



Facultad de Ciencias Exactas

DIVERSIDAD GENÉTICA DE LA REGIÓN NO PSEUDOAUTOSÓMICA DEL CROMOSOMA X EN LA POBLACIÓN BONAERENSE

TRABAJO FINAL

Laboratorio de Procesos Biotecnológicos

Brenda Bezus

Directora: Cecilia I. Catanesi

Tutora: G. Paula Di Santo Meztler

Año 2016



AGRADECIMIENTOS

-Quisiera agradecer en primer lugar a las autoridades del IMBICE, por permitirme realizar mi trabajo final en sus instalaciones, y a todas las personas que trabajan allí, por la calidez y buena onda de todos los días.

-A las autoridades de la UNLP y de la Facultad de Ciencias Exactas, por la educación recibida.

-A mi directora, Cecilia, por su enseñanza, paciencia y dedicación en todo este tiempo, por sus consejos, y por aclarar siempre todas mis dudas.

-A mi tutora, Paula, por la ayuda en las técnicas de laboratorio de todos los días, y por los consejos.

-A los jurados de este trabajo, que aceptaron gentilmente a evaluarme.

-A mis padres, por ayudarme para seguir la carrera que elegí y aconsejarme siempre, y a mamá especialmente por sus consejos con el formato de este trabajo.

-A las chicas de “Diversidad Genética”, Pau, Rebe y Diana, por el compañerismo y la amistad de todos los días. A Lau, por el poquito tiempo compartido, y por toda la información prestada, que me ayudó muchísimo para este trabajo. También a Miri y Meli, por el tiempo compartido en el laboratorio.

-A todos los que donaron voluntariamente una muestra de saliva para que este trabajo fuese posible, y a la gente que me ayudó a conseguir más muestras.

-A las amigas que encontré en esta etapa de mi vida, compañeras de estudio y de dispersión, especialmente a Luli, Glo, Ine, Tati, Bel, Emi.. También a las amigas de toda la vida, Vane y Guada, por sus palabras alentadoras de todos los días y los mates a la hora del descanso.

-A mis hermanos, Andrés y Pablo, especialmente a Pablo por su ayuda con la parte histórica de este trabajo, y por su compañía en las horas de estudio.

-A Maxi, por su paciencia y por estar siempre. A Gaby y Luisa, por sus consejos para la nueva etapa que comienza.

-Al resto de mi familia, en especial a mi abuela y a mi tía Marita, por estar siempre interesadas en el progreso de mi carrera.

-A todas las personas que me acompañaron en esta etapa de mi vida, y a las personas que leyeron mi trabajo. Especialmente a mi mamá, Maxi y Vane, por tomarse el tiempo de leerlo e ir dándome opiniones y consejos durante la escritura.

ÍNDICE

INTRODUCCIÓN.....	1
Genética de poblaciones.....	1
El principio de Hardy-Weinberg.....	1
Marcadores moleculares.....	4
Polimorfismo de Nucleótido Simple o <i>Single Nucleotide Polimorphism</i> (SNP).....	5
Polimorfismos de Inserción-Delección o <i>Small insertions and deletions</i> (INDELS).....	6
Polimorfismos de secuencia repetida.....	7
Cromosoma X.....	7
Población estudiada.....	12
Población de la República Argentina.....	12
Población de la Provincia de Buenos Aires.....	15
OBJETIVOS.....	18
HIPÓTESIS.....	19
MATERIALES Y MÉTODOS.....	20
Población analizada.....	20
Extracción de ADN.....	21
Marcadores analizados.....	22
Polimorfismos de Nucleótido Simple.....	22
Polimorfismos de Inserción-Delección.....	31
Cálculos estadísticos.....	34
Frecuencias Alélicas.....	34
Frecuencias Genotípicas.....	35
Frecuencias Haplotípicas.....	36
Análisis de Desequilibrio de Ligamiento.....	36
Heterocigosis Observada.....	36
Heterocigosis Media.....	37

Heterocigosis Esperada y Equilibrio de Hardy-Weinberg.....	37
Diversidad Génica.....	38
Variabilidad y estructura poblacional.....	38
Índice FST de Wright.....	39
Análisis de Varianza Molecular (AMOVA).....	39
Análisis de la Estructura Poblacional.....	40
RESULTADOS.....	41
Variación dentro de la población.....	41
Polimorfismos de Nucleótido Simple.....	41
Frecuencias Alélicas.....	44
Frecuencias Genotípicas.....	47
Desequilibrio de Ligamiento.....	49
Equilibrio de Hardy-Weinberg.....	49
Haplotipos.....	49
Polimorfismos de Inserción-Delección.....	51
Frecuencias Alélicas.....	53
Frecuencias Genotípicas.....	55
Desequilibrio de Ligamiento.....	56
Equilibrio de Hardy-Weinberg.....	56
Haplotipos.....	56
Variación entre Buenos Aires y otras poblaciones.....	57
Polimorfismos de Nucleótido Simple.....	57
Índice de FST entre poblaciones (para todos los X-SNPs).....	57
Análisis de Varianza Molecular (AMOVA).....	58
Índice de FST y AMOVA para los X-SNPs de forma individual.....	59
Comparación con otras poblaciones del mundo.....	61
Polimorfismos de Inserción-Delección.....	65
Índice de FST entre poblaciones (para todos los X-INDELS).....	65
Análisis de Varianza Molecular (AMOVA).....	66

Índice de FST y AMOVA para los X-Indels de forma individual.....	67
Comparación con otras poblaciones del mundo.....	68
Análisis conjunto de todos los marcadores (INDELS + SNPs).....	70
Haplotipos.....	70
Desequilibrio de Ligamiento.....	70
Índice de FST.....	70
Análisis de Varianza Molecular (AMOVA).....	70
Análisis de la estructura poblacional.....	71
DISCUSIÓN.....	74
Variación dentro de la población de Buenos Aires.....	74
Variación entre la población de Buenos Aires y otras poblaciones argentinas.....	77
Variación entre la población de Buenos Aires y otras poblaciones del mundo.....	79
Variabilidad de Indels y SNPs.....	80
CONCLUSIONES.....	82
BIBLIOGRAFÍA.....	83
Referencias electrónicas.....	87
APÉNDICE.....	88
Apéndice 1.....	88
Apéndice 2.....	90
Apéndice 3.....	91
Apéndice 4.....	92
Apéndice 5.....	94
Apéndice 6.....	95
Apéndice 7.....	96
Apéndice 8.....	97

INTRODUCCIÓN

El estudio de la diversidad biológica y la evolución de las especies se ha valido clásicamente del análisis de rasgos morfológicos y la agrupación en grandes grupos, según la observación de diferencias y similitudes. Con el descubrimiento del ADN, y el advenimiento de técnicas que permiten su estudio, se generó una revolución en el estudio taxonómico. No sólo se pueden diferenciar especies gracias a su estudio, sino que cada individuo dentro de una especie presenta características genéticas únicas, una “huella” genética. La variación genética que existe dentro de las poblaciones es estudiada por la genética de poblaciones.

GENÉTICA DE POBLACIONES

La genética de poblaciones estudia la variación genética existente dentro y entre las poblaciones, y las fuerzas evolutivas que generan cambios en ellas. La variación genética existe gracias a la presencia de varios alelos en diferentes *loci*. Para evaluarla experimentalmente, se miden las frecuencias alélicas en los diferentes *loci* y se realiza una comparación. Estas frecuencias pueden cambiar ante factores como mutación, selección, deriva génica y flujo génico (Griffiths, 2002).

Una población ideal se define como un grupo de individuos de una misma especie, que se cruza y vive en un mismo lugar al mismo tiempo, y está aislado de las demás poblaciones. El entrecruzamiento al azar, es decir, no selectivo, convierte a la población en panmíctica.

El principio de Hardy-Weinberg

El principio de Hardy-Weinberg se postula para una población ideal, y establece que la variación genética a través de las generaciones es constante si no existen fuerzas distorsionadoras. Estas fuerzas podrían ser endogamia, deriva génica, flujo génico, mutación y selección.

Teniendo en cuenta el modelo de segregación de Mendel, las frecuencias genotípicas para un determinado locus bialélico serán los productos de las frecuencias alélicas. Si para un determinado locus existen los alelos A y a, con frecuencias alélicas de p y q, respectivamente, las frecuencias genotípicas serán el producto de las frecuencias alélicas que se combinan, como se muestra en la Figura 1 y la Tabla 1.

		Frecuencias masculinas	
		(A) p	(a) q
Frecuencias femeninas	(A) p	AA $p \times p = p^2$	Aa $q \times p = pq$
	(a) q	Aa $p \times q = pq$	aa $q \times q = q^2$

Figura 1. Frecuencias alélicas masculinas y femeninas para los alelos A y a en una población dada, y las frecuencias genotípicas resultantes en la descendencia (Pierce, 2010).

Tabla 1. Frecuencias genotípicas resultantes de la Figura 1.

Genotipo	A/A	A/a	a/a
Frecuencia genotípica	$p \times p = p^2$	$2 \times p \times q = 2pq$	$q \times q = q^2$

Esto demuestra que para una población en la que se cumplen las suposiciones del principio de Hardy-Weinberg, las frecuencias alélicas se mantienen constantes generación a generación, y se dice que la población está en equilibrio de Hardy-Weinberg para ese locus.

Las frecuencias alélicas determinan las frecuencias genotípicas, en la Figura 2 se puede ver la relación general entre las frecuencias alélicas y genotípicas para un locus en particular cuando la población se ajusta al equilibrio.

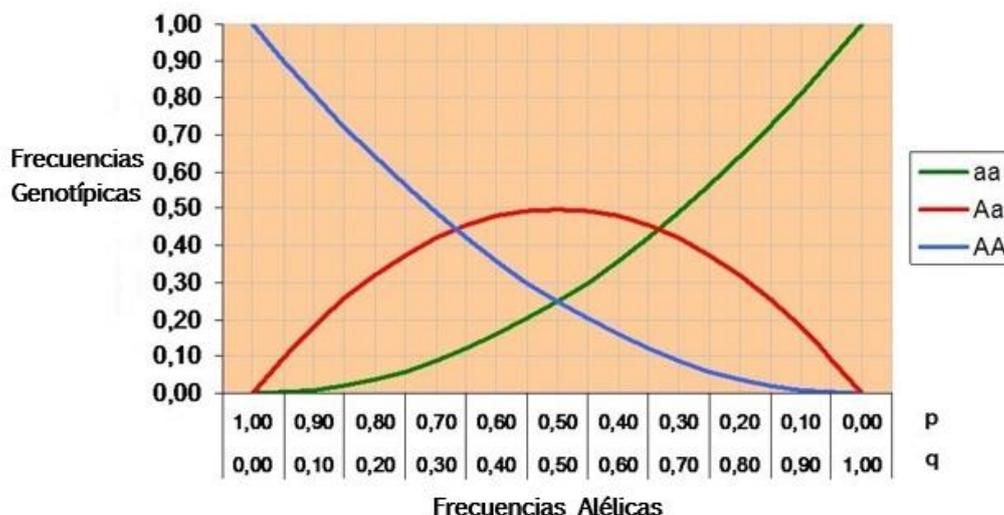


Figura 2. Curvas que muestran las proporciones de homocigotas (a/a y A/A) y heterocigotas (A/a) en función de las frecuencias alélicas de A (p) y a (q).

Si la población está en equilibrio, las frecuencias alélicas en hombres y mujeres son iguales, y conociéndolas pueden conocerse las frecuencias genotípicas.

Las poblaciones se mantienen en equilibrio siempre y cuando no existan fuerzas que provoquen cambios en la variación genética dentro de la población. Los factores que generan estos cambios son: endogamia, deriva génica, flujo génico, mutación y selección (Hedrick, 2000).

La endogamia se refiere al cruzamiento entre individuos muy cercanos genéticamente, y siempre conlleva una reducción de la variación dentro de la población.

La deriva génica es el cambio azaroso en las frecuencias alélicas que ocurre cuando las poblaciones son muy pequeñas. Se generan fluctuaciones en las diferentes generaciones que pueden llevar a la pérdida de alelos. Por este motivo también se considera que la deriva génica es un factor que reduce el valor de variación dentro de una población.

El flujo génico se relaciona con los movimientos entre grupos de poblaciones, generando un intercambio genético. En esta categoría se incluyen las migraciones.

Dependiendo de una situación en particular, la variación puede aumentar o disminuir, pero generalmente genera una fluctuación del equilibrio (Hedrick, 2000).

La mutación es la aparición de alelos nuevos en una población, y es considerada como un factor que lleva a un aumento de la variación genética.

La selección se refiere a la supervivencia diferencial de algunos organismos con caracteres que le ofrecen cierta ventaja para adaptarse al ambiente en donde viven, y provoca cambios en las frecuencias de ciertos genotipos en la población (Griffiths, 2002).

Las poblaciones humanas existentes en su mayoría no son ideales, se cruzan entre ellas y no son panmícticas. Sin embargo, ante la ausencia total o temporal de fuerzas distorsionadoras, los alelos de las poblaciones siempre avanzan hacia el equilibrio de Hardy-Weinberg. Cuando existen fuerzas que generan una variación en las frecuencias alélicas, se produce un reajuste de las frecuencias alélicas en la población, y se llega a un nuevo equilibrio en las siguientes generaciones. En los alelos que se encuentran en el cromosoma X también ocurre un reajuste hacia el equilibrio, pero debido a su forma de herencia, el camino hacia el equilibrio suele ser más largo (3-4 generaciones) (Hedrick, 2000).

MARCADORES MOLECULARES

Durante la evolución ocurren diferentes mecanismos que conducen a mutaciones en el genoma, esto es, el cambio de bases nitrogenadas por otras o inserciones o deleciones dentro del genoma. Cuando alguna de éstas llega a estar presente en más del 1% de la población, se le denomina polimorfismo genético. Esto implica la existencia de dos o más alelos en el mismo *locus*. Si en un determinado *locus* se pueden encontrar diferentes variantes alélicas dentro de una población, entonces se dice que ese locus es polimórfico (Lewin, 2008).

Un marcador molecular es un sitio polimórfico en un locus del genoma que puede seguirse a través de las generaciones. Los marcadores pueden encontrarse en zonas codificantes o no codificantes. En el caso de encontrarse en una zona

codificante, la variante polimórfica puede generar un cambio sinónimo o no sinónimo. La zona no codificante constituye la mayor parte del genoma (aproximadamente el 95%) (Martinez *et al.*, 2010; Checa Caratachea, 2007; International Human Genome Sequencing Consortium, 2004). Los polimorfismos pueden encontrarse tanto en cromosomas autosómicos como sexuales (cromosomas X e Y) o en el genoma mitocondrial.

