aún, cuando en los escenarios de bajo F_{ST} los modelos P y Q mostraron más potencia para detectar QTL que el modelo sin corrección (*naive*), su incapacidad para reducir la tasa de FP no los hace recomendables.

Las funciones de distribución de los valores-p estimadas en este trabajo sugieren que los modelos K y QK son los de mejor desempeño en la identificación de asociaciones fenotipo-genotipo cuando se consideran tanto las tasas de error tipo I como las tasa de error tipo II. Este resultado coincide con los presentados por Yu *et al.* (2006) quienes trabajaron con tres variables fenotípicas medidas en 277 líneas endocriadas de maíz genotipadas con 553 SNP en

un escenario de estructura genética media a alta. El modelo QK fue también el elegido para el MA realizado con los datos moleculares reales considerados en este trabajo por Hansey *et al.* (2011). El modelo QK propuesto por Yu *et al.* (2006) fue inicialmente usado en poblaciones humanas y en especies alógamas. No obstante, Stich y Mechinger (2009) probaron este modelo en especies con distintos sistemas de reproducción usando colecciones de germoplasma no sólo de maíz sino también de papa, remolacha, nabo y arabisopsis, concluyendo que el modelo QK había resultado apropiado para todas las especies evaluadas.

Otros estudios han mostrado, al igual que los resultados observados en el presente trabajo, que en contextos de bajo LD y escasa estructuración genética, la incorporación de los CP como efecto fijo (P) puede producir una sobre-parametrización del modelo que conduce a un incremento en la tasa de FP (Peña Malavera *et al.*, 2014). En nuestro estudio, al igual que lo informado por Malosetti *et al.* (2007) y Gutiérrez *et al.* (2011), el modelo P (modelo de efectos fijos de CP que representan estructura) produjeron un gran número de falsos positivos o asociaciones espurias. Este resultado debería ser frecuente en situaciones donde la estructura poblacional es de poca magnitud y el número de componentes necesario para resumirla, de manera estadísticamente significativa, resulta alto. En tal contexto, el modelo lineal a estimar usa una cantidad alta de parámetros para describir poca variabilidad entre los genotipos moleculares. Sin embargo, Cappa *et al.* (2013) al estudiar el comportamiento de diversos modelos de MA en poblaciones de eucalipto de Argentina y Uruguay concluyeron que, para cuatro de los seis caracteres estudiados, el modelo P resultó una estrategia buena para controlar la marcada estructura familiar que existía entre los genotipos. Es importante notar que la EGP era alta y fue bien identificada en el trabajo con eucaliptus; tanto con *Structure* como con ACP donde los dos primeros componentes principales explicaron cerca del 50% de la variabilidad de los perfiles moleculares. Sólo con dos CP se pudo ordenar claramente los genotipos en tres grupos esperados según el conocimiento previo de sus procedencias. Con alto nivel de EGP y pocos CP, el modelo P puede resultar una buena estrategia para la detección de QTL verdaderos.

La incorporación de los componentes principales como covariables de efectos aleatorios (modelo PA) disminuyó también la tasa de falsos positivos respecto a los modelos

naive y P, pero con menor pérdida de potencia que K y QK. Gutiérrez *et al.* (2011) probaron métodos de corrección de la estructura poblacional, incluyendo entre ellos al modelo PA encontrando comportamientos similares entre este modelo y el modelo Q que usaba como covariables las columnas de salida del programa *Structure* (Pritchard *et al.*, 2000). En este trabajo, el modelo PA siempre controló mejor la tasa FDR que el modelo Q. Es de destacar que la implementación del control por estructura con análisis de componentes principales es significativamente más eficiente en tiempo computacional que el uso de información resultante del programa *Structure*. Wang *et al.* (2012) también mostraron la efectividad del modelo con componentes principales aleatorias (PA) para contemplar la estructura poblacional previa al MA y disminuir los FP.

