



OPTIMIZACIÓN DE MARCADORES INFORMATIVOS DE ANCESTRÍA (AIMS) PARA CARACTERIZAR LA COMPOSICIÓN DE LA POBLACIÓN ARGENTINA



OPTIMIZATION OF ANCESTRY INFORMATIVE MARKERS (AIMS) TO CHARACTERIZE THE COMPOSITION OF THE ARGENTINE POPULATION

Patíño J.¹, Lozano M.^{1,2}, Russo M.G.^{3,4}, Dejean C.B.^{1,2}, Avena S.^{1,4}

¹ Centro de Ciencias Naturales, Ambientales y Antropológicas (CCNAA), Universidad Maimónides, Buenos Aires, Argentina.

² Sección de Antropología Biológica, Facultad de Filosofía y Letras, Universidad de Buenos Aires (UBA), Buenos Aires, Argentina.

³ Instituto de Arqueología, Facultad de Filosofía y Letras, UBA, Buenos Aires, Argentina.

⁴ Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, Argentina.

Corresponding author:
Magdalena Lozano
lozanomagdalena@gmail.com

ORCID 0009-0007-8702-6826

—

Cite this article as:

Patíño J., Lozano M., Russo M.G., Dejean C.B., Avena S. 2025. OPTIMIZATION OF ANCESTRY INFORMATIVE MARKERS (AIMS) TO CHARACTERIZE THE COMPOSITION OF THE ARGENTINE POPULATION. BAG: Journal of Basic and Applied Genetics XXXVI (2).

—

Received: 07/16/2025

Revised version received: 09/30/2025

Accepted: 10/18/2025

General Editor: Elsa Camadro

DOI: 10.35407/bag.2025.36.02.01

ISSN online version: 1852-6233

RESUMEN

La población argentina presenta una compleja estructura genética resultante de procesos históricos de mestizaje entre componentes euroasiáticos occidentales, indígenas y africanos subsaharianos, con marcadas diferencias regionales a lo largo del país. Con el objetivo de caracterizar la composición de la población del Área Metropolitana de Buenos Aires (AMBA) se optimizó un panel de 30 marcadores informativos de ancestría (AIMs), combinando técnicas de PCR-APLP múltiplex, PCR-RFLP e inserción Alu. Se analizaron 50 muestras de donantes de sangre no emparentados, provenientes de hospitales públicos y privados, utilizando los programas STRUCTURE y ADMIX para estimar contribuciones ancestrales individuales y poblacionales. Encontramos predominio del componente euroasiático occidental (84,0%), seguido del indígena (12,5%) y subsahariano (3,4%). Al optimizar y estandarizar la técnica se excluyeron cinco AIMs en desvío del equilibrio de Hardy-Weinberg; a través de distintas pruebas pudimos identificar que esto no alteraba significativamente las estimaciones, validando la robustez del panel seleccionado. Este trabajo demuestra que los 30 AIMs seleccionados permiten estimaciones confiables de ancestría en poblaciones mestizas, proponiendo una herramienta molecular de bajo costo y de alta eficiencia para determinar la ancestralidad genética individual y poblacional, superando limitaciones de métodos previos basados en distinta cantidad y calidad de marcadores. Es de destacar que estas aplicaciones trascienden la antropología biológica, ya que la identificación de la ancestría local es relevante, también, para estudios biomédicos y farmacogenéticos en Argentina, donde suele asumirse erróneamente una homogeneidad europea.

Palabras clave: diversidad genética, marcadores autosómicos, mestizaje.

ABSTRACT

The Argentine population exhibits a complex genetic structure resulting from historical admixture processes among Western Eurasian, Indigenous, and Sub-Saharan African components, with marked regional variations across the country. To characterize the population composition of the Buenos Aires Metropolitan Area (AMBA), we optimized a panel of 30 ancestry informative markers (AIMs), combining multiplex APLP-PCR, RFLP-PCR, and Alu insertion techniques. We analyzed 50 unrelated blood donor samples from public and private hospitals, using STRUCTURE and ADMIX programs to estimate ancestral contributions at the individual and population levels. Results showed predominance of the Western Eurasian component (84.0%), followed by Indigenous (12.5%) and Sub-Saharan African (3.4%) ancestries. Through technique optimization and standardization, we excluded five AIMs in Hardy-Weinberg disequilibrium; various tests confirmed that this exclusion did not significantly alter the estimates, validating the panel's robustness. This work demonstrates that the selected 30 AIMs provide reliable ancestry estimates in admixed populations, offering a cost-effective and highly efficient molecular tool for determining individual and population genetic ancestry that overcomes the limitations of previous methods based on varying marker quantity and quality. Notably, these applications extend beyond biological anthropology, as local ancestry identification is also relevant for biomedical and pharmacogenetic studies in Argentina, where European homogeneity is often erroneously assumed.

Key words: genetic diversity, autosomal markers, admixture.

INTRODUCCIÓN

Las poblaciones latinoamericanas se caracterizan por presentar proporciones variables de ascendencia euroasiática occidental, indígena y africana (subsahariana), con marcadas diferencias regionales. El componente autóctono es más numeroso en las regiones donde los pueblos originarios practicaban la agricultura intensiva, con una gran densidad de población, como en los Andes Centrales o México, mientras que el componente subsahariano lo es donde el tráfico de esclavizados ha sido muy intenso durante la etapa colonial para el desarrollo de la economía de plantación, particularmente en Brasil y el Caribe (Avena *et al.*, 2013).

A la llegada de los conquistadores españoles, en lo que actualmente es Argentina había una mayor cantidad de población autóctona en el norte, especialmente en el noroeste. El proceso de mestizaje se inició ya con la conquista. Desde ese momento, el componente mestizo, el criollo, el indígena y, aunque en menor medida, el mulato, el pardo y el africano constituyeron el núcleo del pueblo argentino en formación, al que luego se sumarían las masivas migraciones europeas desde finales del siglo XIX (Martínez Sarasola, 2011).

Luego de la independencia de los países iberoamericanos en el siglo XIX, se incrementó la inmigración desde Europa. Este proceso fue particularmente importante en Argentina, donde arribaron 6.400.000 inmigrantes, siendo el periodo de afluencia más intensa el comprendido entre 1881 y 1914 con la llegada de 4.200.000 personas (Texidó, 2008). La mayor cantidad de extranjeros se radicó en el centro del país; en 1914 constituían el 49% de la población de la Capital Federal, el 35% en Santa Fe, el 34% en la Provincia de Buenos Aires, el 32% en Mendoza y el 20% en Córdoba (Sáez Capel, 2001). A partir de la década de 1930, la caída de los precios agrícolas tras la fuerte crisis económica que se había originado en EE.UU. impulsó el desplazamiento de familias desde el campo hacia las ciudades, particularmente Buenos Aires y Rosario, en busca de nuevas oportunidades laborales. Desde 1940, la industrialización incentivó estas migraciones internas y de países limítrofes hacia grandes centros urbanos, especialmente al Área Metropolitana de Buenos Aires (Belini y Korol, 2019).

Dado estos distintos procesos poblacionales, la estimación antropogenética de la composición de la población argentina se constituye en un tema de relevancia. Además de su impacto en el área de la identidad, el conocimiento del acervo genético de la población resulta de aplicabilidad en la práctica médica. En diversas estimaciones sobre la prevalencia de patologías con bases genéticas, a falta de datos, suele asumirse que en nuestro país se registrarían los mismos valores que en Europa (Avena *et al.*, 2024). Las estimaciones de ancestría de una población permiten

considerar factores de riesgo para determinados individuos en relación a patologías con bases genéticas, las cuales pueden variar su prevalencia en los distintos grupos humanos (Peterson *et al.*, 2019). Desde los trabajos señeros de Lewontin (1972) es conocido que, si bien la mayor diversidad genética humana se debe a diferencias entre individuos, hay una fracción menor pero significativa que se explica por variación interpoblacional (Carnese, 2019) que, en ciertos casos, puede relacionarse con genes asociados a determinadas enfermedades. Esto explicaría el interés que existe, especialmente en los últimos años, por analizar las distintas prevalencias de estas patologías y su relación con la ancestría. Por ejemplo, Park *et al.* (2017) analizaron las frecuencias de varios tipos de cáncer en distintas poblaciones a nivel mundial, encontrando mayores valores de tumores de próstata en africanos, de mamas en personas de ancestría europea, y colorrectal, esófago y pulmón en asiáticos del este. Reitz *et al.* (2023) han relevado variantes genéticas relacionadas al Alzheimer más frecuentes en personas de origen europeo y asiático oriental. Hoffman *et al.* (2019) señalan que en Estados Unidos una mayor incidencia de cáncer de mama se asocia a la ancestría europea, y entre las mujeres de origen latinoamericano, aquellas que tenían mayor porcentaje de ancestría indígena mostraban un menor riesgo de desarrollar este cáncer.

En farmacogenética también se está prestando una creciente atención a la variabilidad en la respuesta a las drogas según el origen geográfico de las personas, tanto por su eficacia como en su toxicidad. Por ejemplo, la dosis apropiada en promedio de tacrolimus luego de trasplante hepático es un 40% menor en “hispanos de Estados Unidos” respecto a la población “blanca”. Además, los “hispanos” tienen 1,3 veces más probabilidades de presentar insuficiencia renal que los estadounidenses “blancos” (Chamzas *et al.*, 2024). Otro ejemplo se ha reportado en el tratamiento de pterigión con doxiciclina, que ha sido eficaz para las personas de origen europeo, pero no para las “hispanas” (Rúa *et al.*, 2012). Se ha descrito un mayor efecto tóxico de la citarabina utilizada contra la leucemia mieloide aguda en afroamericanos (Wheeler *et al.*, 2011; Di Francia *et al.*, 2021) y del uso de vincristina en niños “hispanos” con leucemia linfoblástica (Mc Clain *et al.*, 2017), entre otros muchos casos.