Una de las formas de determinar la variabilidad genética dentro y entre las poblaciones es comparando las frecuencias alélicas de diferentes marcadores genéticos. Los tipos de marcadores genéticos que se utilizan en genética de poblaciones son: Polimorfismos de Nucleótido Simple (SNP), Polimorfismos de Inserción-Delección (INDEL), y Polimorfismos de secuencia repetida, entre los que se encuentran los Minisatélites (VNTR) y los Microsatélites (STR).

Polimorfismo de Nucléotido Simple o *Single Nucleotide Polimorphism* (SNP)

Es un tipo de polimorfismo que corresponde a la variación en un solo nucleótido. Esta variación se puede tratar de una sustitución de una base por otra, una delección o una inserción de un solo nucleótido. La mayoría de los SNP son bialélicos, esto significa que existen dos variantes para el locus en cuestión. Son muy frecuentes en el genoma, y generalmente poseen una tasa de mutación baja, por lo que son marcadores útiles para construir mapas genéticos (Aranguren-Méndez *et al.*, 2005). En la Figura 3 se muestra un ejemplo de polimorfismo de nucleótido simple.

Los SNPs se pueden estudiar con técnicas moleculares como PCR alelo específica, que consiste en el diseño de dos *primers Reverse* diferentes, cada uno de los cuales amplifica sólo en presencia de uno de los alelos. También se puede utilizar la técnica de RFLP (*Restriction fragment length polymorphism*) si el SNP forma parte de un sitio de restricción, ya que la enzima de restricción reconocerá y cortará el ADN en una variante alélica, pero no lo hará en la otra variante (Griffiths, 2002). Por último se puede recurrir a la secuenciación del segmento que incluye el SNP.

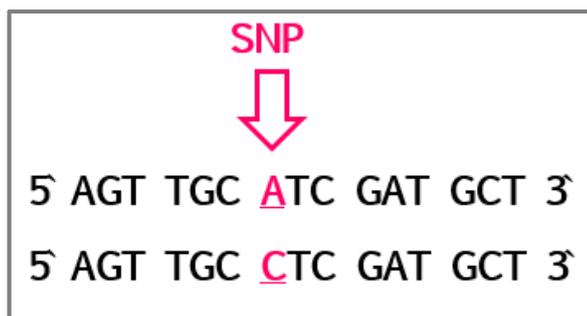


Figura 3. Ejemplo de un polimorfismo de nucleótido simple, en el que hay un alelo A y un alelo C.

Polimorfismos de Inserción-Delección o *Small insertions and deletions* (INDELs)

Son inserciones o deleciones de fragmentos de ADN de uno o más nucleótidos que se encuentran distribuidos a lo largo de todo el genoma. Algunos pueden presentar diferencias grandes de tamaño (hasta 10000 pares de bases) (Mills *et al.*, 2006), aunque por lo general se trata de pocos pares de bases (Weber *et al.*, 2002). En la Figura 4 se muestra un ejemplo de un INDEL.

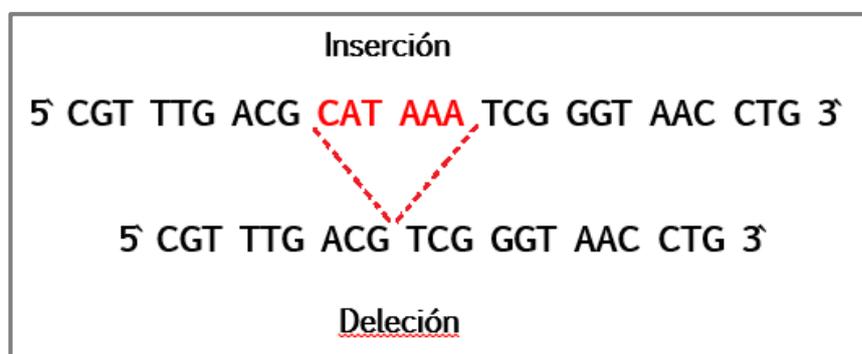


Figura 4. Ejemplo de un polimorfismo de Inserción-Delección. La inserción en este caso es de 6 pares de bases (pb).

Los Indels analizados en este trabajo constituyen inserciones/deleciones de secuencias de ADN pequeñas, desde 4 a 20 pb.

Polimorfismos de secuencia repetida

Existen polimorfismos que comprenden un motivo o secuencia nucleotídica que se repite un número variable de veces (Checa Caratachea, 2007). Dependiendo del tamaño de la repetición, puede tratarse de Minisatélites o *Variable Number of Tandem Repeats* (VNTR), cuyas repeticiones tienen entre 10 y 60 pares de bases, o Microsatélites o *Short Tandem Repeats* (STR) cuyos motivos tienen entre 1 y 7 pares de bases. Por lo general se encuentran en regiones no codificantes (Aranguren-Méndez *et al.*, 2005). En la Figura 5 se representa un ejemplo de este último. Estos tipos de polimorfismo poseen una tasa mutacional mayor a los polimorfismos de secuencia única antes mencionados.

5 CGA TCG CGA CGA CGA CGA CGA AGG TCG 3'

Figura 5. Ejemplo de un STR en el que la secuencia CGA está repetida cinco veces.

CROMOSOMA X

Los humanos poseemos 46 cromosomas, en los que está localizada casi toda nuestra información genética. Veintitrés de ellos provienen del progenitor masculino y 23 de la progenitora femenina. Del total de los cromosomas, 44 son autosómicos, mientras que un par de cromosomas son los sexuales. Existen dos tipos de cromosomas sexuales en los humanos: el cromosoma X y el cromosoma Y.

Los cromosomas X e Y son muy diferentes y pueden distinguirse bien en un cariotipado (observación microscópica de los cromosomas). El cromosoma X tiene un tamaño aproximado de 150 Megabases (Mb), mientras que el cromosoma Y sólo posee 60 Mb (Schaffner, 2004); esta diferencia de tamaño puede apreciarse en la Figura 6. Otra diferencia entre ellos es la tasa mutacional: el cromosoma X posee una tasa mutacional baja, mientras que la del cromosoma Y es alta. En este sentido, los cromosomas sexuales se ubican en los extremos, dejando a los cromosomas autosómicos, que poseen una tasa mutacional moderada, en una posición intermedia. En cuanto a la tasa de recombinación, el cromosoma X posee un valor intermedio,

ya que recombina más que el cromosoma Y (que prácticamente no recombina), pero menos que los autosómicos, ya que el cromosoma X sólo recombina en mujeres (Schaffner, 2004).

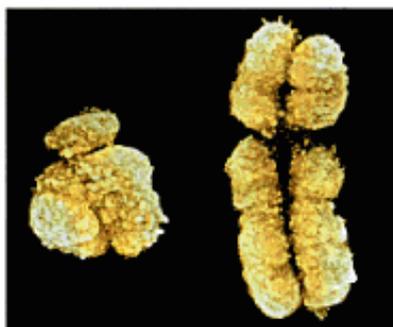


Figura 6. Relación de tamaño y forma de los cromosomas Y (izquierda) y X (derecha) (Pierce, 2010).

Tanto el cromosoma X como el cromosoma Y divergieron de un cromosoma autosómico hace aproximadamente 300 millones de años (Schaffner, 2004); si bien los caminos evolutivos fueron muy diferentes. Por lo tanto, aunque los cromosomas X e Y no son homólogos, aún poseen unas pequeñas zonas de homología, en las que pueden aparearse. Éstas se llaman regiones pseudoautosómicas, y se encuentran en ambas puntas de los cromosomas X e Y, como se puede ver en la Figura 7 (Pierce, 2010).

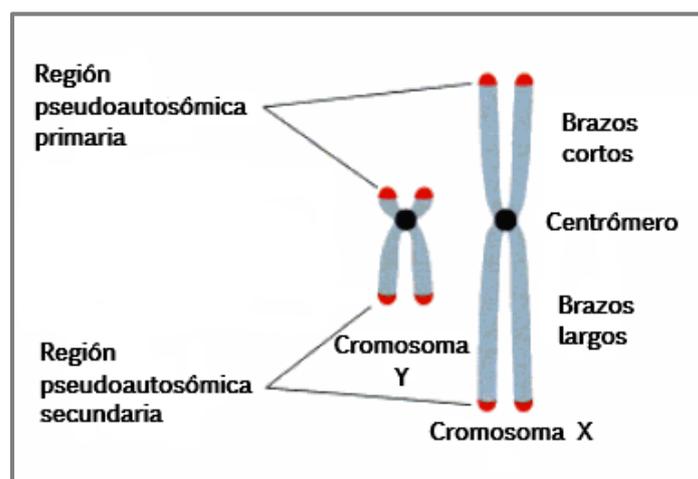


Figura 7. Regiones pseudoautosómicas en los cromosomas X e Y. Los cromosomas X e Y sólo son homólogos en las regiones pseudoautosómicas, que son esenciales para el apareamiento de los cromosomas X e Y en la meiosis de los varones (Pierce, 2010).

Este trabajo se centra en el estudio de la variación del Cromosoma X. Así como todos los cromosomas, el cromosoma X consta de dos brazos, el p (de “*petit*”, pequeño en francés) que es el brazo corto, y el q (“*queue*”, cola) que es el más largo.

La herencia del cromosoma X es única: se comporta como un cromosoma autosómico en el caso de las mujeres, ya que ellas heredan uno de cada progenitor; y como cromosoma sexual y de herencia materna para el caso de los hombres, ya que únicamente poseen un cromosoma X (que se ve acompañado por un cromosoma Y heredado vía paterna). La Figura 8 representa el modo de herencia del cromosoma X. Es así que, en el caso de los hombres, la herencia del cromosoma X genera un haplotipo, al igual que el cromosoma Y y el material genético mitocondrial. En el caso de las mujeres, en cambio, el cromosoma X se presenta en doble copia al igual que los cromosomas autosómicos.

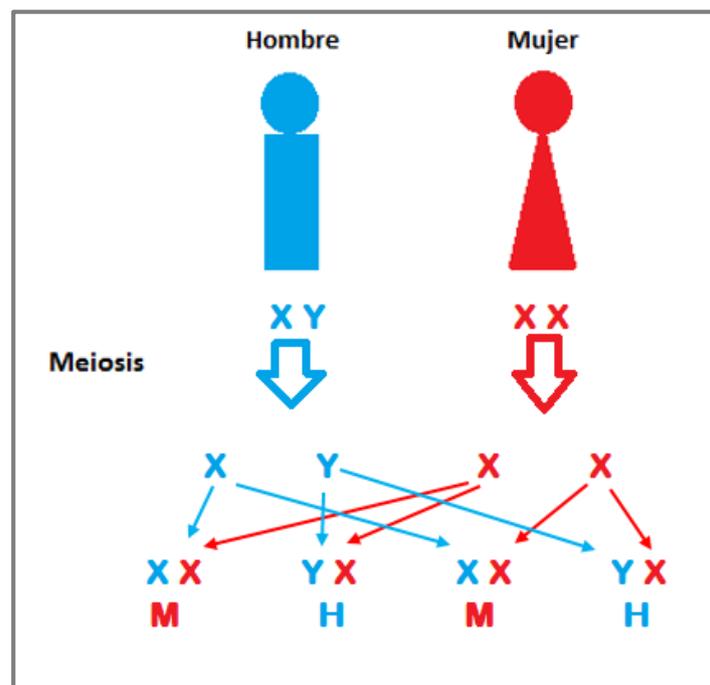


Figura 8. Se muestra la forma de herencia de los cromosomas sexuales (no se muestran los cromosomas autosómicos). La combinación de gametas puede dar lugar a hombres (H) y mujeres (M).

Si bien las mujeres poseen dos cromosomas X, sólo uno de ellos es metabólicamente activo en cada célula somática, mientras que el segundo se encuentra inactivado desde la etapa embriogénica, según la hipótesis de Lyon (Lyon, 1961).

Estas características lo hacen único, y los marcadores moleculares ubicados en el cromosoma X han sido usados extensamente en varias aplicaciones. Para todas ellas, es imprescindible conocer la variabilidad de los marcadores usados en la población en cuestión. Las utilidades que han tenido los marcadores en el cromosoma X son variadas, como las detallan Szibor *et al.* (2007):

- Existen varias enfermedades y trastornos bien conocidos que se encuentran ligados al cromosoma X, como la hemofilia, la distrofia muscular progresiva, la deficiencia de G6PD (glucosa 6-fosfato deshidrogenasa) y deficiencias en la visión de ciertos colores. Éstas son enfermedades genéticamente recesivas, asociadas al cromosoma X y por lo general se presentan en hombres por su condición de hemicigotas (también en mujeres homocigotas para los alelos defectuosos, pero es menos común). Los marcadores en el cromosoma X ligados o directamente relacionados con estas enfermedades entonces tienen una enorme importancia en el área de la genética clínica.
- En el caso de análisis de muestras biológicas, los marcadores del cromosoma X pueden tener una utilidad similar a la de los marcadores ubicados en cromosomas autosómicos, dependiendo del caso. El poder de discriminación (un índice utilizado en ciencia de identificación) de los marcadores de cromosoma X aumenta si se trata de identificar ADN femenino en muestras masculinas (como sería el caso de buscar ADN femenino debajo de las uñas de un sospechoso masculino). En este caso, son más útiles que los marcadores autosómicos.
- Los marcadores en cromosomas sexuales son especialmente eficientes para resolver casos de parentesco difíciles, en los que sólo se cuenta con familiares remotos. Estos casos se encuentran en un contexto post-guerra y migración mundial. Además, en ciertos casos, la identificación de las víctimas de guerras o desastres en masa es más eficiente utilizando marcadores moleculares del

cromosoma X, ya que el poder de exclusión (una manera de evaluar la eficiencia de un marcador en identificación de personas) resulta mejor (más alto).

- En la evaluación de paternidades, los marcadores en el cromosoma X podrían tener una aplicación en los casos padre/hija mujer, mientras que no tienen utilidad alguna en los casos padre/hijo varón. Para análisis de maternidad de hijas mujeres tendrían el mismo peso que los marcadores autosómicos, mientras que sí sería de mayor utilidad en los casos madre/hijo varón.
- Los marcadores en el cromosoma X serían de especial utilidad para determinar si dos mujeres son hermanas, aún si no se dispone del ADN paterno, pero sí el materno. Esto es debido a que las hermanas generalmente heredan haplotipos parcialmente coincidentes de su madre, pero reciben un haplotipo idéntico de su padre.