BIBLIOGRAFÍA

- Aranzana M.J., Kim S., Zhao K., Bakker E., Horton M., Jakob K., Lister C., Molitor J., Shindo C., Tang C., Toomajian C., Traw B., Zheng H., Bergelson J., Dean C., Marjoram P., Nordborg M. (2005) Genome-wide association mapping in *Arabidopsis* identifies previously known flowering time and pathogen resistance genes. *PLoS Genet.* 1 (5): e60.
- Balzarini M., Di Rienzo J. (2004) Info-Gen: Software estadístico para análisis de datos genéticos. Universidad Nacional de Córdoba, Córdoba.
- Balzarini M., Di Rienzo J., Tablada M., Gonzalez L., Bruno C., Córdoba M., Robledo W., Casanoves F. (2012) Estadística y Biometría. Córdoba: Encuentro Grupo de Editores. pag.380. ISBN 978-987-591- 301-1
- Benjamini Y., Hochberg Y. (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B* 57: 289-300.
- Breseghele F., Sorrells M. (2006) Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. *Genetics* 172: 1165-1177.
- Cappa E.P., El-Kassaby Y.A., Garcia M.N., Acuña C., Borralho N.M.G., Grattapaglia D., Marcucci Poltri S. (2013) Impacts of population structure and analytical models in genome-wide association studies of complex traits in forest trees: A Case Study in *Eucalyptus globulus*. *PLoS ONE* 8 (11): e81267.
- Core Team (2013) R: A language and environment for statistical computing. Computing (Editor), Vienna, Austria.
- D'hoop B., Paulo M., Mank R., Eck H., Eeuwijk F. (2008) Association mapping of quality traits in potato (*Solanum tuberosum* L.). *Euphytica* 161(1-2): 47-60.
- Excoffier L., Hofer T., Foll M. (2009) Detecting *loci* under selection in a hierarchically structured population. *Heredity* 103 (4): 285-298.
- Gower J.C. (1967) Multivariate analysis and multidimensional geometry. *Journal of the Royal Statistical Society, Series D (The Statistician)* 17 (1): 13-28.
- Gutiérrez L., Cuesta-Marcos A., Castro A., von Zitzewitz J., Schmitt M., Hayes P. (2011) Association mapping of malting quality quantitative trait *loci* in winter barley: positive signals from small germplasm arrays. *Plant Gen.* 4 (3): 256-272.
- Hansey C.N., Johnson J.M., Sekhon R.S., Kaeppler S.M., Leon N.D. (2011) Genetic diversity of a maize association population with restricted phenology. *Crop Sci.* 51 (2): 704-715.
- Hotelling H. (1936) Relations between two sets of variables. *Biometrika* 28: 321-377.
- Kang H.M., Zaitlen N.A., Wade C.M., Kirby A., Heckerman D., Daly M.J., Eskin E. (2008) Efficient control of population structure in model organism association mapping. *Genetics* 178 (3): 1709-1723.
- Kraakman A.T.W., Martínez F., Mussiraliev B., Eeuwijk F.A., Niks R.E. (2006) Linkage disequilibrium mapping of morphological, resistance, and other agronomically relevant traits in modern spring barley cultivars. *Molecular Breeding* 17 (1): 41-58.
- Malosetti M., Linden C., Vosman B., van Eeuwijk F. (2007)

- A mixed-model approach to association mapping using pedigree information with an illustration of resistance to *Phytophthora infestans* in potato. *Genetics* 175: 879-889.
- Muñoz-Amatriaín M., Cuesta-Marcos A., Endelman J.B., Comadran J., Bonman J.M., Bockelman H.E., Chao S., Russell J., Waugh R., Hayes P., Muehlbauer G. (2014) The USDA Barley Core Collection: Genetic diversity, population structure, and potential for genome-wide association studies. *PLoS ONE* 9 (4): e94688.
- Peña Malavera A. (2015) Aproximaciones estadísticas para el mapeo asociativo en estudios genéticos. Universidad Nacional de Córdoba, Córdoba, 117 pp.
- Peña Malavera A., Gutierrez L., Balzarini M. (2014) Componentes principales en mapeo asociativo. *JBAG* 25: 32-40.
- Price A.L., Patterson N.J., Plenge R.M., Weinblatt M.E., Shadick N.A., Reich D. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* 38 (8): 5.
- Pritchard J., Stephens M., Donnelly P. (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.
- Remington D., Thornsberry J., Matsuoka Y., Wilson L., Whitt S., Doebley J., Kresovich S., Goodman M.M., Buckler E.S. (2001) Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proc. Natl. Acad. Sci. USA* 98 (20): 11479-11484.
- Sargolzaei M., Schenkel F. (2009) QMSim: a large-scale genome simulator for livestock. *Bioinformatics* 25: 680-681.
- Stich B., Melchinger A. (2009) Comparison of mixed-model approaches for association mapping in rapeseed, potato, sugar beet, maize, and *Arabidopsis*. *BMC Genomics* 10 (1): 1-14.
- Stich B., Melchinger A., Heckenberger M., Möhring J., Schechert A., Piepho H.P. (2008) Association mapping in multiple segregating populations of sugar beet (*Beta vulgaris* L.). *Theor. Appl. Genet.* 117 (7): 1167-1179.
- Thornsberry J., Goodman M., Doebley J., Kresovich S., Nielsen D., Buckler E. (2001) Dwarf8 polymorphisms associate with variation in flowering time. *Nat. Genet.* 28 (3): 286-289.
- Tracy C.A., Widom H. (1994) Level-spacing distributions and the Airy kernel. *Comm. Math. Phys.* 159 (1): 151-174.
- Wang M., Jiang N., Jia T., Leach L., Cockram J., Waugh R., Ramsay L., Thomas B., Luo Z. (2012) Genome-wide association mapping of agronomic and morphologic traits in highly structured populations of barley cultivars. *Theor. Appl. Genet.* 124 (2): 233-246.
- Wright S. (1951) The genetical structure of populations. *Ann. Eugen.* 15: 31.
- Yu J., Pressoir G., Briggs W., Bi I., Yamasaki M., Doebley J., McMullen M.D., Gaut B.S., Nielsen D.N., Holland J.B., Kresovich S., Buckler E.S. (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* 2: 203-208.
- Zhang Z., Ersoz E., Lai C.Q., Todhunter R.J., Tiwari H.K., Gore M.A., Bradbury P., Yu J., Arnett D., Ordovas J., Buckler E. (2010) Mixed linear model approach adapted for genome-wide association studies. *Nat. Genet.* 42 (4): 355-360.
- Zhu C., Gore M., Buckler E., Yu J. (2008) Status and prospects of association mapping in plants. *Plant Genome* 1 (1): 16.