El análisis de los procesos histórico-demográficos de nuestra población guarda relación con los resultados obtenidos en los estudios de marcadores informativos de ancestralidad (AIMs, por la sigla en inglés de *Ancestry Informative Markers*), los cuales han mostrado que la mayoría de las personas tiene en su ancestría individual al menos dos y hasta tres componentes a nivel continental: euroasiático, autóctono y/o subsahariano (Corach *et al.*, 2010; Avena *et al.*, 2012; Muzzio *et al.*, 2018; Parolin *et al.*, 2019). Por este motivo, es altamente

probable que la asociación entre el aspecto externo y la ancestría no resulte la misma que en las poblaciones europeas. Este tema ha sido muy abordado en Brasil; Suárez-Kurtz (2008), por ejemplo, realizó un estudio en el que las personas participantes se autclasificaron como “blancas”, “mulatas” y “negras”. No obstante, al realizar la determinación de AIMs, se observó que las proporciones de ancestría europea fueron del 86%, 68% y 43%, respectivamente. Debe destacarse el hecho de que algunos individuos “blancos” tenían más ancestría africana que ciertas personas “negras”, y viceversa, concluyéndose que, en el contexto de ese país, el color de la piel llamativamente era un débil predictor de la ancestralidad. Esto fue reforzado por Pena *et al.* (2020), quienes señalan que en ese país no se emplea el término “raza” sino el de “color”, el cual se basa en una evaluación fenotípica. Los autores plantean que, a diferencia de Estados Unidos, la razón por la que se prefiere la palabra “color” a “raza” probablemente se deba a que captura los aspectos continuos de los fenotipos de la población brasileña. Este sistema de identificación basado principalmente en el aspecto implica clasificar a los individuos según la presencia de ciertos alelos de un pequeño número de genes que influyen en la apariencia física, ignorando el resto del genoma. Los estudios demográficos han demostrado que allí ocurre el apareamiento selectivo basado en el color. La asociación inicial entre el color y los AIMs inevitablemente decaerá con las generaciones debido a la mezcla genética, produciendo una población con grupos de color distintos y, sin embargo, con niveles no tan diferentes en sus ascendencias. Pena *et al.* (2020) concluyen que la correlación entre el color de la piel y la ascendenciagenómica en Brasil es muy imperfecta: a nivel individual, no se puede predecir la ascendencia europea, africana o amerindia por fenotipo y enfatizan que sólo se puede conocer la ancestría genética determinando AIMs. Si bien este caso no es directamente extrapolable a la Argentina, sí nos advierte que no es correcto adoptar acríticamente los esquemas de clasificación utilizados en Estados Unidos en Latinoamérica, pues en ésta la ancestría de una persona no debería presumirse por su fenotipo o aspecto físico, por el contrario debería ser estimada mediante la determinación de AIMs.

La relación entre la autopercepción, el fenotipo y la ancestría genética adquiere especial relevancia en la práctica médica. Por ejemplo, en un estudio realizado con 544 médicas/os de nuestro país (Ruderman *et al.*, 2020), el 64,5% indagaba sobre el origen poblacional o étnico de sus pacientes. De este grupo, el 70,3% preguntaba al paciente por su origen geográfico/étnico/nacionalidad auto-reportada, o bien por el de sus padres y/o abuelos. El 6,2% se basaba en los rasgos físicos externos (por ejemplo, su color de piel), en el apellido y/o en el acento del paciente para asignarle un origen geográfico/étnico/nacional.

Estudios con AIMs en la población argentina

El análisis de la estructura genética mediante AIMs nos permite estimar las proporciones de cada componente característico de la población argentina (euroasiático occidental, indígena y africano subsahariano). Desde hace unos años se ha incrementado el número de trabajos que emplean AIMs con alto poder de discriminación, los cuales han permitido observar las diferencias con respecto a la contribución ancestral en poblaciones mestizadas (Salzano y Sans, 2014). A título de ejemplo, podemos citar a Galanter *et al.* (2012) quienes desarrollaron un panel de 446 AIMs optimizados para estudiar la ancestría en poblaciones de Latinoamérica.

En el caso específico de Argentina, en la ciudad de Buenos Aires, Avena *et al.* (2001) estimaron un 13% de ancestría autóctona y un 2,2% de aporte africano, al tipificar nueve marcadores (ABO, Rh, MNSS, P, Diego, Duffy, Kell, Kidd, Lutheran, Alotipos: Gm y Km). Martínez-Marignac *et al.* (2004), utilizaron un set de cinco marcadores para estudiar la ancestría individual en La Plata. Estos autores reportaron que el porcentaje de ancestría europea oscilaba entre el 48% y el 75%, el de ancestría autóctona entre el 20% y el 45%, y que todas las personas presentaban más de un 4% de aporte africano. Fejerman *et al.* (2005), utilizando 12 AIMs específicos para africanos en una muestra de ciudad de Buenos Aires, encontraron que el aporte promedio era de 2,2% pero pudo observarse que estaba concentrado en el 10% de la muestra. Seldin *et al.* (2007), a partir del estudio de 78 SNPs, calcularon en 19% el componente indígena y en 2% el subsahariano; los autores afirmaban analizar todo el país, pero en realidad se trataba de cinco ciudades de la pampa húmeda. Además, el reducido tamaño de la muestra con el que trabajaron (94 individuos) no resulta adecuado para dar cuenta de las diferentes regiones de un país como Argentina, con una gran extensión y diversidad. Corach *et al.* (2010) utilizaron un enfoque regional determinando 24 AIMs, y encontraron diferencias significativas en la ancestría indígena: un 28% en la región sur, un 15% en la región central y un 17% en la región noroeste del país. Di Fabio Rocca *et al.* (2016), a partir de 11 AIMs, estimaron la mezcla genética en la ciudad de Salta, determinando un aporte mayoritario de ancestría indígena (50,2%), seguido por el aporte europeo (46,3%) y subsahariano (3,5%). Para la región central de la Argentina, empleando 10 AIMs, se reportaron valores de contribución euroasiática de 55,1% en Córdoba y de 58,2% en San Luis (García *et al.*, 2015); para la región patagónica, con 46 AIMs, se reportaron valores de contribución euroasiática de 62,1% (Parolin *et al.*, 2019). En este último trabajo se aprecia que puede haber importantes diferencias al interior de una misma provincia (Chubut), con un rango que va del 45% en Comodoro Rivadavia (Avena *et al.*, 2009) y Esquel (Avena *et al.*, 2010) al 75% en Puerto Madryn.

Debe señalarse que los estudios citados emplean distintos números de marcadores, los cuales además tienen distinto poder informativo. Para especificar, y a título de ejemplo, deberían evitarse los marcadores como el rs2274533, utilizado por Seldin *et al.* (2007), que presenta una frecuencia de 0,85 en europeos, de 0,07 en indígenas y de 0,49 en africanos del oeste, ya que una frecuencia cercana a esta última es esperable tanto en una población con un aporte predominantemente subsahariano, como en una población en la que este componente esté ausente pero haya valores similares de aporte europeo e indígena. Casos como éste no sólo no realizan ningún aporte, sino que introducen una distorsión en la estimación de ancestría.

Considerando que, por un lado, el número de AIMS utilizados puede afectar las estimaciones individuales de ancestría, y por el otro, su cantidad puede estar constreñida por factores económicos, resultaba necesario calcular el número mínimo de AIMS a partir del cual las estimaciones tienen una confiabilidad aceptable. A través de datos publicados sobre 99 AIMS (Avena *et al.*, 2012), nuestro equipo tomó conjuntos de 10, 20, 30 y 50 marcadores seleccionados al azar en una muestra de 441 individuos. El análisis de la variabilidad de las estimaciones de ancestría permitió concluir que resultaba necesario contar con al menos un número mínimo de 30 marcadores biparentales (Russo *et al.*, 2016). En concordancia con este resultado, Ruiz Narváez *et al.* (2011), empleando 30, 60, 100, 150 y 200 AIMS de un total de 1757, concluyeron que 30 marcadores evidenciaban una alta correlación de estimación de ancestría (0,89). Adicionalmente, podemos señalar el trabajo de Ruiz Linares *et al.* (2014), donde se evaluó la ancestralidad en poblaciones de cinco países latinoamericanos: México, Colombia, Perú, Chile y Brasil. Estos autores también sostienen que utilizando 30 AIMS se obtiene una correlación de 0,7 para las estimaciones individuales, cotejadas con un panel de 50.000 marcadores, y valores casi idénticos para las estimaciones a nivel poblacional.

A partir de estas consideraciones nos propusimos poner a punto en nuestro laboratorio una herramienta molecular de bajo costo y de alta eficiencia para determinar la ancestralidad genética individual y poblacional, cuyo objetivo es la aplicación y uso en futuras investigaciones tanto en el campo de la antropología biológica como en el de la biomedicina.

MATERIALES Y MÉTODOS

Se analizaron 50 muestras recolectadas al azar de dadores de sangre no relacionados biológicamente, provenientes del Hospital Italiano (privado, N=25) y del Hospital de Clínicas (público, N=25), ambos del Área Metropolitana de la Ciudad de Buenos Aires (AMBA).