También se ha propuesto utilizar ciertos alelos de marcadores moleculares como marcadores étnicos o pertenecientes a poblaciones particulares. Glesmann (2015) propuso una combinación de alelos para ciertos marcadores moleculares como distintiva de la población nativa chaqueña, a diferencia de la población criolla de la misma provincia.

Entre los marcadores moleculares que se encuentran en un mismo cromosoma puede existir un ligamiento, y es muy importante analizar su existencia antes de elegir el conjunto de marcadores a utilizar con el propósito que sea (Kaessmann *et al.*, 2002; Szibor *et al.*, 2003; Aranguren-Méndez *et al.*, 2005; Ajmone Marsan P. & Oldenbroek K., 2010). Los alelos ligados no segregan independientemente, por lo que es importante realizar un análisis de desequilibrio de ligamiento entre marcadores, que evalúa si existe una asociación no azarosa de alelos en diferentes loci. Debido a su más alta tasa de mutación, los marcadores del tipo STR suelen mostrar menos ligamiento que los marcadores del tipo SNP (Weber & Wong, 1993; Brinkmann *et al.*, 1998; Szibor *et al.*, 2003).

Se ha determinado que diferentes poblaciones poseen diferente grado de ligamiento: la población africana posee mucho menos ligamiento que las poblaciones

europeas. Esto puede deberse a un cuello de botella o “efecto fundador” ocurrido durante la aparición del hombre europeo a partir del africano (Reich *et al.*, 2001). Es así que la cantidad de ligamiento existente nos puede dar información evolutiva muy valiosa. Además, la cantidad y los patrones de ligamiento entre marcadores del cromosoma X pueden ser eficientes en revelar diferencias étnicas (Kaessmann *et al.*, 2002).

POBLACIÓN ESTUDIADA

En el presente trabajo se analiza la variación genética dentro del cromosoma X en la población de la provincia de Buenos Aires. Para entender y poder interpretar esta variación es importante conocer la historia poblacional de Argentina y más específicamente de la provincia de Buenos Aires.

Población de la República Argentina

La República Argentina se encuentra ubicada geográficamente en el sur del continente americano, y es el segundo país más extenso de América del sur. Limita hacia el oeste y el sur con Chile, hacia el norte con Bolivia y Paraguay, y hacia el este con Brasil, Paraguay, el Río de La Plata y el Mar Argentino.

La población del país es resultado de la mezcla de diferentes etnias y culturas que fue ocurriendo a través de los años, y que permitió alcanzar el número poblacional actual de 40 millones de habitantes (INDEC, Censo 2010).

Inicialmente la población en el territorio argentino estaba compuesta únicamente por indígenas americanos: Mocoví, Pilagá, Toba, Wichí, Warpe, Clachaquí, Quilmes, Atacamas, Araucanos, Tehuelche, Onas, Yamanás, entre otros (Solá, 2009). Esto se dio durante la época precolombina, es decir, en el período previo al proceso de expansión ultramarina europea y la consiguiente conquista y colonización de América. Con la llegada de Colón a América en 1492 se inició una nueva etapa, que implicó la llegada de hombres españoles para poblar América, la fundación de ciudades, y la explotación de las riquezas existentes en el nuevo continente (Floria *et al.*, 1992).

La población indígena en el territorio del Virreinato del Río de La Plata a principios del siglo XVI se estimaba en un millón y medio de habitantes, pero como las regiones más pobladas de éste eran el Alto Perú y Uruguay, la población autóctona del actual territorio argentino puede calcularse que se acercaba a los 300.000 individuos (Floria *et al.*, 1992). Esta escasa población estaba dividida entre los numerosos grupos étnicos, cada cual con características fenotípicas y culturales distintas.

A partir de los primeros asentamientos en lo que es el actual territorio de la ciudad de Buenos Aires, empezaron a arribar españoles, mayoritariamente de sexo masculino. Algunos conquistadores se unieron con hijas de jefes indios, los cuales pasaron a ser sus aliados y parientes. Como consecuencia de la escasez de mujeres hispánicas se originó el mestizaje entre hombres españoles y mujeres aborígenes, y el sistema de concubinato casi poligámico que ocurrió en los primeros años facilitó la rápida propagación del mestizaje (Floria *et al.*, 1992).

Hacia el siglo XVII hizo su aparición la población negra esclava traída desde África, empezando a producirse el mestizaje afro-europeo, que dio lugar al mulato. También empezó a aumentar la población femenina española. El proceso de mestizaje siguió su curso, pero mayormente con características clandestinas, es decir, extramatrimoniales.

En el censo nacional de 1869 la población de la República Argentina constaba de un millón setecientos mil habitantes, de los cuales el 28% vivía en la provincia de Buenos Aires. Además, el 47% de los pobladores de esta provincia eran extranjeros. Otras ciudades con densidad poblacional alta eran Córdoba, Rosario, y, en menor medida, Paraná. A partir de 1850 se inició un lento proceso de inmigración principalmente de italianos, españoles y franceses, a la provincia de Buenos Aires.

La promoción de la agricultura, la ganadería y la red de transportes para la posterior industrialización y crecimiento del país incentivó la inmigración desde todos los lugares del mundo. A partir de 1880, hubo una inmigración promedio mínima de 10.000 personas por año, valor que se incrementó a 64.000 y 112.000 en las siguientes dos décadas (Floria *et al.*, 1992).

Del total de inmigrantes europeos que recibió Argentina en el período 1880-1940, la mayoría fueron italianos (45%) y españoles (33%), pero también llegaron franceses, alemanes, polacos, rusos, turcos, judíos, ucranianos, británicos, portugueses, yugoslavos, griegos, belgas, suecos, e inmigrantes no europeos (estadounidenses, árabes, japoneses), entre otros.

Luego de la segunda Guerra Mundial, alrededor de la década de 1940, la inmigración europea prácticamente cesó, y se vio reemplazada por las migraciones internas. Esta se expresó en un proceso de movilización de la población desde las zonas rurales a las urbanas, proceso incentivado por la industrialización que se estaba llevando a cabo en el país. Posteriormente, y hasta la fecha, la inmigración en el país proviene principalmente de países limítrofes y Perú. Por otra parte, la población nativa que habita nuestro país en la actualidad llega a más de 600.000 individuos (ECPI 2004-2005).

La población actual de Argentina es entonces el resultado de generaciones de mezcla entre varios grupos, a modo de resumen (Avena *et al.*, 2012):

- los indígenas americanos que residían originalmente en esta parte de Sudamérica,
- los conquistadores españoles que empezaron a arribar desde principios del siglo XVI,
- africanos que fueron traídos como esclavos desde fines de ese mismo siglo,
- las grandes inmigraciones europeas que arribaron entre los años 1870 y 1950,
- migraciones desde áreas rurales a ciudades (1930-1980),
- más recientemente, las inmigraciones desde otros países latinoamericanos, especialmente Bolivia, Paraguay y Perú (INDEC, Censo 2010).

Los análisis de ancestría materna (ADN mitocondrial) y paterna (cromosoma Y) evidencian lo acontecido en la historia de Argentina: el ADN mitocondrial posee un alto componente nativo americano mientras que el cromosoma Y muestra un alto porcentaje de componente europeo (Avena *et al.*, 2012).

Avena *et al.* (2012) analizaron los porcentajes ancestrales a partir de estudios de marcadores moleculares en varios puntos del país, obteniendo un promedio de 65% de ancestría europea, 31% indígena americano, y 4% africano en la Argentina. Sin embargo, se ha encontrado que los porcentajes varían mucho entre las diferentes regiones de Argentina: el noroeste (NOA) posee un mayor porcentaje de ancestría amerindia y un menor porcentaje europeo, comparados con Buenos Aires, el sur y el nordeste (NEA) del país.

Población de la Provincia de Buenos Aires

La provincia de Buenos Aires es una de las veintitrés provincias que componen la República Argentina, y su capital es la ciudad de La Plata. Se encuentra ubicada en el centro-este el país, limitando al norte con las provincias de Entre Ríos y Santa Fe, con las provincias de Córdoba, La Pampa y Río Negro al oeste, y al sur y al este con el Mar Argentino.

Es la provincia más poblada de la República Argentina al presente, conteniendo el 38,95% de la población total del país (INDEC, Censo 2010).

Por ser la provincia que contiene a la ciudad autónoma de Buenos Aires, y por su numerosa población, refleja una parte importante de la historia étnica del país.

En la época precolombina estuvo habitada por los Pampas, un pueblo cazador (Floria *et al.*, 1992). Más adelante, la zona que comprende actualmente al Gran Buenos Aires ¹ (la cual tenía importancia por su puerto) fue elegida como capital del Virreinato del Río de La Plata.

Al ser Buenos Aires la provincia en la que se encuentra la capital, fue la que más recibió las migraciones externas e internas, y así mismo la que más creció en población hasta llegar al día de hoy, con quince millones seiscientos mil personas (INDEC, Censo 2010).

¹ El Gran Buenos Aires es el área metropolitana conformada por la Ciudad de Buenos Aires y los partidos circundantes, en la cual se registra una alta concentración poblacional.

En cuanto al Gran Buenos Aires, es la tercera metrópolis más habitada en América Latina, luego de México y San Pablo. La composición genética de sus habitantes está fuertemente influenciada por dos eventos históricos que acontecieron (Avena *et al.*, 2012):

- la inmigración europea, principalmente italianos y españoles, entre los años 1870 y 1950, que se mezcló con la población ya mestizada proveniente de la colonización.
- Una segunda ola de inmigración en los años 1940s, cuando el desarrollo industrial en la ciudad atrajo población de otras provincias y países limítrofes. Estos migrantes poseían un componente genético amerindio mayor, contribuyendo a una mayor mezcla en la población ya mestiza de Buenos Aires.

En estudios de ancestría se ha encontrado un valor de 76% de ancestría europea para la provincia de Buenos Aires, la cual difiere mucho del porcentaje obtenido para el Noroeste del país: 33% (Avena *et al.*, 2012). Sin embargo, el 1,97% de la población actual de Buenos Aires está conformada por nativos y descendientes de nativos, que viven en su mayor parte (93%) en zonas urbanas, no en comunidades aisladas (INDEC, Censo 2010).

En este contexto, ante una población tan mixta en cuanto a su población, es importante conocer la variación de cada compartimiento genómico en particular. El cromosoma X constituye una parte del genoma con un modo de herencia único, que aún no ha sido analizado en profundidad en la población bonaerense.

En este trabajo se eligieron 15 marcadores moleculares ubicados en una región del brazo largo del cromosoma X que presenta menor densidad de genes codificantes en comparación con otras regiones del X. De esta manera, se estima que dicha región tiene mayor posibilidad de generar variabilidad genética, y ofrecerá la posibilidad de hallar cambios de secuencia. Se incluyeron en este estudio 10 polimorfismos del tipo SNP y 5 del tipo Inserción-delección. Todos ellos se encuentran en regiones no codificantes y de herencia específica del cromosoma X (no

pseudoautosómicos), y han sido tipificados en otras poblaciones de la Argentina (Glesmann, 2015; Birdi *et al.*, 2013), datos que aportan en el conocimiento de la variación de este cromosoma en la población estudiada.

OBJETIVOS

El **objetivo general** del proyecto es caracterizar la población bonaerense en cuanto a su diversidad genética del cromosoma X.

Objetivos particulares:

- Conocer la variación de 5 polimorfismos del tipo Indel y 10 del tipo SNP específicos de cromosoma X, en la población de la provincia de Buenos Aires.
- Evaluar las similitudes y diferencias de dicha variación con datos de otras provincias argentinas y con bibliografía científica sobre otras poblaciones del mundo.

HIPÓTESIS

Los hechos histórico-demográficos que generaron y moldearon la población argentina actual pueden evidenciarse en la variación de marcadores de cromosoma X. La variación hallada en los marcadores del Cromosoma X puede reflejar un componente europeo de nuestra población, evidenciando un importante proceso de flujo genético desde esa procedencia ocurrido sobre la población de la provincia de Buenos Aires.

También podría evidenciarse, a partir de los marcadores de Cromosoma X, un componente nativo americano, en relación con el origen mixto de la población argentina en general.

MATERIALES Y MÉTODOS

POBLACIÓN ANALIZADA

Se tomaron muestras de personas habitantes de diversos partidos de la provincia de Buenos Aires; los detalles se muestran en la Tabla 2. Las edades de los donantes varían entre 18 y 50 años y se alcanzó un número muestral total de 116 personas (45 hombres y 71 mujeres). Previo a la toma de muestras de saliva a los donantes se les entregó una carta informativa sobre el proyecto, y se respondieron inquietudes. A continuación los participantes completaron y firmaron un consentimiento informado. En el Apéndice 1 se encuentran la carta informativa y el consentimiento. Se solicitó a los participantes que se realizaran un buche bucal con 2 ml de agua mineral, volcándolo luego en un tubo plástico tipo falcon. Cada muestra se tomó por duplicado, luego se le agregaron dos volúmenes de etanol 96% para inhibir actividad bacteriana, se mezcló el contenido y se guardó en freezer hasta el día de la extracción de DNA.

Tabla 2. Muestras tomadas en cada partido perteneciente a la provincia de Buenos Aires.

Partido	Número muestral tomado
La Plata	91
Berisso	3
Ensenada	2
Azul	13
Buenos Aires	4
Coronel Brandsen	2
Tandil	1
Total	116

EXTRACCIÓN DE ADN

La extracción de ADN a partir de las muestras de saliva se realizó según el protocolo de la técnica con cloruro de Litio (LiCl). Los pasos a seguir fueron los siguientes:

1. Una vez descongelados los tubos preservados en freezer, se equilibraron con etanol 96% y se centrifugaron a 4000 rpm durante 10 minutos a una temperatura de 7°C.
2. Se descartó el sobrenadante y se agregó a cada tubo 150 µl de buffer TE (Tris-EDTA) para lavar el pellet. Luego se centrifugó nuevamente a 4000 rpm durante 10 minutos a 7°C.
3. Se descartó el sobrenadante y se agregaron 300µl de Buffer de digestión y 10 µl de proteinasa K. Se incubó hasta el segundo día en baño térmico a 37°C (16 horas).
4. Se agregaron 30 µl de LiCl 5M y se agitó 1 minuto.
5. Se agregaron 600 µl de SEVAG (alcohol isoamílico y cloroformo, 1:24) y se dejó en agitador 30 minutos.
6. Se centrifugó durante 20 minutos a 4000 rpm, a una temperatura de 7°C. Se retiraron los tubos de la centrífuga con precaución para no mezclar las fases. Con ayuda de pipetas Pasteur se pasó la fase superior a otro tubo, sin tocar la proteína que quedó entre las dos fases.
7. Se agregaron a la fase acuosa 2 volúmenes de etanol absoluto frío y se mezcló suavemente buscando la formación de un ovillo.
8. Cuando se formó el ovillo, éste se pasó a un tubo tipo eppendorf con pipeta automática de 1 ml y se lavó con 50 µl de etanol 70 (700 µl de etanol absoluto frío y 30 µl de agua Milli-Q). En los casos en los que no se formó el ovillo, se centrifugó la totalidad del volumen presente en el tubo (cuando fue necesario, se centrifugó más de una vez).
9. Se centrifugaron los tubos eppendorf durante 10 minutos a 13000 rpm.
10. Se descartó el sobrenadante y se lavó el pellet con 50 µl de etanol 70, Se centrifugó nuevamente 10 minutos a 13000 rpm. Se descartó el etanol y luego

de la evaporación absoluta del mismo, se resuspendió el pellet en 100-150 μ l de TE (Tris-EDTA pH 8), con ayuda de un vórtex.