AGRADECIMIENTOS

El presente trabajo es parte de la tesis de Andrea Peña Malavera para el cumplimiento de los requisitos del Doctorado en Ciencias de la Ingeniería de la Universidad Nacional de Córdoba y del programa de becas de posgrado del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET).

OBITUARY

TIRANTI, Iván Nicolás

Junín 10/8/1927 – Río Cuarto 26/8/2016



A los 89 años falleció en la ciudad de Río Cuarto quien fuera orientador del grupo de genética y mejoramiento vegetal y animal de la Facultad de Agronomía y Veterinaria (FAV) de la Universidad Nacional de Río Cuarto (UNRC). A partir de junio de 1975 y hasta diciembre de 1998 fue Profesor Titular con Dedicación Exclusiva y desde 1999 fue Profesor Consulto vitalicio. En la Sociedad Argentina de Genética fue Presidente (1989-1991), Miembro del Consejo Asesor (1971-1974, 1981-1989, 1993-2002) y Socio Honorario desde septiembre de 2001.

Se graduó como Ingeniero Agrónomo en la Universidad Nacional de La Plata (1953), obtuvo sus grados académicos de *MSci* en Texas A&M University (1963), con una tesis sobre citogenética de híbridos interespecíficos del género *Gossypium*, y de *PhD* en la California University- Davis (1970) con una tesis sobre diploides de alfalfa derivados de cruzamientos entre tetraploides y diploides.

Trabajó en la Estación Experimental Roque Sáenz Peña del Instituto Nacional de Tecnología Agropecuaria (1957-59 y 1960-67); en la Chacra Experimental Miramar del Ministerio de Asuntos Agrarios de la Provincia de Buenos Aires (1959-60); en la Universidad Nacional del Sur, Departamento Agronomía (1970-75), donde compartió la docencia y estableció una entrañable amistad con el recordado Ovidio Núñez.

Ingresó a la UNRC como profesor de Genética y Fitotecnia de Agronomía, luego en Mejoramiento Vegetal y más tarde en Genética General de Medicina Veterinaria. En sus inicios en la UNRC y durante varios años, colaboró junto a otros docentes de Agronomía con la asignatura Genética de las carreras de Biología y Microbiología, Facultad de Ciencias Exactas Físico-Químicas y Naturales. Además, dirigió varios estudiantes en su tesis de doctorado, maestría y fue supervisor de becarios y ayudantes de investigación de la UNRC.

Sus trabajos de investigación fueron en centeno, pastos nativos y alfalfa mientras estuvo en Agronomía, posteriormente en rearrreglos cromosómicos en suinos y finalmente en genética apícola, con publicaciones nacionales e internacionales y numerosas comunicaciones en congresos. Fue colaborador permanente de la revista técnico-científica de la UNRC y participó activamente de la administración universitaria como Consejero Docente en la FAV; asimismo fue miembro del Consejo de Investigación de la UNRC, Director de la Escuela de Posgraduación, entre otros cargos de gestión universitaria. Fue evaluador de proyectos internos y externos, de becarios y revisor de artículos para revistas de su especialidad.

Todos los que tuvimos relación laboral con Iván lo recordaremos siempre. Fue un orientador permanente en los quehaceres académicos de docencia, experimentación a campo e investigación en laboratorio, en la gestión universitaria y la formación de personal para las actividades universitarias. Lector incansable, con un anecdotario propio de quien vivió su pasar por la vida con la satisfacción de bridar toda su formación en beneficio de sus compañeros de trabajo, de la Universidad, de su familia y de la sociedad donde estaba inserto. Fuente de toda consulta técnica o cultural, amante del cine, la ópera, la natación y el buen café que molía en molinillo propio esparciendo su exquisito aroma alrededor de la oficina que ocupaba. Siempre agradeceremos su permanente predisposición para enseñar, instruir y educar.

Ing. Agr. (MSc.) Víctor A. Ferreira



BAG
Journal of Basic & Applied Genetics