Las muestras forman parte del banco de ADN del Equipo de Antropología Biológica (UMai-UBA). Todos los participantes accedieron a donar voluntariamente sangre periférica a través de un consentimiento informado para realizar estudios genéticos aprobado por el Comité de Ética del Hospital Italiano de Buenos Aires (protocolo n°1356). Las muestras fueron anonimizadas.

Se aplicó la técnica de polimorfismos de productos de longitud amplificados en múltiplex (APLP) para SNPs (por las siglas en inglés de *Single Nucleotide Polymorphisms*; mutación puntual en un cambio de base nucleotídica) mitocondriales, propuesta por Umetsu *et al.* (2005), y modificada por Jurado Medina *et al.* (2014). Se desarrolló un protocolo de amplificación alelo-específico para la tipificación de SNPs biparentales de bajo costo, consistente en un número reducido de marcadores, pero con un alto poder de discriminación entre las tres parentales (euroasiática occidental, indígena y africana subsahariana).

Diseño de cebadores alelo-específicos

Se construyeron tres cebadores para cada SNP. Un cebador se ensambla específicamente al nucleótido de la secuencia del alelo de referencia y otro al alelo con la variación; ambos difieren únicamente en su longitud debido a la adición de una cola de nucleótidos no complementarios. El tercer cebador delimita el tamaño del fragmento, permitiendo obtener amplicones de tamaños no superpuestos para las dos variantes alélicas.

Para la búsqueda de la secuencia de ADN del SNP de interés se utilizó la base de datos dbSNP del NCBI (*National Center for Biotechnology Information, U.S. National Library of Medicine*) (<https://www.ncbi.nlm.nih.gov/snp/>). En la puesta a punto del diseño experimental se adicionaron en los casos necesarios las colas de cinco o seis bases en las secuencias no complementarias. Por último se controlaron los cebadores utilizando el programa BLAST que pertenece al NCBI – Pub Med (disponible en http://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome), y una vez validados como óptimos se utilizaron para la genotipificación de la población de estudio.

Relevancia y descripción de los marcadores seleccionados

Al inicio del trabajo experimental se consideraron 40 AIMS para realizar diferentes combinaciones en múltiplex hasta encontrar la más adecuada. Se realizó una selección detallada tomando marcadores de estudios previos en donde se genotipificaron poblaciones latinoamericanas (Shriver *et al.*, 2003; Bonilla *et al.*, 2004; Fejerman *et al.*, 2005; Seldin *et al.*, 2007; Duque *et al.*, 2012; Doldan, 2014; Patiño, 2020). Se descartaron los que presentaron en la fase experimental bandas inespecíficas, no amplificaron alguna de las dos

variantes alélicas y/o no se visualizaron adecuadamente en el gel de electroforesis. Inicialmente, se agruparon los cebadores de cada AIM según las temperaturas de *melting* (T_m) y según los fragmentos amplificados para evitar su solapamiento en la lectura electroforética. Fueron seleccionados los 30 AIMS más apropiados, que se describen sucintamente en la Tabla 1, considerando que esta es la cantidad mínima aceptable de marcadores para la determinación de ancestría individual (Ruiz Narváez *et al.*, 2011; Ruiz Linares *et al.*, 2014; Doldan, 2014; Russo *et al.*, 2016). Veintinueve fueron determinados a través del análisis de SNPs, correspondiendo el restante a una mutación que implica la inserción o delección de una secuencia Alu de 300 pb (Gómez Perez *et al.*, 2007). Para evitar el enmascaramiento, los marcadores elegidos presentan una diferencia en sus frecuencias mayor al 40% entre una población parental respecto a las otras, siendo a su vez la diferencia entre las dos restantes menor al 5% (Galanter *et al.*, 2012; Patiño, 2020).

En la Tabla 1 se presentan las frecuencias alélicas de los marcadores seleccionados en las parentales europea, indígena y afro subsahariana, según las publicaciones de referencia y las bases de datos genómicas (HGDP-CEPH: Human Genome Diversity Project, ALFRED y 1000 Genomes). No fue posible disponer de genotipos de referencia para la parental indígena en 18 AIMS, por lo cual se realizaron simulaciones de los genotipos teniendo en cuenta las frecuencias alélicas poblacionales reportadas en la bibliografía (Tabla 1). Esto se debe a que la mayoría de las bases genómicas como NCBI, HGDP-CEPH y 1000 Genomes, contienen datos genéticos de poblaciones consignadas como indígenas (Colombia, México, Puerto Rico) pero que han sido recopilados a partir de poblaciones cosmopolitas, como es el caso de Medellín en Colombia. Estas muestras de poblaciones cuyos individuos provienen de ciudades que han atravesado varios procesos migratorios no reflejan necesariamente el genotipo indígena de la región.

Amplificación de ADN

La determinación por PCR-APLP se realizó utilizando en la mezcla de reacción 4 μ l de buffer de reacción Green Promega (5X), 0,8 μ l de Cl_2Mg 25 mM, 0,45 μ l de la mezcla de cebadores de cada múltiplex, 0,08 μ l de una mezcla de los cuatro desoxirribonucleótidos trifosfato 10 mM, 0,06 μ l de GoTaq Polimerasa Promega y 14 μ l de agua calidad biología molecular para completar 20 μ l de volumen final de reacción. A cada tubo se le agregaron 0,5 μ l de ADN genómico (50ng/ μ l) y para cada set de cebadores se estableció una concentración óptima.

La determinación por PCR-RFLP se realizó utilizando en la mezcla de reacción 3 μ l de buffer de reacción Green Promega (5X), 1 μ l de Cl_2Mg 25 mM, 0,6 μ l de la mezcla de cebadores correspondientes, 0,07 μ l de una mezcla de los cuatro desoxirribonucleótidos trifosfato 10 mM, 0,07 μ l

de GoTaq Polimerasa Promega y 8,26 μ l de agua calidad biología molecular para completar 15 μ l de volumen final de reacción. A cada tubo se le agregaron 2 μ l de ADN genómico (50ng/ μ l) y para cada set de cebadores se estableció una concentración óptima.

Las distintas PCR cumplieron la configuración de etapas detallada en la Tabla 2. La concentración de cebadores puede verse detallada en la Tabla 3 para las PCR-APLP y múltiplex, mientras que en la Tabla 4 se detallan las condiciones de las PCR-RFLP.

Estimaciones de ancestría

Como primer paso se calcularon las frecuencias alélicas mediante el complemento de Microsoft Excel GenAIEx 6.5 (Peakall y Smouse, 2012), con el cual también se evaluó si éstos se encontraban en equilibrio de Hardy–Weinberg.

La estimación de la ancestría a nivel individual se realizó mediante el programa STRUCTURE 2.3.4. (Pritchard *et al.*, 2000). Luego de realizar cinco corridas para cada K (de 1 a 5) se usó el programa CLUMPP (*Cluster Matching and Permutation Program*) para reorganizar las asignaciones de pertenencia a grupos (*clusters*) provenientes de las múltiples ejecuciones. Con el programa STRUCTURE HARVESTER, que emplea el método de Evanno *et al.* (2005), se obtuvo un valor del estadístico delta K de 1,67 para K=3, resultando éste el número óptimo de *clusters*. El periodo de *burn-in* fue de 100.000 y las probabilidades posteriores se estimaron a partir de 1.000.000 iteraciones bajo el modelo *admixture*, utilizando un alfa de 0,33 (1/k, siguiendo a Wang *et al.*, 2008).

A nivel poblacional se realizaron dos aproximaciones distintas. Por un lado, se utilizó el método de identidad génica mediante el programa ADMIX 95, basado en el método de los cuadrados mínimos desarrollado por Chakraborty (1985) (disponible en: <http://www.genetica.fmed.edu.uy/software.htm>), que constituye una forma directa de estimación al utilizar las frecuencias alélicas de los marcadores tanto en las poblaciones parentales como en la muestra de estudio. Por otra parte, se obtuvo un promedio de las estimaciones de ancestría individual calculadas con el ya citado programa STRUCTURE, que es un método indirecto de estimación de mezcla génica poblacional.

Las comparaciones de diferentes ancestrías poblacionales generadas a partir de distintos conjuntos de datos se validaron mediante el análisis estadístico de comparación de proporciones, mientras que las ancestrías individuales se compararon mediante el cálculo del coeficiente de correlación de Pearson. En este último caso, debido a las múltiples comparaciones, para evaluar la significación de la correlación se empleó la corrección de Bonferroni (siguiendo a Russo *et al.*, 2016). Todas las pruebas estadísticas se realizaron con el programa InfoStat (Di Rienzo *et al.*, 2003).