Las muestras se cuantificaron con el equipo Nanodrop 2000 (Thermo Fisher Scientific Inc.).

MARCADORES ANALIZADOS

Polimorfismos de nucleótido simple

Se analizaron 10 marcadores de tipo SNP. Todos fueron amplificados mediante la técnica de reacción en cadena de la polimerasa (PCR). Ésta consiste en la amplificación de un fragmento deseado de ADN, gracias a la incorporación de cebadores oligonucleotídicos o *primers Forward* y *Reverse* diseñados artificialmente para hibridar en una zona específica del genoma. La enzima Taq Polimerasa se agrega para polimerizar las hebras de ADN a partir de los moldes, utilizando los cebadores y dNTPs. Se realizan sucesivos ciclos de desnaturalización, hibridación o *annealing* y extensión. La desnaturalización consiste en aumentar la temperatura de la mezcla a unos 94°C para que las hebras de ADN se separen. Al ser la Taq Polimerasa termorresistente, conserva su actividad a pesar de la temperatura alta. Durante el *annealing* se baja la temperatura durante unos segundos para permitir que los cebadores hibriden, delimitando la zona de interés del genoma. La extensión de la cadena de ADN dura un minuto aproximadamente a la temperatura de mejor actividad de la polimerasa, de modo que es cuando se sintetiza el ADN a partir del molde y los cebadores. Como en cada ciclo se genera el doble de la cantidad presente de ADN molde, la producción de éste es exponencial durante los sucesivos ciclos, y luego de 34 ciclos en el tubo hay una gran cantidad del amplicón de interés (Garbus *et al.*, 2010). La Figura 9 representa los tres primeros ciclos de la PCR, y cómo aumenta el número de copias de la región de ADN deseada.

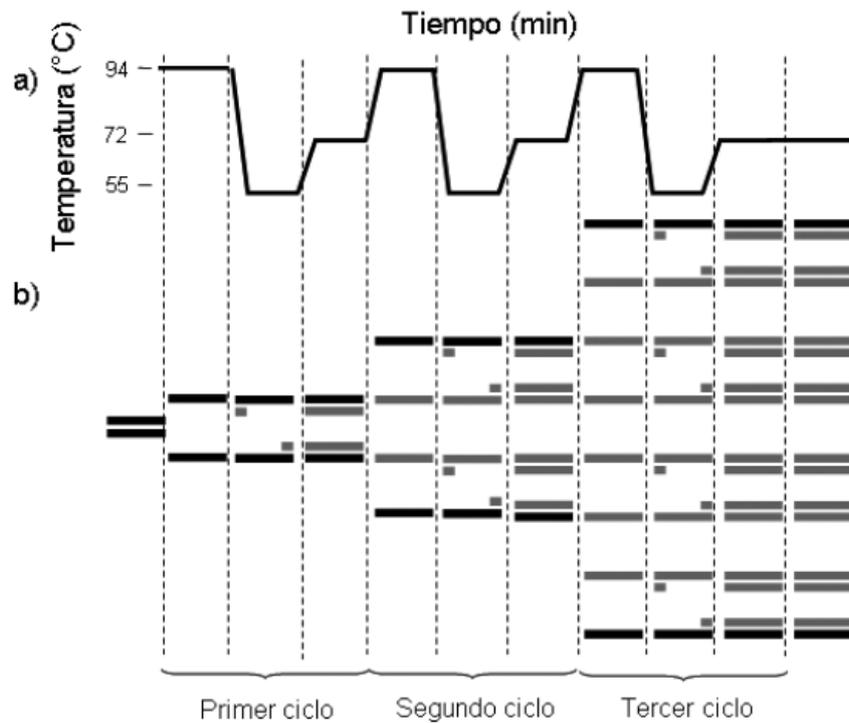


Figura 9. Representación de una reacción de PCR de tres ciclos. A) Se muestra un gráfico de la temperatura en función del tiempo. En cada ciclo se suceden tres etapas: desnaturalización (94°C), hibridación (aprox 55°C, dependiendo de los *primers* o cebadores) y polimerización (72°C). Se muestra gráfico de temperatura en función del tiempo, en minutos. B) Comportamiento de un fragmento de ADN blanco, a través de los ciclos. En cada ciclo la cantidad de ADN blanco se duplica, de manera tal que al cabo de n ciclos el número de copias es 2^n (Garbus *et al.*, 2010).

Con respecto a los polimorfismos de tipo SNP, en la Figura 10 se muestra las localizaciones de los SNPs analizados en el cromosoma X.

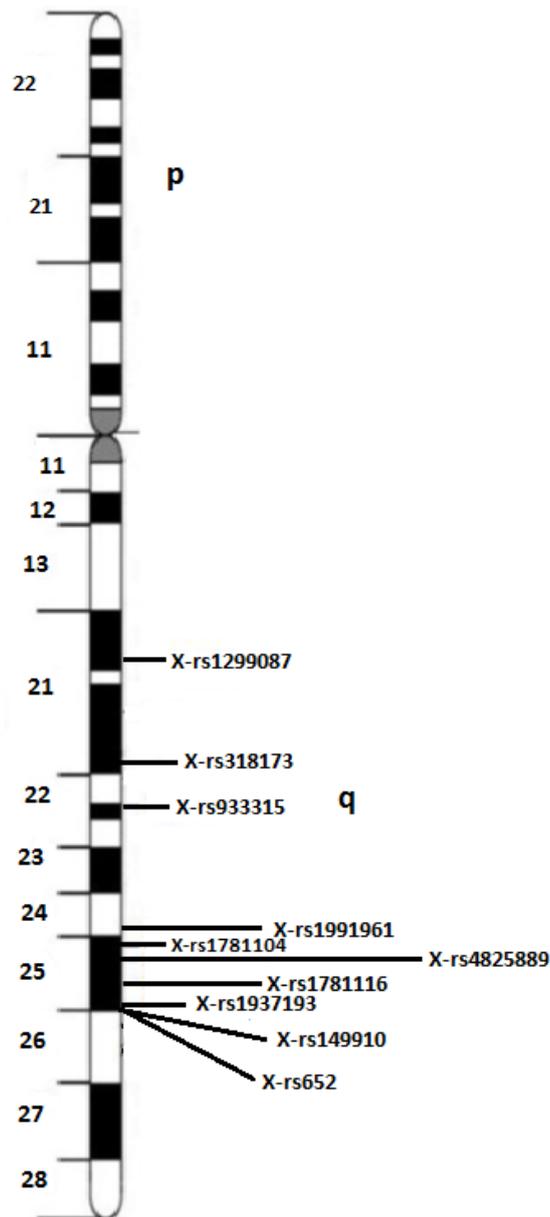


Figura 10. Ubicación citogenética de los X-SNPs analizados. (Glesmann, 2015)

Algunos pares de SNPs se coamplificaron en reacciones de dúplex, lo que permitió leer dos resultados a partir de una misma reacción de PCR. Para ello, se juntaron pares de marcadores con temperaturas de *annealing* similares y se ensayaron diferentes combinaciones de concentraciones de *primers* hasta encontrar las adecuadas.

Para algunos de los marcadores o dúplex de los marcadores usados, se utilizó un *Touch Down*. Este procedimiento consiste en la realización de los primeros ciclos con una temperatura de *annealing* mayor, aumentando la especificidad de la

amplificación. Luego, se va disminuyendo la temperatura de *annealing* en los sucesivos ciclos.

Del total de los marcadores analizados, nueve de ellos fueron amplificados con la técnica de PCR alelo específica. En ésta, se realizan dos reacciones de PCR por individuo. En ambas se utiliza el mismo primer *Reverse*, pero el primer *Forward* para cada reacción tiene la última base complementaria al alelo en cuestión, convirtiéndolo en específico de alelo.

Todas las reacciones de PCR se realizaron con un volumen final por reacción de 10µl, y las siguientes concentraciones de reactivos:

- 1X de buffer de PCR (50 mM de Tris-HCl pH 8, 100 mM de ClNa, 0,1 mM de EDTA, 1 mM de DTT, 50% de glicerol y 1% de Tritón®X-100),
- 2mM de MgCl₂,
- 0,2mM de dNTPs,
- 0,025 U de Taq ADN Polimerasa (INBIO Highway, Argentina),
- agua Milli-Q para completar los 10µl finales,
- Las concentraciones de *primers* (sintetizados por la empresa Invitrogen, Thermo Fisher Scientific Inc.), y ADN que se indican en la Tabla 3.

Las secuencias de los *primers* utilizados fueron diseñadas en nuestro laboratorio con el programa Primer 3 y se encuentran en el Apéndice 2.

Los ciclados se realizaron en cicladores MPI (La Plata, provincia de Buenos Aires, Argentina) y se especifican en detalle en la Tabla 4.

Las características de los amplicones y los SNPs se muestran en la Tabla 5.

Tabla 3. Concentraciones de *primers* (en picomoles por reacción) y ADN (en nanogramos por reacción) de las reacciones de PCR realizadas.

	Dúplex rs933315(a)- rs1991961(b)	Dúplex rs318173(a)- rs4825889(b)	Dúplex rs1937193(a)- rs149910(b)	rs1781116	rs652	rs1781104
Primer (pmol/r)	(a): 2,5 (b): 3,0	(a): 2,0 (b): 4,5	(a): 2,0 (b): 2,5	4,5	1,4	3,5
ADN (nmol/r)	20	20	20	20	15	20

Tabla 4. Características de los ciclados de las reacciones de PCR. A modo de ejemplo, un *Touch Down* de 1°C-60°C-2 ciclos significa que se empezó con una temperatura de *annealing* definida según el ciclado, y se fue disminuyendo la temperatura en 1°C cada dos ciclos, hasta la temperatura de 60°C. Luego, la temperatura de *annealing* se mantuvo en 60°C el resto de los ciclos.

	Dúplex rs933315- rs1991961		Dúplex rs318173- rs4825889		Dúplex rs1937193- rs149910		rs1781116		rs652		rs1781104	
	Temp (°C)	Tiempo (s)	Temp (°C)	Tiempo (s)	Temp (°C)	Tiempo (s)	Temp (°C)	Tiempo (s)	Temp (°C)	Tiempo (s)	Temp (°C)	Tiempo (s)
Desnaturalización inicial	94	120	94	240	94	120	94	120	94	120	94	120
Desnaturalización	93	45	94	40	93	45	93	45	93	45	93	45
<i>Annealing</i>	64	60	64	55	66	60	66	60	66	60	64	60
Extensión	72	60	72	60	72	60	72	60	72	60	72	60
Extensión final	72	300	72	420	72	300	72	300	72	300	72	300
Nº de ciclos	34		34		34		34		34		34	
<i>Touch down</i>	1°C- 60°C-2 ciclos		1°C- 60°C-2 ciclos		1°C- 62°C-2 ciclos		1°C- 62°C- 2 ciclos		1°C- 62°C- 2 ciclos		1°C- 60°C- 2 ciclos	

Referencias: Temp= temperatura.

Tabla 5. Característica de los amplicones y los SNPs analizados.

SNP	rs318173	rs4825889	rs933315	rs1991961	rs149910	rs1937193	rs1781116	rs652	rs1781104	rs1299087
Alelos	G-T	A-G	A-G	G-T	G-T	A-T	C-T	G-T	A-C	A-G
Alelo ancestral	T	A	A	T	G	A	T	T	C	G
Ubicación cromosómica	Xq21.3	Xq25	Xq22.3	Xq24	Xq25	Xq25	Xq25	Xq25	Xq25	Xq21.1
Tamaño del amplicón (pb)	101	255	175	150	125	145	106	180	141	156

Referencias: pb= pares de bases

Los resultados de las PCR alelo específicas se determinaron mediante electroforesis en geles de agarosa al 1,6%. En ésta, se separan fragmentos de ADN por tamaño, aplicando un voltaje entre los terminales positivos y negativos de una cuba. Se utilizó como buffer de corrida TAE 1% (Tris-Base 40mM; ácido acético; EDTA 1mM, pH 8). Cada muestra se sembró con 1µl de una solución de 2,5 ml de buffer de carga (0,25% de xilen cianol, 0,25% de azul de bromofenol y 30% de sacarosa) con 1 µl de GelRed™ (Biotium, USA) para visualizar el ADN. El voltaje se aplicó con fuentes BIO-RAD (Power Pac™, USA) y fue de 85-90 V durante 1 hora aproximadamente, utilizando una cuba horizontal MPI (La Plata, Argentina). El gel se reveló a la luz UV en un analizador de imágenes GelDocXR (BIO-RAD, USA).

Debido a que se utilizó PCR alelo específica, por cada muestra se sembraron dos calles en el gel. La presencia de cierta banda observada en agarosa indica la presencia del alelo correspondiente en el individuo. Se observa la presencia de una de las dos bandas si se trata de una mujer homocigota o un hombre hemicigota, y se observan ambas bandas en el caso de una mujer heterocigota. Se corrió un control de tamaño como referencia en cada corrida.

Para resolver los dúplex rs933315-rs1991961 y rs1937193-rs149910 se armaron geles de agarosa 1,6 % con 1,5 µl de GelRed™ incorporado en la agarosa fundida, y se sembraron 3µl de amplicón con 2 µl de buffer de carga. Se corrió el gel a 75 V durante 1 hora y media, y se reveló con luz UV en el analizador. Los resultados dudosos se corrieron en acrilamida, que posee un mayor poder de resolución.