Tabla 1. Marcadores (30) informativos de ancestría con código de referencia del polimorfismo de nucleótido único (rs dbSNP, por su sigla en inglés) en la base de datos NCBI (National Center for Biotechnology Information, U.S. National Library of Medicine) y frecuencias alélicas según publicaciones de referencia y bases de datos genómicas: HGDP-CEPH (Human Genome Diversity Project), ALFRED (the Allele Frequency Database) y 1000 Genomes, en población europea (EURO), indígena y afro subsahariana (SUBSA).

dbSNP rs	AIMs	Antecedentes	Alelo Ancestral	Frecuencia según REFERENCIA		
				EURO	INDÍGENA	SUBSA
734780	EURO		C	0.080	0.730	0.740
1426654	SLC24A5		A	1.000	0.050	0.020
2278354	ROPN1L	Se encuentran asociados con la pigmentación de la piel, evidenciando el aumento o disminución de la expresión proteica (Lee et al., 2010; Leite, 2011)	T	0.900	0.240	0.280
17203	WI-11153		G	0.133	0.819	0.785
2695	WI-11909		G	0.881	0.181	0.805
1042602	TYR192	Han sido utilizados en fenotipos relacionados con el envejecimiento en adultos (Reiner et al., 2005)	A	0.449	0.034	0.050
285	LPL		T	0.492	0.442	0.971
1800498	DRD2	Han sido utilizados en estudios de mezcla poblacional (Kang et al., 2011)	T	0.670	0.045	0.135
2816	WI-7423	Han sido utilizados como indicadores en estudios de asociación genética entre dos tipos de diabetes y enfermedad renal en etapa terminal (Keene et al., 2008)	T	0.476	0.058	0.000
2288697	KLHL29	Han sido utilizados en la incidencia de hipertensión en mujeres jóvenes relacionada a la herencia genética y al índice de masa corporal (Taylor et al., 2010)	A	0.030	0.850	0.010
4646	CYP19-E2	Han sido utilizados en la incidencia de cáncer de mama (Cocca et al., 2016; Beltran et al., 2016)	A	0.287	0.741	0.332
3138522	APO	Se relaciona con el metabolismo de lípidos, promoviendo el flujo de colesterol de los tejidos hacia el hígado para la excreción.	inserción	0.925	0.977	0.420
7041	GC*S	Han sido utilizados indicadores de polimorfismos de receptores de la vitamina D en la población asiática oriental (Wang et al., 2018), europea (Miller et al., 2018) e iraní (Nazemismalman et al., 2019)	G	0.607	0.542	0.069
2814778	FY*NULL	Es considerando como una ventaja adaptativa y resistencia al parásito <i>Plasmodium vivax</i> , causante de una forma grave de malaria (Mansilla, 2010)	FY NULL (C)	0.002	0.000	0.999
2065982	AMER1		C	0.060	0.810	0.090
13069719	FOXP1		C	0.850	0.090	0.840
12469563	AMER4		C	0.030	0.760	0.000
16897585	AMER2		A	0.000	0.780	0.000
2384319	KIF3C		G	0.070	0.830	0.060
2763	WI-9231		G	0.147	0.548	0.129
10510511			A	0.020	0.920	0.000
3870336	DAG1	No se ha descrito su función biológica	A	0.090	0.940	0.090
2042314			A	0.150	1.000	0.140
7211306	OGFOD3		A	0.280	0.900	0.290
9290363	MECOM		G	0.090	0.100	0.880
8009244	DLGAP5		A	0.890	0.910	0.140
5025718	NOTCH2		C	0.870	0.880	0.030
596985	ROR1		T	0.970	0.990	0.130
1871534	SLC39A4		C	0.010	0.030	0.950
984303	PTPRO		A	1.000	0.980	0.210

REFERENCIA	Frecuencia según HGDP			Frecuencia según ALFRED			Frecuencia según 1000 genomes		
	EURO	INDÍGENA	SUBSA	EURO	INDÍGENA	SUBSA	EURO	INDÍGENA	SUBSA
<i>Duque et al, 2012</i>							0.111	0.432	0.726
<i>Seldin et al, 2007</i>					0.062		0.996	0.589	0.074
<i>Duque et al, 2012</i>							0.900	0.620	0.340
<i>Shriver et al, 2003</i>							0.138	0.520	0.800
<i>Shriver et al, 2003</i>	0.780	0.160	0.780				0.730	0.490	0.750
<i>Shriver et al, 2003</i>	0.370	0.020	0.010		0.027		0.372	0.238	0.012
<i>Shriver et al, 2003</i>					0.277		0.517	0.461	0.933
<i>Shriver et al, 2003</i>				0.614	0.031	0.146	0.613	0.378	0.148
<i>Shriver et al, 2003</i>	0.510	0.050	0.010		0.044		0.512	0.274	0.016
<i>Duque et al, 2012</i>							0.042	0.377	0.018
<i>Shriver et al, 2003</i>					0.812		0.290	0.472	0.284
<i>Shriver et al, 2003</i>				0.960	0.966	0.690			
<i>Shriver et al, 2003</i>	0.600	0.470	0.140		0.502		0.583	0.537	0.094
<i>Shriver et al, 2003</i>					0.010		0.006	0.078	0.964
<i>Seldin et al, 2007</i>					0.860		0.054	0.406	0.077
<i>Seldin et al, 2007</i>							0.814	0.540	0.730
<i>Duque et al, 2012</i>							0.023	0.328	0.009
<i>Duque et al, 2012</i>							0.008	0.340	0.003
<i>Seldin et al, 2007</i>	0.010	0.890	0.013				0.108	0.406	0.075
<i>Shriver et al, 2003</i>							0.125	0.352	0.157
<i>Ruiz Linares et al, 2014</i>	0.200	0.880	0.000				0.030	0.400	0.003
<i>Ruiz Linares et al, 2014</i>	0.900	0.890	0.100				0.080	0.390	0.102
<i>Ruiz Linares et al, 2014</i>	0.120	0.960	0.080				0.153	0.490	0.130
<i>Seldin et al, 2007</i>							0.337	0.560	0.327
<i>Seldin et al, 2007</i>					0.100		0.048	0.112	0.847
<i>Duque et al, 2012</i>							0.889	0.791	0.114
<i>Seldin et al, 2007</i>							0.865	0.779	0.068
<i>Seldin et al, 2007</i>	0.990	0.960	0.210				0.981	0.904	0.266
<i>Seldin et al, 2007</i>					0.015		0.007	0.073	0.915
<i>Seldin et al, 2007</i>							0.993	0.924	0.253

Tabla 2. Etapas de las PCR para APLP y RFLP. Se detalla la temperatura (°C), cantidad de ciclos y tiempo de cada etapa.

Técnica Molecular	Etapa	Temperatura (°C)	Ciclos	Tiempo
PCR-APLP	Desnaturalización inicial	94	1	5 minutos
	Desnaturalización	94		1 minuto
	Hibridación	Variable	30-35	1 minuto
	Polimerización	72		1 minuto
	Extensión final	72	1	5 minutos
PCR - RFLP	Desnaturalización inicial	94	1	5 minutos
	Desnaturalización	94		45 segundos
	Hibridación	Variable	33	45 segundos
	Polimerización	72		45 segundos
	Extensión final	72	1	10 minutos

Tabla 3. Cebadores utilizados en cada múltiplex y APLP. Se detallan los cebadores utilizados para cada marcador dentro de las distintas múltiplex, junto con el tamaño esperado para cada amplicón, la concentración de cebadores para una preparación 20 veces concentrada, y la temperatura (°C) y cantidad de ciclos para su PCR. En letras minúsculas y negrita se indican las bases adicionadas para diferenciar los amplicones por tamaño.

Múltiplex	dbSNP rs	Marcador	Cebador	Secuencia (5'-3')	Tamaño del amplicón	Concentración cebador 20x	T° de hibridación / cantidad de ciclos
Múltiplex I	7211306	OGFOD3	F1/G	ATGTATAATCTcGGCAAACG	70	5 µM	52° / 35 ciclos
			F2/A	atttaATGTATAATCTcGtCAAAC A	75	5 µM	
			R	AACATTAtACACaCAGGACC		10 µM	
	13069719	FOXP1	F1/C	TTAGCATGGACCATTTTTCAC	81	4 µM	
			F2/T	tataatTTAGCATGtACCATTTTC AT	86	4 µM	
			R	GATTACCGAGGATATTTTAGG		4 µM	
	2288697	KLHL29	F1/A	TGGATAGATCCtAcAGACAA	95	20 µM	
			F2/G	ttaatTGGATAGATCCtAcAGACA G	100	15 µM	
			R	ACTCACTAATAaTCTTAGGG		15 µM	
	2065982	AMER1	F1/C	TTCAAGTTCTTCctAAGGAAAC	155	6 µM	
			F2/T	aatatTTCAAGTTCTTCctAAGGA AAT	160	6 µM	
			R	CAAGATGGGATTTCTTaCCTGCC		6 µM	
Múltiplex II	1426654	SLC24A5	F1/A	TGTCTCAGGATGTTaCAGGCA	81	5 µM	52° / 35 ciclos
			F2/G	ttagatTGTCTCAGGATGTTaCAG GCG	87	5 µM	
			R	ATgTACCTAGGAgAGCAGTA		5 µM	
	8009244	DLGAP5	F1/G	TATTACCCATCTTAGCAAGG	100	9 µM	
			F2/A	tataatTATTACCCATCTTaCAAG A	105	5 µM	
			R	GCCAATGCTTTTTTGACACC		5 µM	
	12469563	AMER4	R1/C	ACCCACCctATGAGATAAAAG	166	15 µM	
			R2/T	octacACCCACCtATGAGATAAA AA	171	15 µM	
			F	CAAAGTATTAaTGTTCCGTAA		15 µM	
	16897585	AMER2	F1/G	CAAGGCAATTATCTTgTTCCG	201	9 µM	
			F2/A	ttattaCAAGGCAATTATCTTgTT CCA	207	5 µM	
			R	ATTAGCATCTCCTTCGCAGC		5 µM	

Tabla 3 (continúa). Cebadores utilizados en cada múltiplex y APLP. Se detallan los cebadores utilizados para cada marcador dentro de las distintas múltiplex, junto con el tamaño esperado para cada amplicón, la concentración de cebadores para una preparación 20 veces concentrada, y la temperatura (°C) y cantidad de ciclos para su PCR. En letras minúsculas y negrita se indican las bases adicionadas para diferenciar los amplicones por tamaño.