El marcador rs1299087 fue tipificado por la técnica PCR-RFLP. En este método, uno de los alelos del SNP genera un sitio de restricción enzimático, mientras que si está el otro alelo no se presenta dicho sitio. Por lo tanto, podemos determinar la identidad del alelo sometiendo el producto de PCR a la digestión con la enzima correspondiente. En este caso, el alelo G genera el sitio de corte para la enzima MspI (HpaII). El sitio de restricción se muestra en la Figura 11.

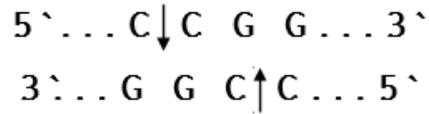


Figura 11. Sitio de corte de la enzima de restricción MspI.

La amplificación se realizó mediante una PCR estándar, con una concentración de primer de 2,7 pmol por reacción, y 20 ng de DNA (volumen final 10 μ l).

El ciclado se muestra en la Tabla 6.

Tabla 6. Ciclado de la PCR del SNP rs1299087.

	Temp (°C)	Tiempo (s)
Desnaturalización inicial	94	120
Desnaturalización	93	45
<i>Annealing</i>	54	60
Extensión	72	60
Extensión final	72	300
Nº de ciclos	34	

Referencias: Temp= temperatura.

El amplicón generado es de un tamaño de 156 pb. Para la digestión con la enzima de restricción MspI (Thermo Fisher Scientific Inc.) se realizaron reacciones de un volumen final de 16 μ l, utilizando el buffer Tango (Thermo Fisher Scientific Inc.), 4,8 unidades enzimáticas por reacción y 6 μ l del amplicón. Se incubó a 37°C *overnight* (16 horas aproximadamente).

Se visualizó el resultado en gel de acrilamida al 8% neutro. Éste se preparó mezclando el instagel 8% (0,32g/l de bisacrilamida, 0,02 g/l de acrilamida ultrapura, preparado con agua Milli-Q) con un volumen de Temed y dos volúmenes de Persulfato de Amonio al 40%. Se utilizó como buffer de corrida TBE (Tris-Ácido bórico 45mM-

EDTA 1mM) 0,5X. Se agregaron 2µl de buffer de carga a cada tubo de digestión y se sembraron 6µl. Se utilizó como control de corrida una escalera de 100pb (INBIO Highway, Argentina). Se corrió a 250V y 20mA durante una hora y media. Luego se realizó la tinción con una solución de Gel Red (5µl de Gel Red en 80 ml de buffer TAE) durante 10 minutos y se reveló a la luz UV.

Luego de la digestión, pueden obtenerse fragmentos de 107 pb y 49 pb si la enzima encontró el sitio de restricción (alelo G). Los patrones de digestión nos permiten determinar si el individuo es homocigota para uno de los alelos (se ven o bien bandas de 107 y 49 pb, o bien una banda de 156) o heterocigota (sólo posible en el caso de mujeres, se observan las tres bandas).

Polimorfismos de Inserción-Delección

Se analizaron cinco marcadores del tipo Indel. Estos son fragmentos de diversos tamaños que pueden o no estar presentes en el cromosoma X. Cuando se amplifica por PCR la zona, se obtienen entonces amplicones de dos diferentes tamaños dependiendo de la presencia o ausencia del fragmento.

Se realizaron reacciones de PCR estándar, utilizando 15 ng de ADN por reacción y 2,0 pmol de *primers* por reacción (10µl finales). Los indels MID 1540 y MID 3754 se realizaron en dúplex, y se utilizaron concentraciones de *primers* de 3,5 pmol y 2,0 pmol por reacción, respectivamente. Además, en la reacción de dúplex se utilizaron 20 ng de ADN por reacción.

Los ciclados se realizaron en termocicladores MPI (La Plata, provincia de Buenos Aires, Argentina), y se detallan en la Tabla 7.

Tabla 7. Condiciones de los ciclados utilizados en las reacciones de PCR de los Indels analizados.

	MID 1705		MID 3756		MID 193		Dúplex MID 1540- MID 3754	
	Temp (°C)	Tiempo (s)	Temp (°C)	Tiempo (s)	Temp (°C)	Tiempo (s)	Temp (°C)	Tiempo (s)
Desnaturalización inicial	94	120	94	120	93	300	94	120
Desnaturalización	93	40	93	35	94	30	93	40
<i>Annealing</i>	59	55	60	60	56	60	59	55
Extensión	72	60	72	60	72	60	72	60
Extensión final	72	600	72	600	72	300	72	600
Nº de ciclos	35		38		34		35	
<i>Touch down</i>	1°C-54°C-1 ciclos		1°C-56°C-2 ciclos		1°C-48°C-2 ciclos		1°C-54°C-2 ciclos	

Referencias: Temp= temperatura.

Las secuencias de los *primers* utilizados se muestran en el Apéndice 3, así como también los alelos y los números de identificación de los X-Indels analizados.

Se verificaron los amplicones en agarosa 1,6% y luego todas las reacciones positivas se resolvieron en gel de acrilamida 8%. Éste tiene una mayor resolución que los geles de agarosa, permitiendo visualizar las inserciones y deleciones como bandas independientes.

Se corrieron volúmenes de entre 0,7 y 1,0 μ l de amplicón, dependiendo de la intensidad de la banda observada en agarosa. A cada tubo se le agregaron 2 μ l de buffer de carga. Se corrió una escalera de 100 pb (INBIO Highway, Argentina) como control de tamaño. Los geles se tiñeron con una solución de Gel Red (5 μ l de Gel Red en 80 ml de TAE 1X) durante 10 minutos y se reveló el gel con luz UV en el equipo analizador de imágenes GelDocXR (BIO-RAD, USA).

En la Tabla 8 se muestran los tamaños de los amplicones y las características de los X-Indels analizados.

Tabla 8. Características de los X-Indels analizados y tamaños de los amplicones.

Indel	MID 1705	MID 3756	MID 3754	MID 1540	MID 193
Ubicación cromosómica	Xq25	Xq25	Xq24	Xq28	Xq27
Tamaño de la inserción (pb)	4	13	12	20	4
Tamaño amplicones (pb)	133-137	111-124	129-141	231-251	101-105

Referencias: pb= pares de bases

En la Figura 12 se muestra la localización cromosómica de los mismos.

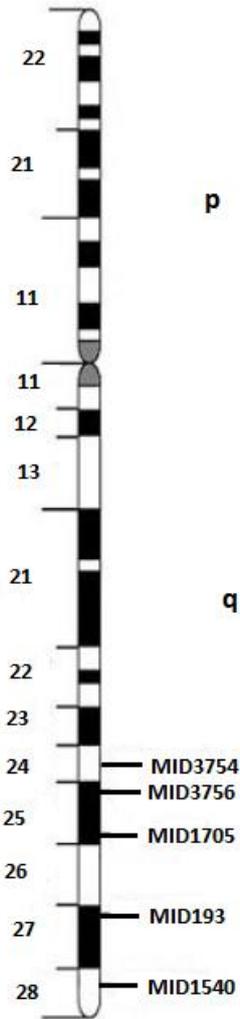


Figura 12. Ubicación citogenética de los X-Indels analizados (Glesmann, 2015).

CÁLCULOS ESTADÍSTICOS

Se calcularon los valores de frecuencias alélicas, genotípicas, y haplotípicas. Los valores de frecuencias con sus desvíos estándar se graficaron con el programa Prism version 6.07 for Windows, GraphPad Software, La Jolla California USA, (www.graphpad.com).

Frecuencias alélicas

La frecuencia alélica de un alelo es la cantidad de veces que se encuentra en la población analizada, respecto del total de *loci*. Para calcular las frecuencias alélicas se contaron los alelos de cada locus y se utilizó la fórmula:

$$P_i = \frac{n_i}{T} ,$$

donde P_i es la frecuencia del alelo i , n_i es el número de alelos i observados en la población, y T es la cantidad total de alelos analizados.

El desvío estándar (σ) de las frecuencias alélicas es una medida de la desviación típica de la distribución de muestreo. Se calcula mediante la siguiente expresión:

$$\sigma = \sqrt{\frac{p_i(1-p_i)}{2N}} ,$$

donde N es el número de individuos de la población y P_i es la frecuencia del alelo i .

Frecuencias genotípicas

Se define frecuencia genotípica como la proporción en la que se encuentra un genotipo en la población. En el caso de marcadores del cromosoma X, las frecuencias genotípicas sólo pueden ser calculadas para las mujeres, ya que los hombres son hemicigotas. Se calcularon las frecuencias genotípicas para cada uno de los marcadores, con la siguiente fórmula:

$$P_{ij} = \frac{n_{ij}}{N} ,$$

donde n_{ij} es el genotipo con los alelos i y j , y N es el número de individuos diploides.

También se calculó el desvío estándar como:

$$\sigma = \sqrt{\frac{p_{ij}(1-p_{ij})}{N}} ,$$

donde P_{ij} es la frecuencia del genotipo ij , y N el número total de genotipos.

Frecuencias haplotípicas

Se corrieron los datos obtenidos de muestras masculinas en el programa Arlequín 3.5.2.2 (Excoffier *et al.*, 2010) para calcular el número de haplotipos presente. Las frecuencias haplotípicas se calcularon con la fórmula:

$$P_i = \frac{n_i}{N} ,$$

donde n_i es la cantidad de veces que se observa el haplotipo i en el total de los N hombres analizados.

Análisis de desequilibrio de ligamiento

En general, existe una asociación entre la distancia física entre dos zonas del cromosoma, el número de nucleótidos, y la cantidad de recombinación encontrada entre ellos. Sin embargo, dentro de un mismo cromosoma existen zonas de mayor tasa de recombinación (llamadas “zonas calientes”) y otras de menor tasa (Hedrick, 2000). Dos alelos muy cercanos tienden a segregar juntos durante la división meiótica si no existe demasiada recombinación entre ellos. Así, el ligamiento entre dos alelos depende de la distancia a la que se encuentran y de la tasa de recombinación en esa zona. Si dos alelos de dos marcadores tienen alta probabilidad de segregar juntos se dice que los marcadores están ligados.

El análisis de desequilibrio de ligamiento se realizó también con el programa Arlequín 3.5.2.2, y se realizó entre Indels, entre SNPs, y entre ambos tipos de marcadores.

Heterocigosis observada

Se utilizó la siguiente fórmula para calcular la proporción de heterocigosis en la población:

$$H = \frac{nh}{N} ,$$

donde n_h es el número de mujeres heterocigotas y N el número total de mujeres analizadas.

Heterocigosis media

La heterocigosis media es el promedio de las heterocigosis observadas para cada uno de los marcadores analizados, y nos da una idea de la variabilidad genotípica de la población. Se calcula como:

$$H = \frac{\sum h_i}{k} ,$$

donde h_i es la heterocigosis para cada marcador en la población, y k es el número total de marcadores analizados.

Heterocigosis esperada y equilibrio de Hardy-Weinberg

En una población que está en equilibrio de Hardy-Weinberg, las frecuencias alélicas se mantienen constantes de una generación a otra, y nos permiten predecir las frecuencias genotípicas.

El ajuste al equilibrio de Hardy-Weinberg se calculó utilizando los genotipos femeninos en el programa Arlequín 3.5.2.2. La heterocigosis esperada según el equilibrio de Hardy-Weinberg se calcula como la unidad menos la sumatoria de los cuadrados de las frecuencias alélicas:

$$H_e = 1 - \sum p_i^2 ,$$

donde H_e es la heterocigosis esperada y P_i son las frecuencias alélicas.

El programa realiza un test de significancia comparando las heterocigosis observadas y esperadas. Si el marcador está en equilibrio, las heterocigosis deben ser similares. En cambio en un marcador que no está en equilibrio, ambos valores de heterocigosis son significativamente distintos. El valor de probabilidad p nos indica si las diferencias observadas entre ambas heterocigosis son debidas al azar. Un valor $p < 0,05$

indica que las diferencias observadas entre ambas heterocigosis son significativas, y que no se deben al azar.

Diversidad génica

La diversidad génica es el equivalente a la heterocigosis esperada en hombres, los cuales poseen sólo un cromosoma X. Está definido como la probabilidad de que dos haplotipos tomados al azar en la población sean diferentes. El programa Arlequín estima la diversidad génica, su varianza y su desviación estándar.

$$\hat{H} = \frac{n}{n-1} \left(1 - \sum_{i=1}^k p_i^2\right)$$

donde H es la diversidad génica, n el número de copias génicas en la población muestral, k el número de haplotipos, y p_i la frecuencia del haplotipo de orden i.

Variabilidad y estructura poblacional

La variabilidad génica poblacional posee dos componentes:

- La variabilidad intrapoblacional, que es la que existe entre los individuos de la misma población, se calcula sumando todos los valores de heterocigosis observada para los marcadores analizados.
- La variabilidad interpoblacional, la cual mide la diferencia genética entre poblaciones. Por lo general es menor que la variabilidad que existe dentro de la misma población.

La variabilidad total observada es, entonces, la suma de ambas variabilidades. Las estimaciones de variabilidad y estructura poblacional se realizaron utilizando los programas Arlequín 3.5.2.2 (Excoffier *et al.*, 2010) y Structure 2.3.4 (Pritchard *et al.*, 2000).

Índice FST de Wright

Este índice mide qué tan diferentes genéticamente son las poblaciones, comparando los valores de heterocigosis de las poblaciones tomadas en conjunto con la heterocigosis de cada población particular, de acuerdo con la fórmula:

$$F_{st} = \frac{\sum P_i (1 - P_i) \cdot F_{sti}}{\sum P_i (1 - P_i)}$$

Un valor de F_{st} cercano a 0 indica que las frecuencias alélicas en las poblaciones que se están comparando son muy similares. Cuando el F_{st} toma valores mayores indica que las poblaciones están más alejadas genéticamente porque sus distribuciones de frecuencias son distintas.

En este trabajo se utilizó el programa Arlequín para comparar los datos obtenidos con datos previos de la población de la provincia de Chaco (Glesmann, 2015) incluyendo comunidades aborígenes Wichí y Mocoví, Resistencia y Misión Nueva Pompeya. Además, para el estudio de los X-Indels de forma individual, también se realizó una comparación con datos de la ciudad de Viedma (Bridi *et al.*, 2013). Se utilizaron sólo estas poblaciones porque son las únicas en Argentina tipificadas para estos marcadores y presentes en la bibliografía.

Para algunos de los marcadores individuales se realizó una comparación con datos de frecuencias alélicas de otras partes del mundo, obtenidos en el sitio web 1000 Genomes (<http://www.1000genomes.org>).

Análisis de varianza molecular (AMOVA)

El análisis de varianza molecular es un método para detectar diferencias entre poblaciones utilizando marcadores moleculares. Se llevó a cabo entre grupos de poblaciones armados *a priori* para los X-SNPs, los X-Indels, y para todos los marcadores y es una estimación de la variación que existe dentro de una población, entre las poblaciones agrupadas, y entre los grupos de poblaciones armados. El cálculo de AMOVA se realizó con el programa Arlequín 3.5.2.2.

Análisis de la estructura poblacional

El análisis de estructura poblacional se realizó con el programa Structure 2.3.4 (Pritchard *et al.*, 2000). Se corrieron los datos de las poblaciones de Buenos Aires, Resistencia, Misión Nueva Pompeya, y comunidades Wichi y Mocoví, bajo el supuesto de ancestría mixta y tres posibles orígenes ancestrales ($k=3$).

RESULTADOS

VARIACIÓN DENTRO DE LA POBLACIÓN

Polimorfismos de Nucleótido Simple

Los amplicones de las reacciones de PCR para los X-SNPs se visualizaron en agarosa al 1,6%. En las Figuras 13, 14 y 15 se observan algunas de las visualizaciones obtenidas. Para el caso de las PCR alelo específicas se ve una amplificación de ambos alelos sólo en el caso de mujeres heterocigotas, y se observa una sola banda en el caso de hombres y mujeres homocigotas.

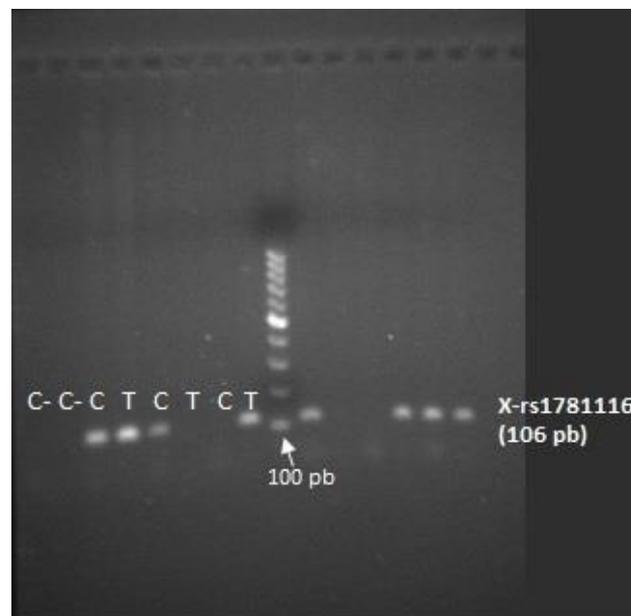


Figura 13. Electroforesis en gel de agarosa 1,6 % de una PCR alelo específica del marcador X-rs1781116. En el medio, una escalera de 100 pares de bases (pb). C- =control negativo, C=amplificación del alelo C, T=amplificación del alelo T.

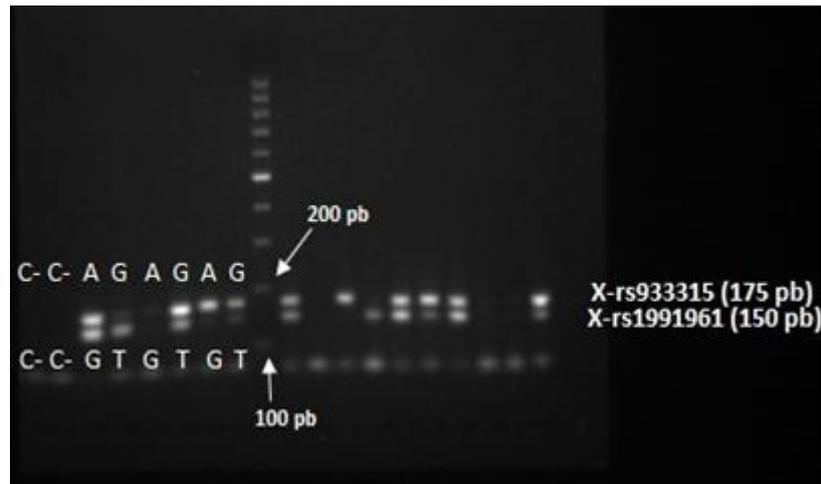


Figura 14. Electroforesis en gel de agarosa 1,6% de amplificaciones del dúplex X-rs933315-X-rs1991961. Se muestran los tamaños de cada amplicón. En el medio, una escalera de 100 pb. C-=control negativo, A=amplificación del alelo A del marcador X-rs933315, G (arriba)=amplificación del alelo G del marcador X-rs933315, G (abajo)=amplificación del alelo G del marcador X-rs1991961, T=amplificación del alelo T del marcador X-rs1991961. Las referencias para el marcador X-rs933315 se ven arriba de las bandas; las referencias para el marcador X-rs1991961 se ven debajo de las bandas que corresponden. Las bandas tenues por debajo de los 100 pb son inespecíficas (corresponden a dímeros de *primers*).

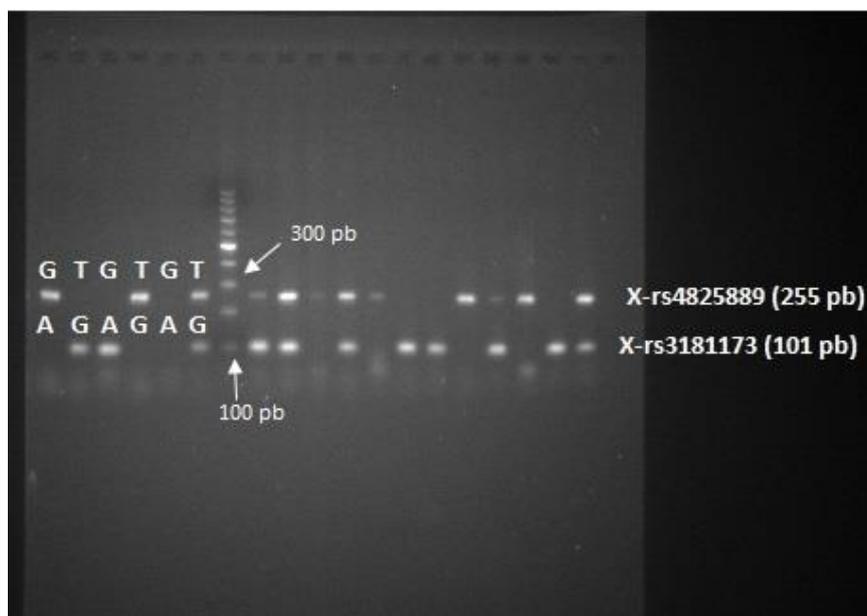


Figura 15. Electroforesis en gel de agarosa de amplicones del dúplex X-rs4825889-X-rs3181173. Se especifican los tamaños de los amplicones. En la calle 7, escalera de 100 pb. G (arriba)=alelo G del marcador X-rs4825889, T=alelo T del marcador X-rs4825889, A=alelo A del marcador X-rs3181173, G (abajo)=alelo G del marcador X-rs3181173.

Los alelos del marcador rs1299087 se determinaron por la técnica PCR-RFLP. Las digestiones se corrieron en geles de poliacrilamida neutra al 8% para observar los fragmentos. Se pueden dar tres casos:

- Si el o los alelos son G, la enzima MspI corta el amplicón de 156 pb dejando visibles sólo los fragmentos de 107 pb y 49 pb.
- Si el o los alelos son A, la enzima MspI no reconoce sitio de restricción y no digiere el amplicón; sólo se ve la banda intacta de 156 pb.
- En el caso de mujeres heterocigotas, la enzima corta sólo en los amplicones que posean el sitio de restricción (alelo G); se observan en la acrilamida fragmentos de 156 pb, 107 pb y 49 pb.

El fragmento de 49 pb es muy pequeño por ende migra más rápido quedando muy abajo en el gel o incluso llegando a caerse en algunas corridas. Por lo tanto, los alelos se tipificaron observando las bandas más grandes, de 156 pb y 107 pb. En la Figura 16 se muestra, a modo de ejemplo, la imagen obtenida de una corrida electroforética en gel del poliacrilamida.

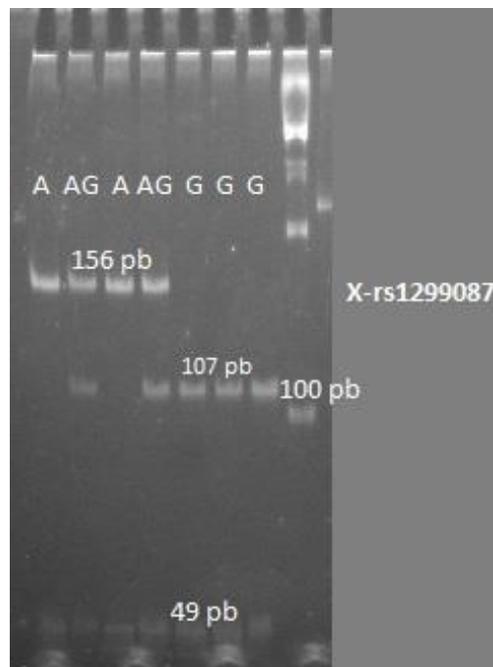


Figura 16. Electroforesis de digestiones del amplicón del marcador X-rs1299087 en poliacrilamida neutra 8%. Si está presente el alelo A se observa una banda de 156 pb (A); si está presente el alelo G se observan bandas de 107 pb y 49 pb (G); si están presentes ambos alelos se observan las tres bandas. En la derecha se encuentra una escalera de 100 pb utilizada como marcador de peso molecular.

Frecuencias alélicas

Se calcularon las frecuencias alélicas para hombres y mujeres por separado, y luego las totales, y se muestran en las Tablas 9 y 10. Para cada valor se calculó también el desvío estándar. Las mayores diferencias de frecuencias entre ambos alelos se vieron en los marcadores rs149910, rs318173 y rs652. En los dos primeros prevalece el alelo ancestral mientras que en el rs652 prevalece el otro alelo (G). En la Figura 17 se representan las frecuencias alélicas totales para cada uno de los X-SNPs.

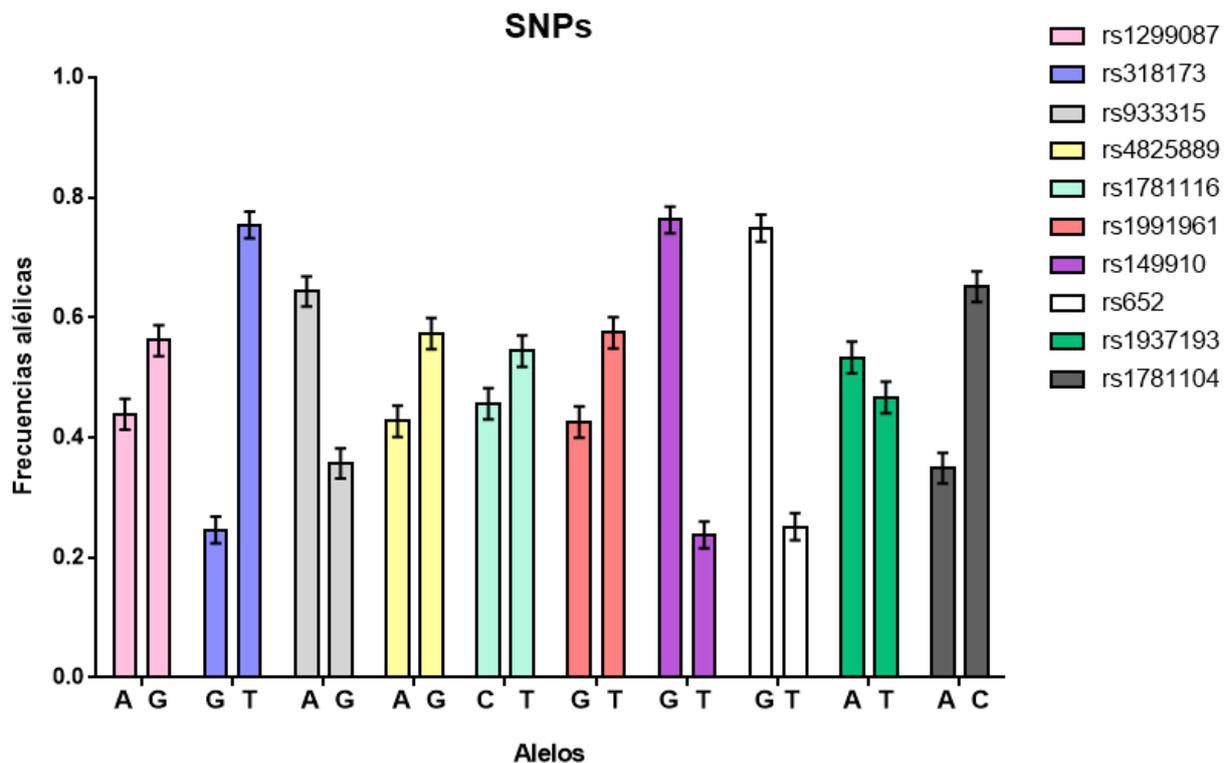


Figura 17. Frecuencias alélicas totales de los X-SNPs analizados. Debajo de las barras se indican los distintos alelos para cada marcador.