			F1/T	CTCATgATCAGCTGCCTATGT	71	4 μM	
	596985	ROR1	F2/C	ttatatCTCATgATCAGCTgCTA TGC	77	4 μM	
			R	GACAGCCCCTGATTGGAGAG		4 μM	
			R1/T	CCAGAgCTCACTAGCTCTCA	100	5 μM	60° / 35 ciclos
	5025718	NOTCH2	R2/C	aattaCCAGAActCACTAGCTCTC G	105	5 μM	
			F	ACACCCTGATAGCCTGGGAC		5 μM	
			F2/C	TGTGAACAGGAtCAGATGACC	181	10 μM	
Múltiplex III	4646	CYP19-E2	F1/A	ttataTGTGAACAGGAtCAGATGA CA	186	10 μM	
			R	AGAAGGCTGaTCAGTACCCA		10 μM	
			F1/G	GAAGCGCAAGGCAtGAAGGG	220	5 μM	60° / 35 ciclos
	2278354	ROPN1L	F2/T	aataaGAAGCGCAAGGCAtGAAGG T	225	5 μM	
			R	GACACTAGATTCTGCTGGGC		5 μM	
			F1/G	GTGTTTTGTTTCTTTAAAtTGGCG AttaGTGTTTTGTTTCTTTAAAtTG GCT	143	15 μM	
	2384319	KIF3C	F2/T		147	15 μM	60° / 35 ciclos
			R	AAAATTGCTAAGGtAAGGCT		15 μM	
			R1/C	CAGGACAGCaTCCCAGTGAG	80	4 μM	
	1871534	SLC39A4	R2/G	aatataCAGGACAGCaTCCctAGT GAC	86	4 μM	
			F	GCTGCAaGGGGGTCACCCAC		4 μM	
			F1/G	CACTTGCCACAtTCAGACG	130	15 μM	
Múltiplex IV	2763	WI-9231	F2/C	ttaataCACTTGCCACAtTCAGA CC	136	15 μM	65° / 35 ciclos
			R	CACCCTGCCAtTCTCCTGG		15 μM	
			F1/G	TGCTCgTCTTCATGCACCCG	151	4 μM	
	984303	PTPRO	F2/A	taagtTGCTCgTCTTCATGCACCC A	156	4 μM	
			R	gGGGTGCAGgGCTGTTCATT		4 μM	
			F1/G	ATTTGTGTTGGGTGGaCAAGG atattATTTGTGTTGGGTGGaCAA GA	47	9 μM	
	2695	WI-11909	F2/A		52	5 μM	
			R	GTCTCCTCTGAGATTTTCTG aattattGCCAGCAGTGGtTATC ACT	107	5 μM	65° / 35 ciclos
			F1/T		100	9 μM	
	734780	EURO	F2/C	GCCAGCAGTGGGTATCACC		9 μM	
			R	GGGAGAGGCACgAGGGgAT		5 μM	65° / 35 ciclos
			R1/G	AAGTGAGAAAGTGTGTCTTC taattAAGTGAGAAAGTGTaTCTT A	130	9 μM	
			R2/T		135	5 μM	
APLP Individuales	9290363	MECOM	F	ATTACAGGTGAAGGTATTC		5 μM	54° / 30 ciclos
			F1/A	CCAGGAATCAAGTCGCCTTAA ataattCCAGaAATCAAGTCGaCT TAC	71	9 μM	
			F2/C		77	5 μM	
	10510511	***	R	AGGGAAGTGATgTACCTACA		5 μM	54° / 35 ciclos
			F1/T	CTGGGTACCATAGACgACAT ataTtaCTGaGTACCATAGACgAC AC	60	9 μM	
			F2/C		66	5 μM	
	3870336	DAG1	R	TGAaGAGaCCTTGACTGCCT		5 μM	54° / 35 ciclos
			F1/A	CAAGACCTTAAAgCCTATGGA attaaCAAGACCTTAAATCCTATG GG	111	9 μM	
			F2/G		116	5 μM	
	2042314	***	R	CCTgGCTCTAGTCTCATAT		5 μM	54° / 35 ciclos

Tabla 4. Cebadores utilizados en PCR-RFLP e indel. Se detallan los cebadores utilizados para cada marcador dentro de las PCR-RFLP, junto con la enzima y las temperaturas (°C) para su correcta digestión. Además, se encuentran detallados los tamaños (pb) de los fragmentos esperados tanto si corta como cuando no corta la enzima.

dbSNP rs	Marcador	Secuencia (5'-3')	T° de melting	Enzima	T° de digestión	Tamaño del fragmento
3138522	APO	AAGTGCTGTAGG CCATTTAGATTA G AGTCTTCGATGA CAGCGTATACAG A	60	-	indel	110 /410
7041	GC*S	AGATCTGAAATG GCTATTATTTTG C GGAGGTGAGTTT ATGGAACAGC	60	HaeIII	37°	200/ 121+79
1800498	DRD2	TGGGGGTGTGAA GAAAAGAG TTTAGTAGCAGA GGAAGGAGTGG	60	aTaq I	65	70 / 40+30
2816	WI-7423	CTCCTTGGCAGG GATTTGT CTGGTATCCACG GTGCAAG	60	Sma I	25°-30°	80 / 46+34
17203	WI-11153	CTTCAAATTGCT TTAAGTAC ATCCAACAGTCA AGGTCTAC	49	Bsa A I	37°	45 / 28+17
285	LPL	GCTTAATTCTCA ATTCAATG CTTTAGACTCTT GTCCAGGT	50	PvuII	37°	189 / 149+40
2814778	FY*NULL	AGGCTTGTGCAG GCAGTG GGCATAGGGATA AGGGAC	54	Sty I	37°	223= 82+77+62 (-) / 82+65+62+12 (+)
1042602	TYR192	TTATGTGTCAAT GGATGCAC GCTTCATGGGCA AAATCAAT	58	Mbo I	37°	69 / 32+37

RESULTADOS

Como se ha señalado anteriormente, habiendo partido de 40 AIMS que cumplían las condiciones de poder informativo ya descritas, se han elegido 30, los cuales fueron seleccionados dado que eran los que mejor podían ser agrupados en las reacciones multiplex, 22 fueron mediante la técnica de PCR-APLP múltiplex, siete mediante PCR-RFLP y uno por determinación de inserción Alu (INDEL). Esto disminuyó el tiempo de tipificación, el gasto y el uso de equipo especializado. En la Tabla 5 se presentan los resultados de las frecuencias alélicas obtenidas en la población de estudio (AMBA) mediante las tres técnicas para los 30 AIMS genotipificados. En la Figura 1 se presentan las frecuencias en las poblaciones parentales.

Para cinco marcadores (DLGAP5, WI-9231, OGFOD3, KIF3C y NOTCH2) se observaron diferencias significativas que indicaron un desvío del equilibrio de Hardy-Weinberg (H-W). Por esta razón, y con el objetivo de analizar el efecto de estos marcadores sobre la estimación de ancestría individual y poblacional, se realizaron análisis y comparaciones posteriores con diferentes conjuntos de marcadores, constituidos de la siguiente manera: 1) grupo A: 30 AIMS; 2) grupo B: 25 AIMS, excluyendo los cinco AIMS que presentaron desvíos del equilibrio H-W; y 3) grupo C: 25 AIMS, incluyendo los cinco marcadores en desvío del equilibrio H-W y excluyendo cinco AIMS con poder discriminatorio semejante a los marcadores que presentaban desvíos del equilibrio: dos del componente euroasiático (WI-7423 y

Tabla 5. Frecuencias alélicas de 30 marcadores biparentales (AIMs, Marcadores Informativos de Ancestría) empleados en la población de estudio (AMBA).

dbSNP rs	AIMs	Alelo referencia	Frecuencia alélica AMBA
734780	EURO	C	0,250
1426654	SLC24A5	A	0,885
2278354	ROPN1L	T	0,865
1800498	DRD2	T	0,560
17203	WI-11153	G	0,570
2816	WI-7423	T	0,520
1042602	TYR192	A	0,390
2695	WI-11909	G	0,694
2288697	KLHL29	A	0,364
2065982	AMER1	C	0,011
13069719	FOXP1	C	0,890
12469563	AMER4	C	0,078
16897585	AMER2	A	0,090
4646	CYP19-E2	A	0,606
2384319	KIF3C	G	0,306
2763	WI-9231	G	0,090
10510511	10510511	A	0,290
3870336	3870336	A	0,180
2042314	2042314	A	0,260
7211306	OGFOD3	A	0,179
9290363	MECOM	G	0,090
8009244	DLGAP5	A	0,920
5025718	NOTCH2	C	0,840
596985	ROR1	T	0,800
1871534	SLC39A4	C	0,080
984303	PTPRO	A	0,963
3138522	APO	inserción	0,990
285	LPL	T	0,450
7041	GC*S	G	0,490
2814778	FY NULL	C	0,011

TYR192), dos del indígena (CYP19-E2 y WI-11909) y uno del componente afro subsahariano (APO).

Los resultados obtenidos con STRUCTURE son mostrados en la Tabla 6. Para el Grupo A también se estimó la mezcla génica utilizando ADMIX, dando valores similares (81,9%, 13,6% y 4,4%, para los componentes euroasiático occidental, indígena y subsahariano, respectivamente). Las ancestrías individuales obtenidas con cada conjunto de 25 AIMs y la calculada a partir de los 30 AIMs para cada componente por

separado (euroasiático, indígena y afro subsahariano) correlacionaron significativamente (Tabla 7).