Tabla 9. Frecuencias alélicas en hombres (N=45), mujeres (N=71), y totales (N=116), para los X-SNPs analizados, junto a sus respectivas desviaciones estándar.

rs1299087	Frecuencias alélicas hombres	rs318173	Frecuencias alélicas hombres	rs933315	Frecuencias alélicas hombres	rs4825889	Frecuencias alélicas hombres	rs1781116	Frecuencias alélicas hombres
A	0,4667 +/- 0,0744	G	0,2889 +/- 0,0676	A	0,6222 +/- 0,0723	A	0,3556 +/- 0,0714	C	0,3182 +/- 0,0702
G	0,5333 +/- 0,0744	T	0,7111 +/- 0,0676	G	0,3778 +/- 0,0723	G	0,6444 +/- 0,0714	T	0,6818 +/- 0,0702
	Frecuencias alélicas mujeres		Frecuencias alélicas mujeres		Frecuencias alélicas mujeres		Frecuencias alélicas mujeres		Frecuencias alélicas mujeres
A	0,4296 +/- 0,0294	G	0,2324 +/- 0,0251	A	0,6500 +/- 0,0285	A	0,45 +/- 0,0297	C	0,5000 +/- 0,0301
G	0,5704 +/- 0,0294	T	0,7676 +/- 0,0251	G	0,3500 +/- 0,0285	G	0,55 +/- 0,0297	T	0,5000 +/- 0,0301
	Frecuencias alélicas totales		Frecuencias alélicas totales		Frecuencias alélicas totales		Frecuencias alélicas totales		Frecuencias alélicas totales
A	0,4385 +/- 0,0257	G	0,2460 +/- 0,0223	A	0,6432 +/- 0,0249	A	0,4270 +/- 0,0257	C	0,4560 +/- 0,0261
G	0,5615 +/- 0,0257	T	0,7540 +/- 0,0223	G	0,3568 +/- 0,0249	G	0,5730 +/- 0,0257	T	0,5440 +/- 0,0261

Tabla 10. Frecuencias alélicas en hombres (N=45), mujeres (N=71), y totales (N=116), para los X-SNPs analizados, junto a sus respectivas desviaciones estándar.

rs1991961	Frecuencias alélicas hombres	rs149910	Frecuencias alélicas hombres	rs652	Frecuencias alélicas hombres	rs1937193	Frecuencias alélicas hombres	rs1781104	Frecuencias alélicas hombres
G	0,4667 +/- 0,0744	G	0,6977 +/- 0,0700	G	0,7333 +/- 0,0659	A	0,6364 +/- 0,0725	Frecuencias alélicas mujeres	0,2667 +/- 0,0659
T	0,5333 +/- 0,0744	T	0,3023 +/- 0,0700	T	0,2667 +/- 0,0659	T	0,3636 +/- 0,0725	C	0,7333 +/- 0,0659
	Frecuencias alélicas mujeres		Frecuencias alélicas mujeres		Frecuencias alélicas mujeres		Frecuencias alélicas mujeres		Frecuencias alélicas mujeres
G	0,4118 +/- 0,0298	G	0,7826 +/- 0,0248	G	0,7535 +/- 0,0256	A	0,5000 +/- 0,0303	A	0,3769 +/- 0,0301
T	0,5882 +/- 0,0298	T	0,2174 +/- 0,0248	T	0,2465 +/- 0,0256	T	0,5000 +/- 0,0303	C	0,6231 +/- 0,0301
	Frecuencias alélicas totales		Frecuencias alélicas totales		Frecuencias alélicas totales		Frecuencias alélicas totales		Frecuencias alélicas totales
G	0,4254 +/- 0,0260	G	0,7624 +/- 0,0224	G	0,7487 +/- 0,0224	A	0,5333 +/- 0,0263	A	0,3486 +/- 0,0255
T	0,5746 +/- 0,0260	T	0,2376 +/- 0,0224	T	0,2513 +/- 0,0224	T	0,4667 +/- 0,0263	C	0,6514 +/- 0,0255

Frecuencias genotípicas

Se calcularon las frecuencias genotípicas con los datos femeninos únicamente, ya que los hombres son hemicigotas para estos marcadores. En las Tablas 11 y 12 se muestran los valores, junto a sus desviaciones estándar. Se observa que los marcadores rs318173, rs652 y rs149910 tienen bajas frecuencias en los genotipos homocigotas G/G, T/T y T/T, respectivamente. También se destaca la frecuencia del genotipo homocigota T/T para el marcador rs149910, que tiene un valor muy similar a la frecuencia del heterocigota G/T. En la Figura 18 se representan las frecuencias genotípicas obtenidas, graficadas con el programa GraphPad Prism 6.

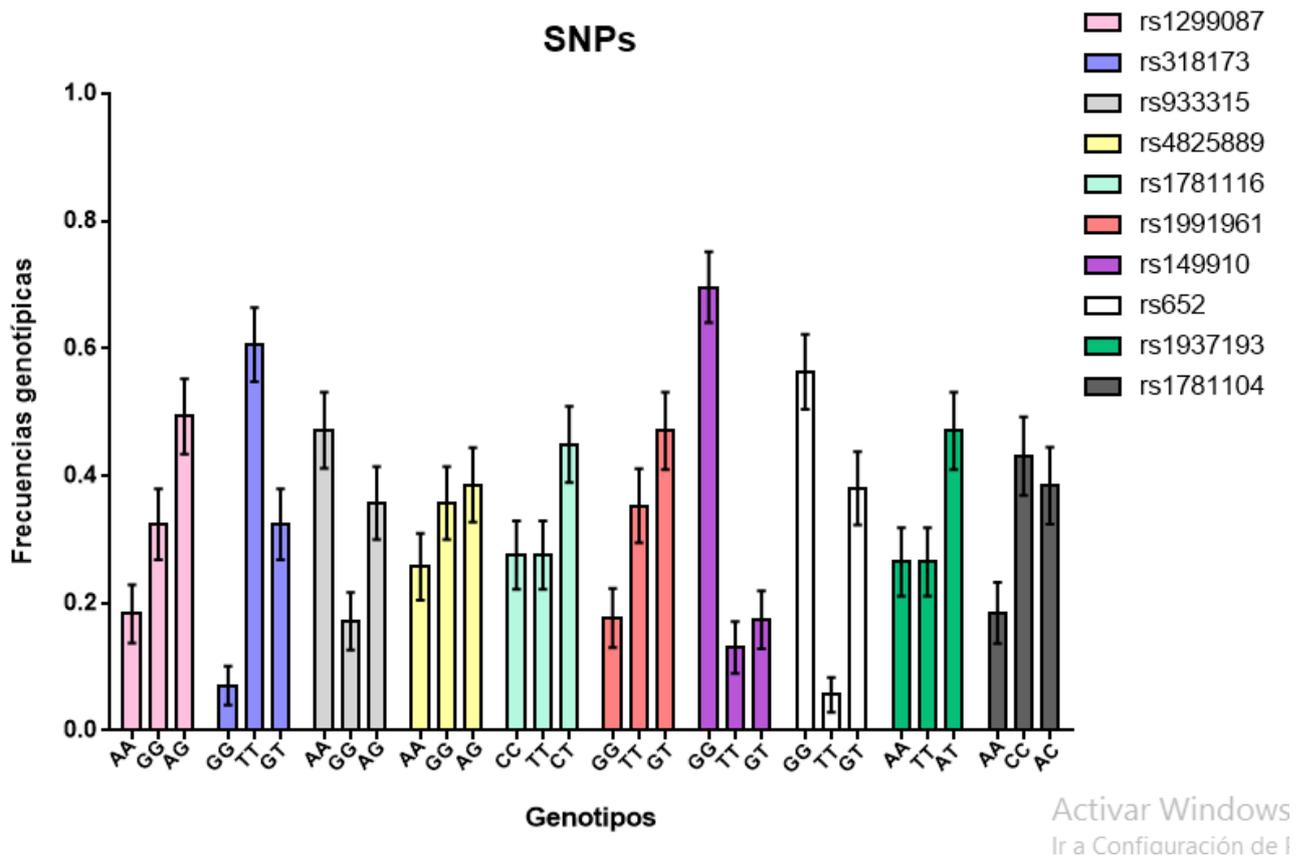


Figura 18. Frecuencias genotípicas obtenidas para cada uno de los X-SNPs. Debajo de cada barra se muestra el genotipo correspondiente.

Tabla 11. Frecuencias genotípicas de los marcadores analizados, calculadas con los datos femeninos (N=71), cada dato con su desviación estándar.

rs1299087	Frecuencias genotípicas	rs318173	Frecuencias genotípicas	rs933315	Frecuencias genotípicas	rs4825889	Frecuencias genotípicas	rs1781116	Frecuencias genotípicas
A-A	0,1831 +/- 0,0459	G-G	0,0704 +/- 0,0304	A-A	0,4714 +/- 0,0597	A-A	0,2571 +/- 0,0522	C-C	0,2754 +/- 0,0538
G-G	0,3239 +/- 0,0555	T-T	0,6056 +/- 0,0580	G-G	0,1714 +/- 0,0450	G-G	0,3571 +/- 0,0573	T-T	0,2754 +/- 0,0538
A-G	0,4930 +/- 0,0593	G-T	0,3239 +/- 0,0555	A-G	0,3571 +/- 0,0573	A-G	0,3857 +/- 0,0582	C-T	0,4493 +/- 0,0599

Tabla 12. Frecuencias genotípicas de los marcadores analizados, calculadas con los datos femeninos (N=71), cada dato con su desviación estándar.

rs1991961	Frecuencias genotípicas	rs149910	Frecuencias genotípicas	rs652	Frecuencias genotípicas	rs1937193	Frecuencias genotípicas	rs1781104	Frecuencias genotípicas
G-G	0,1765 +/- 0,0462	G-G	0,6957 +/- 0,0554	G-G	0,5634 +/- 0,0589	A-A	0,2647 +/- 0,0535	A-A	0,1846 +/- 0,0481
T-T	0,3529 +/- 0,0580	T-T	0,1304 +/- 0,0405	T-T	0,0563 +/- 0,0274	T-T	0,2647 +/- 0,0535	C-C	0,4308 +/- 0,0614
G-T	0,4706 +/- 0,0605	G-T	0,1739 +/- 0,0456	G-T	0,3803 +/- 0,0576	A-T	0,4706 +/- 0,0605	A-C	0,3846 +/- 0,0603

Desequilibrio de ligamiento

Se realizó el análisis de desequilibrio de ligamiento entre los X-SNPs utilizando los datos femeninos. En la Tabla 13 se muestran los valores p de este análisis. Los valores p significativos ($p < 0,05$, marcados con negrita y subrayado) indican que los marcadores se encuentran ligados. Los resultados obtenidos concuerdan con ocho pares de marcadores ligados.

Equilibrio de Hardy-Weinberg

Para este cálculo se evalúan sólo los datos femeninos, que son los que ofrecen genotipos para este tipo de marcadores. En la Tabla 14 se muestran los valores de heterocigosis esperadas y observadas. El marcador X-rs149910 no se ajustó al equilibrio, mientras que el resto de los marcadores sí se encontraron dentro de los valores esperados para el ajuste al equilibrio. Cabe destacar que los marcadores X-rs933315 y X-rs4825889 obtuvieron valores de p cercanos al valor límite de 0,05.

Haplotipos

Se corrieron los datos masculinos sin *missing data* (N=41) y se obtuvieron 34 haplotipos diferentes. La mayor frecuencia obtenida fue de 0,0732, que corresponde a tres individuos con el mismo haplotipo. Los haplotipos se muestran en el Apéndice 4.

La diversidad génica promedio entre todos los loci resultó de 0,4600 +/- 0,2556. La heterocigosis media fue de 0,3857 +/- 0,0933.

Tabla 13. Valores p para el análisis de desequilibrio de ligamiento para los X-SNPs. Los valores en negrita y subrayado indican un ligamiento positivo entre los marcadores ($p < 0,05$). Los marcadores están ordenados en la tabla según su orden en el cromosoma X. N=70, mujeres.

	rs1299087	rs318173	rs933315	rs1991961	rs1781104	rs4825889	rs1781116	rs1937193	rs149910	rs652
rs1299087										
rs318173	0,7573									
rs933315	0,2690	0,8325								
rs1991961	0,1230	0,9054	0,1715							
rs1781104	0,5170	0,6774	<u>0,0375</u>	0,9445						
rs4825889	1,0000	0,2401	0,2825	0,8362	<u>0,0000</u>					
rs1781116	1,0000	0,1685	0,2781	0,7813	<u>0,0000</u>	<u>0,0000</u>				
rs1937193	0,6657	0,5607	0,5178	0,2889	<u>0,0000</u>	<u>0,0000</u>	<u>0,0000</u>			
rs149910	0,1992	0,4745	0,4993	0,2555	0,8054	0,0875	0,1813	0,1745		
rs652	0,5777	0,3949	0,6436	0,3318	0,3871	0,9876	0,4949	0,4820	<u>0,0002</u>	

Tabla 14. Se ven los valores de heterocigosis observada y esperada para cada uno de los X-SNPs analizados, junto a los valores de p y su desviación estándar. En **negrita y subrayado**, el valor de p significativo indica que el marcador en cuestión no se encuentra en equilibrio de Hardy-Weinberg. N=70, mujeres.

Marcador	Nº de genotipos	Heterocigosis observada	Heterocigosis esperada	Valor p
rs1299087	70	0,4857	0,4933	1,0000 +/- 0,0000
rs318173	70	0,3286	0,3629	0,5052 +/- 0,0005
rs933315	69	0,3478	0,4570	0,0618 +/- 0,0002
rs4825889	70	0,3857	0,4986	0,0880 +/- 0,0003
rs1781116	69	0,4493	0,5036	0,4677 +/- 0,0005
rs1991961	68	0,4706	0,4880	0,8060 +/- 0,0004
rs149910	69	0,1739	0,3427	<u>0,0002</u> +/- 0,0000
rs652	70	0,3714	0,3704	1,0000 +/- 0,0000
rs1937193	68	0,4706	0,5037	0,6324 +/- 0,0005
rs1781104	64	0,3750	0,4724	0,1134 +/- 0,0003

Polimorfismos de Inserción-Delección

Los resultados de la amplificación de los X-Indels analizados se resolvieron en geles de poliacrilamida neutra al 8%. Se muestran algunas de las imágenes obtenidas para los diferentes X-Indels (Figuras 19, 20 y 21). Se observa una banda (del tamaño de la inserción o la delección) en el caso de mujeres homocigotas u hombres, y ambas bandas en el caso de mujeres heterocigotas.