Asimismo, no hubo diferencias significativas entre la ancestría poblacional obtenida con AIMs de los grupos A, B y C (Tabla 8). No obstante, pudo observarse que más importante que la inclusión de AIMs en desvío del equilibrio de Hardy-Weinberg fue el efecto de la reducción de cinco marcadores. Esto se visualiza en el aumento del componente minoritario (en este caso el subsahariano), tal como había sido simulado por Russo *et al.* (2016).

Tabla 6. Estimación de mezcla génica poblacional obtenida con el programa STRUCTURE 2.3.4 (Pritchard *et al.* 2000) para el grupo de marcadores A, B y C, usando 50 genotipos de la población de Buenos Aires.

Grupo de marcadores	Componente Euroasiático	Componente Indígena	Componente Afro subsahariano
Grupo A	84,00%	12,50%	3,40%
Grupo B	83,00%	12,40%	4,15%
Grupo C	84,20%	11,70%	4,00%

Tabla 7. Coeficientes de Pearson para la correlación entre ancestría individual calculada con cada conjunto de marcadores, grupo A, B y C. En todos los casos $p < 0,016$ luego de corrección por Bonferroni.

Par de grupos comparados	Componente Euroasiático	Componente Indígena	Componente Afro subsahariano
A-B	0,98	0,99	0,92
A-C	0,96	0,97	0,96
B-C	0,92	0,96	0,86

Tabla 8. Diferencia de proporciones obtenida con el programa Infostat para cada comparación entre porcentajes de ancestría poblacional obtenida con cada grupo de marcadores A, B y C.

Inferencia basada con dos grupos de marcadores	Componente Euroasiático		Componente Indígena		Componente Afro subsahariano	
	Diferencia de proporciones	<i>p</i>	Diferencia de proporciones	<i>p</i>	Diferencia de proporciones	<i>p</i>
A-B	0,01	0,85	0	1	0,01	0,72
A-C	0	1	0,01	0,82	0,01	0,72
B-C	0,01	0,85	0,01	0,82	0	1

DISCUSIÓN

En este estudio identificamos y seleccionamos un conjunto de 30 marcadores informativos para dar cuenta de los aportes europeos, indígena y subsaharianos, cantidad recomendable a partir del análisis ya comentado realizado por Russo *et al.* (2016), en el cual se observó que el error de las estimaciones descendía sensiblemente a medida que se iban incorporando AIMS, pero luego de los 30 marcadores esa disminución se ralentizaba. Por esa razón se decidió que esa cantidad representa una buena relación entre la confiabilidad del resultado, el costo y la capacidad de procesamiento de muestras.

El set es adecuado para la caracterización de la población, con marcadores que no enmascaran ninguno de los componentes presentes. Esto constituye un aporte en el contexto de estudios en ciudades cosmopolitas argentinas, pues fueron elegidos para dar cuenta de los tres componentes a nivel continental, que como ya señalamos se encuentran presentes desde la época colonial, cuando además comenzó el proceso de mestizaje (Martínez Sarasola, 2011). Luego de la consolidación del Estado nacional, importantes contingentes migratorios se asentaron principalmente en la zona central de Argentina en el período 1880-1930, particularmente en Buenos Aires. La ciudad contaba con alrededor de

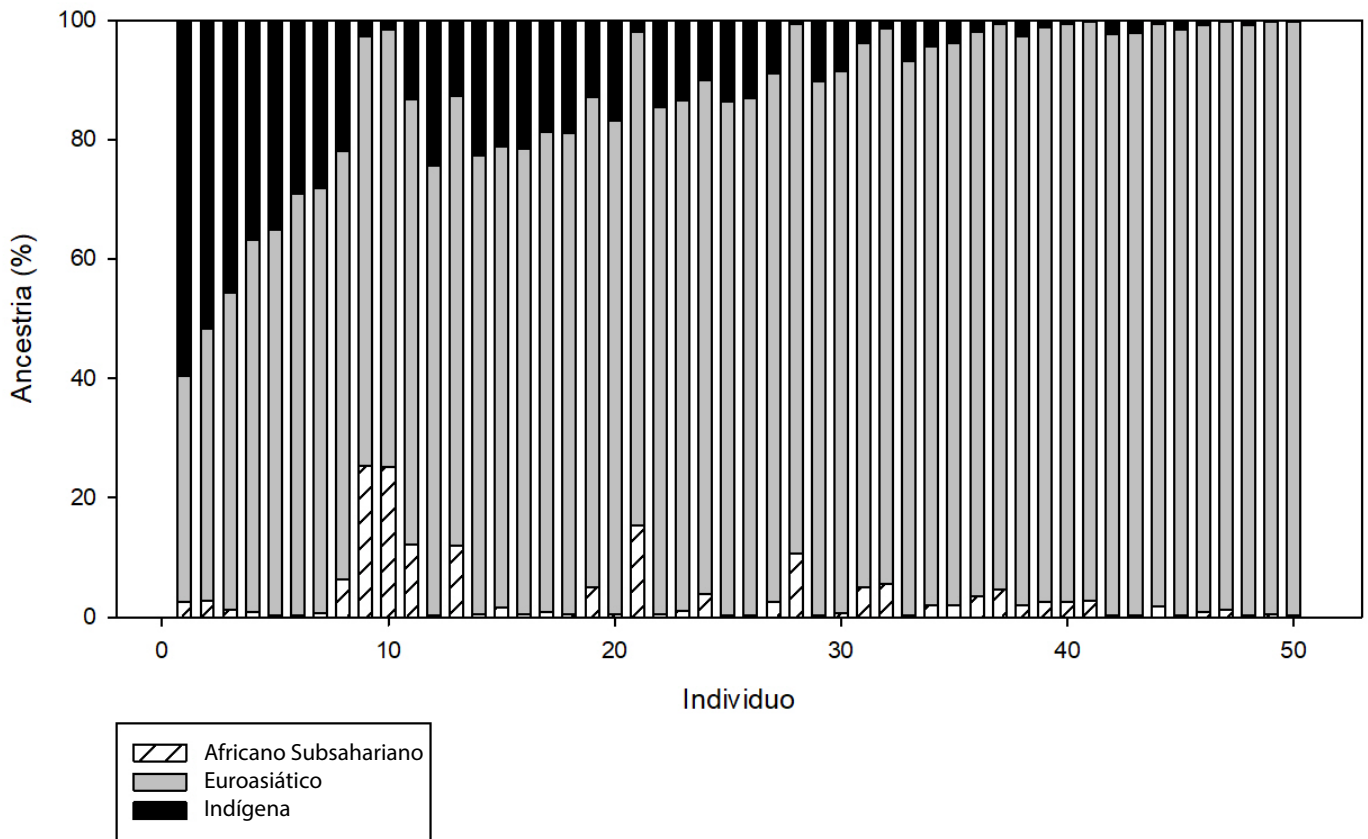


Figura 1. Estimación de los componentes de ancestría individual (Africano Subsahariano, Euroasiático, Indígena) en la población de Buenos Aires. Cada individuo es representado por una barra vertical sobre el eje X.

70.000 habitantes hacia mediados del siglo XIX, y posteriormente recibió una importante inmigración europea, principalmente hasta la finalización de la segunda guerra mundial. De un total aproximado de tres millones y medio de inmigrantes de ese origen al país, un tercio se radicó en la ciudad de Buenos Aires. A partir de la década de 1940 y concomitantemente a la disminución del arribo de europeos, se registró un aumento de la cantidad de migrantes internos y de los países vecinos de elevado componente hispano-indígena (Avena *et al.*, 2006).

El gran flujo de migrantes de origen europeo quedó evidenciado con los resultados obtenidos en el presente estudio, donde se registraron valores del 84%, 12,5% y 3,5%, para los componentes euroasiático occidental, indígena y subsahariano, respectivamente, empleando el programa STRUCTURE. Como mencionamos anteriormente, utilizando el programa ADMIX, esos valores fueron del 82%, 13,6% y 4,4%. Trabajos previos de nuestro equipo empleando este último y sobre muestras provenientes de los mismos hospitales habían arrojado valores similares, del 80,8%, 15,4% y 3,8% para los componentes euroasiático occidental, indígena y subsahariano, respectivamente (Avena *et al.*, 2006).

En relación a la ancestría individual cabe destacar que más del 96% de las muestras analizadas presentaron

un componente europeo mayor al 50%. En cuanto al componente indígena, 22 personas (44%) presentaron más de un 10% de ese aporte. Por último, el componente subsahariano fue el minoritario, pero aun así seis personas (12%) superaron el 10%. Los resultados obtenidos coinciden con trabajos previos realizados por Avena *et al.* (2012) y Bobillo y Corach (2012) en Buenos Aires, mostrando que el componente europeo es mayoritario a nivel poblacional, mientras que a nivel individual se observa una heterogeneidad de los tres componentes ancestrales, evidenciando procesos de intenso mestizaje y justificando la relevancia de realizar la determinación de ancestría a nivel individual.

Hemos observado cinco marcadores que no estaban en equilibrio de Hardy-Weinberg. No necesariamente esto debe ser motivo de exclusión, porque esa diferencia puede deberse a la acción de mecanismos microevolutivos, como puede ser el caso si hay estratificación poblacional. En esos cinco marcadores se ha registrado un déficit de heterocigotos, lo cual puede indicar elección preferencial de pareja. Doldan (2014) analizó los patrones de elección de pareja en 107 matrimonios de la ciudad de La Plata, encontrando una correlación significativamente positiva entre los cónyuges respecto al nivel de instrucción, información genealógica y 10 marcadores genéticos biparentales.

En las comparaciones realizadas entre los resultados obtenidos de ancestría (individual y poblacional) obtenida con 30 AIMS y 25 AIMS (con y sin marcadores en desvío del equilibrio, respectivamente) no se observaron diferencias significativas, no obstante más importante que la inclusión de AIMS en desvíos del equilibrio de Hardy-Weinberg fue el efecto de la reducción de cinco marcadores, porque a medida que se disminuye su cantidad aumenta el componente minoritario (en este caso el subsahariano).

En definitiva, presentamos la selección de un set de 30 AIMS, herramienta molecular de bajo costo y de alta eficiencia, para determinar la ancestralidad genética individual y poblacional para la determinación de la mezcla génica en una población latinoamericana, considerando las particularidades de la población de la Argentina. Se obtuvieron resultados concordantes con estudios que han empleado un número mayor de marcadores, detectando similares aportes continentales de los tres principales componentes, lo que evidencia el poder informativo de los AIMS seleccionados. Esto resulta de aplicación tanto en genética de poblaciones humanas como en el campo de la biomedicina.

BIBLIOGRAFÍA

- Avena S., Di Fabio Rocca F., Postillone M.B., Dejean C. (2013) ¿Existió el crisol de razas en Argentina? Una respuesta desde la antropogenética. En: Gurevich B. (Ed.) *La Complejidad Después De Babel*. Lumiere, Buenos Aires, pp. 281–314.
- Avena S.A., Goicoechea A.S., Dugoujon J.M., Slepoy M.G., Slepoy A.S., Carnese F.R. (2001) Análisis antropogenético de los aportes indígena y africano en muestras hospitalarias de la ciudad de Buenos Aires. *Rev. Argent. Antropol. Biológica* 3: 79–99.
- Avena S., Goicoechea A.S., Rey J., Dugoujon J.M., Dejean C.B., Carnese F.R. (2006) Mezcla génica en una muestra poblacional de la ciudad de Buenos Aires. *Medicina (Mex.)* 66: 113–118.
- Avena S., Parolín M.L., Boquet M., Dejean C.B., Postillone M.B., Trentini Y.A., Rocca F., Mansilla F., Jones L., Dugoujon J., Carnese F.R. (2010) Mezcla génica y linajes uniparentales en Esquel (Pcia. de Chubut): Su comparación con otras muestras poblacionales argentinas. *BAG J. Basic Appl. Genet.* 21: 01–14.
- Avena S., Parolín M.L., Dejean C.B., Fabrykant G., Ríos Part M.D.C., Goicoechea A.S., Dugoujon J.M., Carnese F.R. (2009) Mezcla génica y linajes uniparentales en Comodoro Rivadavia (Provincia de Chubut, Argentina). *Rev. Argent. Antropol. Biológica* 11: 25–41.
- Avena S.A., Russo M.G., Lozano M., Di Fabio Rocca F. (2024) El uso de categorías raciales en la biomedicina argentina. Un análisis de publicaciones nacionales e internacionales. *Rev. Urug. Antropol. Etnografía* 9 (2).
- 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., Beckman K., Burchard E.G., Parolín M.L., Goicoechea A., Acreche N., Boquet M., Ríos Part M.D.C., Fernández V., Rey J., Stern M.C., Carnese R.F., Fejerman L. (2012) Heterogeneity in genetic admixture across different regions of Argentina. *PLoS One* 7, e34695. <https://doi.org/10.1371/journal.pone.0034695>
- Belini C., Korol J. (2019) *Historia económica de la Argentina en los siglos XX y XXI*. Siglo XXI Editores Argentina S.A., Buenos Aires, Argentina.
- Bobillo M.C., Corach D. (2012) Desplazamiento étnico en la región bonaerense. Empleo de marcadores genéticos en la detección del nivel de heterogeneidad genética en la población actual. *Rev. Mus. Antropol.* 5: 45–52. <https://doi.org/10.31048/1852.4826.v5.n1.9124>
- Bonilla C., Parra E.J., Pfaff C.L., Dios S., Marshall J.A., Hamman R.F., Ferrell R.E., Hoggart C.L., McKeigue P.M., Shriver M.D. (2004) Admixture in the Hispanics of the San Luis Valley, Colorado, and its implications for complex trait gene mapping. *Ann. Hum. Genet.* 68: 139–153. <https://doi.org/10.1046/j.1529-8817.2003.00084.x>
- Carnese F.R. (2019) *El mestizaje en la Argentina: indígenas, europeos y africanos: una mirada desde la antropología biológica*. Editorial de la Facultad de Filosofía y Letras, Universidad de Buenos Aires, Buenos Aires, Argentina.
- Chakraborty R. (1985) Gene identity in racial hybrids and estimation of admixture rates. In: Neel J.B.; Ahuja Y. (Eds.) *Genetic Microdifferentiation in Man and Other Animals*. Indian Anthropological Association, Delhi, pp. 171–180.
- Chamzas A., Tellez E., SyBing A., Gobburu J.V.S., Gopalakrishnan M. (2024) Optimizing tacrolimus dosing in Hispanic renal transplant patients: insights from real-world data. *Frontiers in pharmacology*, 15: 1443988. <https://doi.org/10.3389/fphar.2024.1443988>
- Corach D., Lao O., Bobillo C., van Der Gaag K., Zuniga S., Vermeulen M., van Duijn K., Goedbloed M., Vallone P.M., Parson W., de Knijff P., Kayser M. (2010) Inferring continental ancestry of argentineans from autosomal, Y-chromosomal and mitochondrial DNA. *Ann. Hum. Genet.* 74(1): 65–76. <https://doi.org/10.1111/j.1469-1809.2009.00556.x>
- Di Fabio Rocca F., Albeza M.V., Postillone M.B., Acreche N., Lafage L., Parolín M.L., Dejean C.B., Carnese F.R., Avena S. (2016) Historia poblacional y análisis antropogenético de la ciudad de Salta. *Andes* 27(2).
- Di Francia R., Crisci S., De Monaco A., Cafiero C., Re A., Iaccarino G., De Filippi R., Frigeri F., Corazzelli G., Micera A., Pinto A. (2021) Response and Toxicity to Cytarabine Therapy in Leukemia and Lymphoma: From Dose Puzzle to Pharmacogenomic Biomarkers. *Cancers* 13(5): 966. <https://doi.org/10.3390/cancers13050966>
- Di Rienzo J., Casanoves F., Balzarini M., Gonzalez L., Tablada M., Robledo C. (2003) *InfoStat Statistical Software*, Universidad Nacional de Córdoba, Argentina.
- Doldan P. (2014) *Patrones de elección del cónyuge y marcadores genéticos. Un análisis bioantropológico*. Tesis de Licenciatura en Cs. Biológicas, Universidad Maimónides, Buenos Aires.
- Duque C., Parra M.V., Valencia A.V., Bedoya G., Ruiz-Linares A. (2012) Comparación de cuatro programas utilizados en la determinación de la composición genética ancestral de la población antioqueña. *Rev. Colomb. Antropol.* 48: 233–257.
- Evanno G., Regnaut S., Goudet J. (2005) Detecting the number of clusters of individuals using the software structure: A simulation study. *Molecular Ecology* 14(8): 2611–2620. <https://onlinelibrary.wiley.com/doi/10.1111/j.1365-294X.2005.02553.x>
- Fejerman L., Carnese F.R., Goicoechea A.S., Avena S., Dejean C.B., Ward R.H. (2005) African ancestry of the population of Buenos Aires. *Am. J. Phys. Anthropol.* 128: 164–170. <https://doi.org/10.1002/ajpa.20083>
- Galanter J.M., Fernandez-Lopez, J.C., Gignoux, C.R., Barnholtz-Sloan, J., Fernandez-Rozadilla, C., Via, M., ... Carracedo A., for the Lace Consortium (2012) Development of a Panel of Genome-Wide Ancestry Informative Markers to Study Admixture Throughout the Americas. *PLOS Genet.* 8: e1002554. <https://doi.org/10.1371/journal.pgen.1002554>

- García A., Dermarchi D.A., Tovo-Rodrigues L., Pauro M., Callegari-Jacques S.M., Salzano F.M., Hutz M.H. (2015) High interpopulation homogeneity in Central Argentina as assessed by Ancestry Informative Markers (AIMs). *Genet. Mol. Biol.* 38: 324–331. <https://doi.org/10.1590/S1415-475738320140260>
- Gómez Pérez L., Alfonso Sánchez M.A., Pérez Miranda A.M., Martínez de Pancorbo Gómez M., Peña García J.A. (2007) Utilidad de las inserciones Alu en los estudios de mestizaje. *Antropo* 14: 29–36.
- Hoffman J., Fejerman L., Hu D., Huntsman S., Li M., John E.M., Torres-Mejía G., Kushi L., Ding Y.C., Weitzel J., Neuhausen S.L., Lott P., COLUMBUS Consortium, Echeverry M., Carvajal-Carmona L., Burchard E., Eng C., Long J., Zheng W., Olopade O., ... Ziv E. (2019) Identification of novel common breast cancer risk variants at the 6q25 locus among Latinas. *Breast Cancer Research* 21(1): 3. <https://doi.org/10.1186/s13058-018-1085-9>
- Jurado Medina L.S., Muzzio M., Schwab M., Costantino M.L., Barreto G., Bailliet G. (2014) Human Y-chromosome SNP characterization by multiplex amplified product-length polymorphism analysis. *Electrophoresis* 35: 2524–2527. <https://doi.org/10.1002/elps.201400020>
- Lewontin R.C. (1972) The apportionment of human diversity. *Evolutionary Biology* 6: 381–398.
- Martínez Marignac V.L., Bertoni B., Parra E.J., Bianchi N.O. (2004) Characterization of admixture in an urban sample from Buenos Aires, Argentina, using uniparentally and biparentally inherited genetic markers. *Hum. Biol.* 76: 543–557. <https://doi.org/10.1353/hub.2004.0058>
- Martínez Sarasola C. (2011) Nuestros paisanos los indios. Editorial Del Nuevo Extremo S.A., Buenos Aires, Argentina.
- Mc Clain C.A., Bernhardt M.B., Berger A., Bernini J.C., Marquez-Do D., Winslow R., Scheurer M.E., Schafer E.S. (2017) Pharmacogenetic association with neurotoxicity in Hispanic children with acute lymphoblastic leukaemia. *British Journal of Haematology* 181(5): 684–687. <https://doi.org/10.1111/bjh.14678>
- Muzzio M., Motti J.M.B., Paz Sepulveda P.B., Yee M.C., Cooke T., Santos M.R., Ramallo V., Alfaro E.L., Dipierri J.E., Bailliet G., Bravi C.M., Bustamante C.D., Kenny E.E. (2018). Population structure in Argentina. *PloS One* 13(5): e0196325. <https://doi.org/10.1371/journal.pone.0196325>
- Park S.Y., Wilkens L.R., Kolonel L.N., Monroe K.R., Haiman C.A., Marchand L.L. (2017) Exploring Differences in the Aspirin–Colorectal Cancer Association by Sex and Race/Ethnicity: The Multiethnic Cohort Study. *Cancer Epidemiology, Biomarkers y Prevention* 26(2): 162–169. <https://doi.org/10.1158/1055-9965.EPI-16-0560>
- Parolin M.L., Toscanini U.F., Velázquez I.F., Llull C., Berardi G.L., Holley A., Tamburrini C., Avena S., Carnese F.R., Lanata J.L., Sánchez Carnero N., Arce L.F., Basso N.G., Pereira R., Gusmão L. (2019) Genetic admixture patterns in Argentinian Patagonia. *PloS One* 14: e0214830. <https://doi.org/10.1371/journal.pone.0214830>
- Patiño J. (2020) Determinación de Marcadores Informativos de Ancestría mediante Multiplex PCR– APLP. Tesis de Maestría Biología Molecular Médica, Universidad de Buenos Aires, Buenos Aires.
- Peakall R., Smouse P.E. (2012) GenAEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28(19): 2537–2539. <https://doi.org/10.1093/bioinformatics/bts460>
- Pena S.D.J., Santos F.R., Tarazona-Santos E. (2020) Genetic admixture in Brazil. *Am. J. Med. Genet. C Semin. Med. Genet.* 184: 928–938. <https://doi.org/10.1002/ajmg.c.31853>
- Peterson R.E., Kuchenbaecker K., Walters R.K., Chen C.Y., Popejoy A.B., Periyasamy S., Lam M., Iyegbe C., Strawbridge R.J., Brick L., Carey C.E., Martin A.R., Meyers J.L., Su J., Chen J., Edwards A.C., Kalungi A., Koen N., Majara L., Schwarz E., ... Duncan L.E. (2019) Genome-wide Association Studies in Ancestrally Diverse Populations: Opportunities, Methods, Pitfalls, and Recommendations. *Cell* 179(3): 589–603. <https://doi.org/10.1016/j.cell.2019.08.051>
- Pritchard J.K., Stephens M., Donnelly P. (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959. <https://doi.org/10.1093/genetics/155.2.945>
- Reitz C., Pericak-Vance M. A., Foroud T., Mayeux R. (2023) A global view of the genetic basis of Alzheimer disease. *Nature reviews. Neurology* 19(5): 261–277. <https://doi.org/10.1038/s41582-023-00789-z>
- Rúa O., Larrayoz I.M., Barajas M.T., Velilla S., Martínez A. (2012) Oral doxycycline reduces pterygium lesions; Results from a double blind, randomized, placebo controlled clinical trial. *PLoS ONE* 7(12): e52696. <https://doi.org/10.1371/journal.pone.0052696>
- Ruderman A., Paschetta C., González-José R., Ramallo V. (2020) Origen étnico y ancestría genética, ¿qué lugar tienen en la práctica médica en Argentina? *Revista Argentina De Antropología Biológica* 22(2): 020. <https://doi.org/10.24215/18536387e020>
- Ruiz-Linares A., Adhikari K., Acuña-Alonso V., Quinto-Sanchez M., Jaramillo C., Arias W., ... Gonzalez-José R. (2014) Admixture in Latin America: geographic structure, phenotypic diversity and self-perception of ancestry based on 7,342 individuals. *PLoS Genet.* 10: e1004572. <https://doi.org/10.1371/journal.pgen.1004572>
- Ruiz-Narváez E.A., Rosenberg L., Wise L.A., Reich D., Palmer J.R. (2011) Validation of a small set of ancestral informative markers for control of population admixture in African Americans. *Am. J. Epidemiol.* 173: 587–592. <https://doi.org/10.1093/aje/kwq401>
- Russo M.G., Di Fabio Rocca F., Doldán P., Cardozo D.G., Dejean C.B., Seldes V., Avena S. (2016) Evaluación del número mínimo de marcadores para estimar ancestría individual en una muestra de la población argentina. *Rev. Mus. Antropol.* 9(1): 49–56. <https://doi.org/10.31048/1852.4826.v9.n1.12579>
- Sáez Capel J (2001) Los migrantes y la discriminación en Argentina. *Scripta Nova. Revista electrónica de Geografía y Ciencias Sociales* 94 (31). <https://www.ub.edu/geocrit/sn-94-31.htm>
- Salzano F.M., Sans M. (2014) Interethnic admixture and the evolution of Latin American populations. *Genet. Mol. Biol.* 37: 151–170. <https://doi.org/10.1590/s1415-47572014000200003>
- Seldin M.F., Tian C., Shigeta R., Scherbarth H.R., Silva G., Belmont J.W., Kittles R., Gamron S., Allevi A., Palatnik S.A., Alvarellos A., Pairs S., Caprarulo C., Guillerón C., Catoggio L.J., Prigione C., Berbotto G.A., García M.A., Perandones C.E., Pons-Estel B.A., Alarcon-Riquelme M.E. (2007) Argentine population genetic structure: large variance in Amerindian contribution. *Am. J. Phys. Anthropol.* 132: 455–462. <https://doi.org/10.1002/ajpa.20534>
- Shriver M.D., Parra E.J., Dios S., Bonilla C., Norton H., Jovel C., Pfaff C., Jones C., Massac A., Cameron N., Baron A., Jackson T., Argyropoulos G., Jin L., Hoggart C.J., McKeigue P.M., Kittles R.A. (2003) Skin pigmentation, biogeographical ancestry and admixture mapping. *Hum. Genet.* 112: 387–399. <https://doi.org/10.1007/s00439-002-0896-y>
- Suarez-Kurtz G. (2008) The implications of population admixture in race-based therapy. *Clinical Pharmacology & Therapeutics* 83(3): 399–400. <https://doi.org/10.1038/sj.clpt.6100488>
- Texidó E. (2008) Perfil Migratorio de las Migraciones. Organización Internacional para las Migraciones, Buenos Aires, Argentina.

Umetsu K., Tanaka M., Yuasa I., Adachi N., Miyoshi A., Kashimura S., Park K.S., Wei Y.H., Watanabe G., Osawa M. (2005) Multiplex amplified product-length polymorphism analysis of 36 mitochondrial single-nucleotide polymorphisms for haplogrouping of East Asian populations. *Electrophoresis* 26: 91-98. <https://doi.org/10.1002/elps.200406129>

Wang S., Ray N., Rojas W., Parra M.V., Bedoya G., Gallo C., Poletti G., Mazzotti G., Hill K., Hurtado A.M., Camrena B., Nicolini H., Klitz W., Barrantes R., Molina J.A., Freimer N.B., Bortolini M.C., Salzano F.M., Petzl-Erler M.L., Tsuneto L.T., Dipierri J.E., Alfaro E.L., Bailliet G., Bianchi N.O., Llop E., Rothhammer F., Excoffier L., Ruiz-Linares A. (2008) Geographic patterns of genome admixture in Latin American Mestizos. *PLoS Genet.* 4: e1000037. <https://doi.org/10.1371/journal.pgen.1000037>

Wheeler H.E., Gorsic L.K., Welsh M., Stark A.L., Gamazon E.R., Cox N.J., Dolan M.E. (2011) Genome-wide local ancestry approach identifies genes and variants associated with chemotherapeutic susceptibility in African Americans. *PLoS One* 6: e21920. <https://doi.org/10.1371/journal.pone.0021920>

FUENTES DE FINANCIACIÓN

Este trabajo ha recibido soporte económico del subsidio PIP CONICET 2021-2111 y de un Subsidio Intramuros de la Fundación Fiorellino-UMai.

CONTRIBUCIÓN DE LOS AUTORES

Todos los autores enumerados anteriormente han hecho contribuciones sustanciales, directas e intelectuales al trabajo y lo han aprobado para su publicación.

—