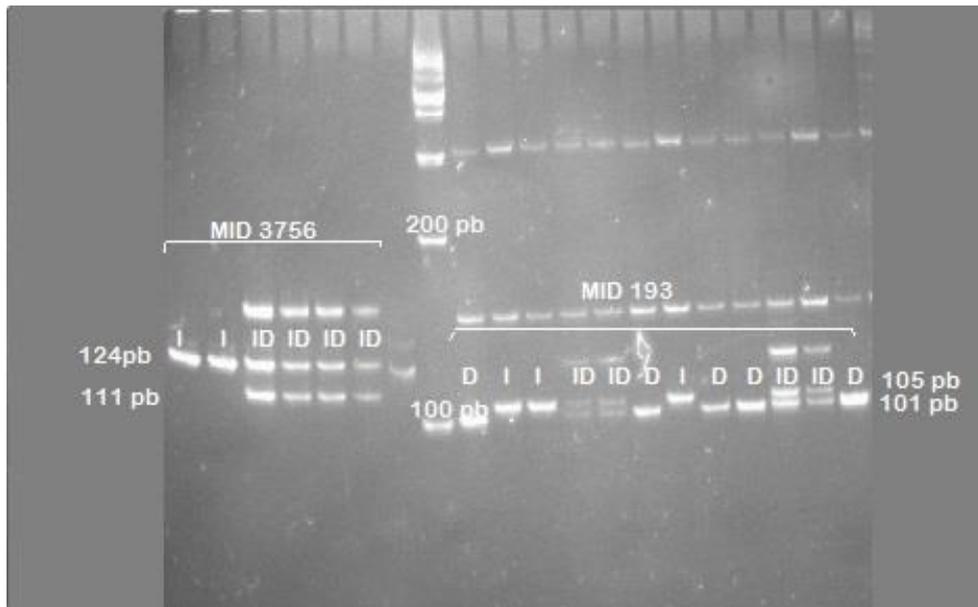


Figura 19. Corrida electroforética en gel de poliacrilamida de los X-Indels MID 3756 (izquierda) y MID 193 (derecha), y una escalera de 100 pb utilizada como marcador de peso molecular. I=banda que corresponde a la inserción, D=banda que corresponde a la delección, ID=ambas bandas. En la imagen se indican los tamaños moleculares de los diferentes amplicones.

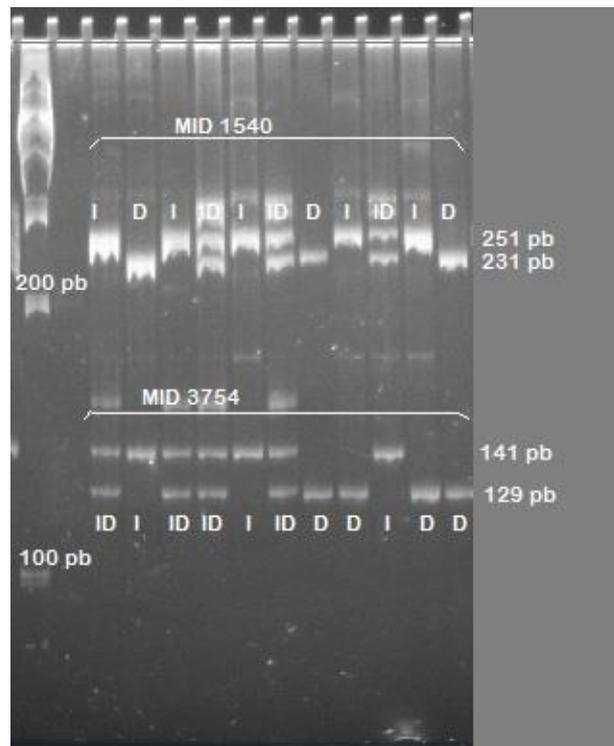


Figura 20. Electroforesis en gel de poliacrilamida del dúplex MID 1540-MID 3754, corridas junto a una escalera de 100 pb utilizada como marcador de peso molecular. I=presencia de la banda que corresponde a la inserción, D=presencia de la banda que corresponde a la delección, ID=presencia de ambas bandas. Se indican en la imagen los tamaños moleculares de las distintas bandas.

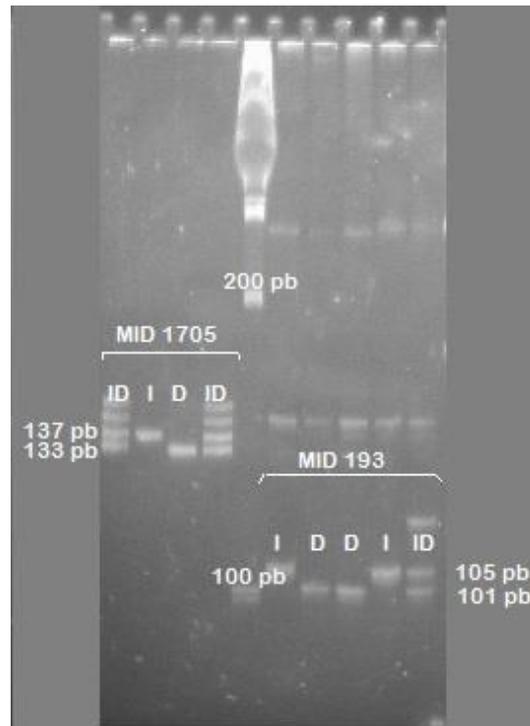


Figura 21. Electroforesis en poliacrilamida neutra al 8 % de los marcadores MID 1705 a la izquierda y MID 193 a la derecha. En el medio, una escalera de 100 pb utilizada como marcador de peso molecular. I=inserción, D=delección, ID=inserción y delección.

Frecuencias alélicas

Se calcularon las frecuencias alélicas para hombres y mujeres por separado, y las totales. La mayor variación de frecuencias entre alelos se observó para el marcador MID 1705, seguido por los marcadores MID193 y MID 3754. En todos los marcadores, sin embargo, se observó una menor frecuencia para el alelo correspondiente a la delección. Todas las frecuencias se muestran en la Tabla 15.

Se muestran las frecuencias alélicas totales, graficadas con el programa GraphPad Prism 6, en la Figura 22.

Tabla 15. Frecuencias alélicas de los X-Indels analizados. Varones: N=45, Mujeres N=71, total: N=116.

Frec. alélicas hombres	MID 1705	MID 3756	MID 3754	MID 193	MID 1540
Inserción	0,8444 +/- 0,0540	0,5778 +/- 0,0736	0,5333 +/- 0,0744	0,5556 +/- 0,0741	0,5333 +/- 0,0744
Delección	0,1556 +/- 0,0540	0,4222 +/- 0,0736	0,4667 +/- 0,0744	0,4444 +/- 0,0741	0,4667 +/- 0,0744
Frec. alélicas mujeres					
Inserción	0,8099 +/- 0,0233	0,5857 +/- 0,0294	0,6786 +/- 0,0279	0,7254 +/- 0,0265	0,5929 +/- 0,0294
Delección	0,1901 +/- 0,0233	0,4143 +/- 0,0294	0,3214 +/- 0,0279	0,2746 +/- 0,0265	0,4071 +/- 0,0294
Frec. alélicas totales					
Inserción	0,8182 +/- 0,0199	0,5838 +/- 0,0256	0,6432 +/- 0,0249	0,6845 +/- 0,0240	0,5784 +/- 0,0257
Delección	0,1818 +/- 0,0199	0,4162 +/- 0,0256	0,3568 +/- 0,0249	0,3155 +/- 0,0240	0,4216 +/- 0,0257

Referencias: Frec= frecuencias.

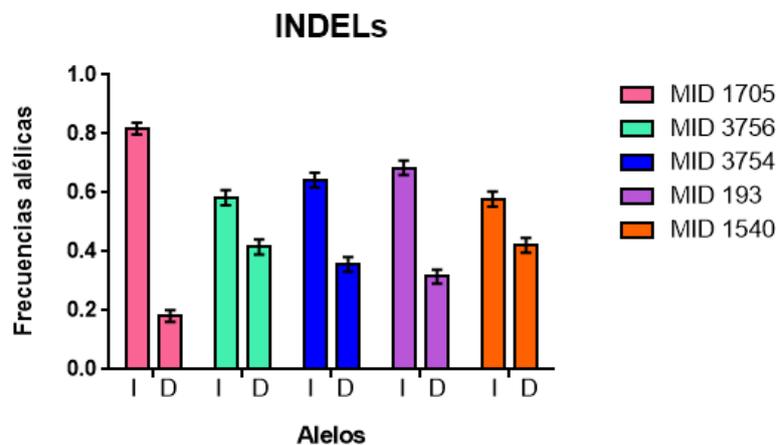


Figura 22. Frecuencias alélicas totales obtenidas en los X-Indels analizados. I=inserción, D=delección.

Frecuencias genotípicas

Se calcularon las frecuencias genotípicas con los datos femeninos, y sus respectivas desviaciones estándar, y se muestran en la Tabla 16. Se observa una baja prevalencia del genotipo homocigota para la delección en los marcadores MID 1705, MID 193 y MID 3754. Asimismo, en todos los marcadores se observó que la frecuencia de los homocigotas para la delección fue menor que las frecuencias del homocigota inserción y el heterocigota. En la Figura 23 se representan las frecuencias genotípicas para los X-Indels.

Tabla 16. Frecuencias genotípicas de los X-Indels analizados (Mujeres, N=71).

Frecuencias genotípicas	MID 1705	MID 3756	MID 3754	MID 193	MID 1540
II	0,6479 +/- 0,0567	0,3571 +/- 0,0573	0,4429 +/- 0,0594	0,3803 +/- 0,0576	0,3429 +/- 0,0567
DD	0,0282 +/- 0,0196	0,1857 +/- 0,0465	0,0857 +/- 0,0335	0,0845 +/- 0,0330	0,1571 +/- 0,0435
ID	0,3239 +/- 0,0555	0,4571 +/- 0,0595	0,4714 +/- 0,0597	0,5352 +/- 0,0592	0,5000 +/- 0,0598

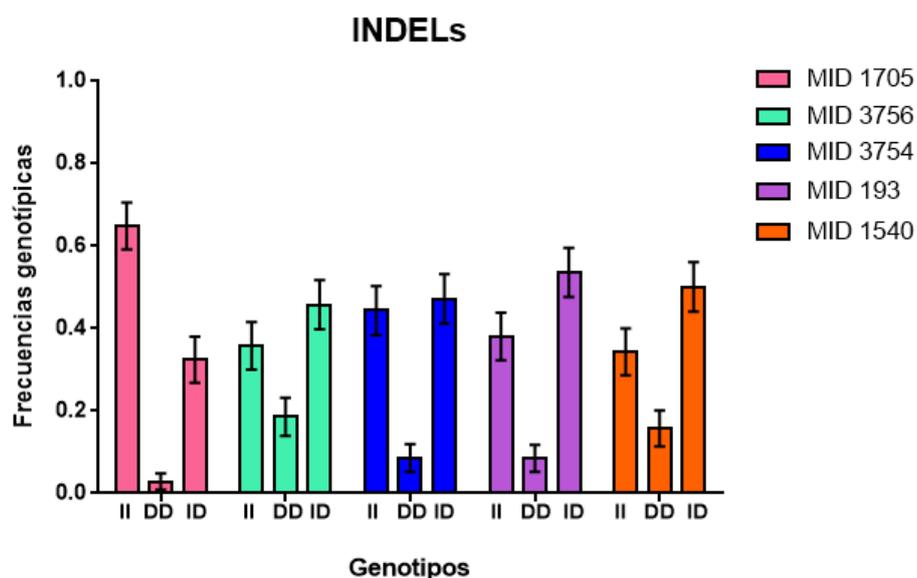


Figura 23. Frecuencias genotípicas de los X-Indels analizados. II=homocigota para la inserción, DD=homocigota para la delección, ID=heterocigota.

Desequilibrio de ligamiento

El análisis de desequilibrio de ligamiento se realizó con los datos femeninos; en la Tabla 17 se muestran los valores p. Dos pares de marcadores resultaron ligados: MID 3756-MID 3754 y MID 1540-MID 3754, ambos con un valor significativo ($p < 0,05$).

Tabla 17. Valores p para el desequilibrio de ligamiento entre los X-Indels analizados. Los valores marcados en negrita y subrayados indican un ligamiento positivo ($p < 0,05$). N=70, mujeres. Los marcadores se encuentran ordenados en la tabla según su ubicación cromosómica.

	MID 3754	MID 3756	MID 1705	MID 193	MID 1540
MID 3754					
MID 3756	<u>0,0097</u>				
MID 1705	0,8270	0,1547			
MID 193	0,7489	0,4801	0,2247		
MID 1540	<u>0,0000</u>	0,1158	0,5826	0,6738	

Equilibrio de Hardy-Weinberg

Los valores de heterocigosis observada y esperada, junto con el valor p, se muestran en la Tabla 18. Se observa que ambos valores de heterocigosis son similares para los cinco X-Indels analizados, y los valores p resultaron mayores a 0,05, por lo que todos los marcadores se encuentran en equilibrio de Hardy-Weinberg en esta población.

Haplotipos

Para el conjunto de los X-Indels, se encontraron 18 diferentes haplotipos entre los 45 hombres, con frecuencias que varían desde 0,0222 (un individuo) a 0,1333 (seis individuos). El haplotipo más frecuente correspondió a la combinación de todos los alelos de inserción para los cinco X-Indels. Los diferentes haplotipos se muestran en el Apéndice 5.

La diversidad genética promedio obtuvo un valor de $0,4172 \pm 0,2596$.

Tabla 18. Heterocigosis observada y esperada para los diferentes marcadores, junto a sus valores p. Ninguno resultó significativo ($p < 0,05$), ajustándose todos los marcadores al equilibrio de Hardy-Weinberg.

Marcador	Nº de genotipos	Heterocigosis observada	Heterocigosis esperada	Valor p
MID 1705	71	0,3235	0,3102	1,0000 +/- 0,0000
MID 3756	70	0,4571	0,4888	0,6305 +/- 0,0005
MID 3754	70	0,4714	0,4394	0,5915 +/- 0,0005
MID 193	71	0,3803	0,4013	0,7659 +/- 0,0004
MID 1540	70	0,5000	0,4862	1,0000 +/- 0,0000
Heterocigosis media		0,4252 +/- 0,0738		

VARIACIÓN ENTRE BUENOS AIRES Y OTRAS POBLACIONES

Polimorfismos de Nucleótido Simple

Para el caso de los X-SNPs, se utilizaron datos de poblaciones ya tipificadas para estos marcadores. Hasta el presente se cuenta con datos de las poblaciones chaqueñas nativas amerindias (Wichí y Mocoví), Resistencia y Misión Nueva Pompeya (ciudades chaqueñas), (Glesmann, 2015). En la comparación de todos los marcadores se calcularon los índices de F_{ST} entre las poblaciones, y se realizó un AMOVA. También se evaluaron las diferencias entre poblaciones marcador a marcador.

Índice de F_{ST} entre poblaciones (para todos los X-SNPs)

El índice de F_{ST} nos muestra qué tan distintas son las poblaciones genéticamente, obteniéndose un valor mayor cuanto mayor es la diferencia. Los valores de F_{ST} obtenidos con el programa Arlequin 3.5.2.2 entre los diferentes pares de poblaciones se muestran en la Tabla 19.

Tabla 19. Valores de FST entre las poblaciones para los X-SNPs analizados. Los valores que resultaron significativos se marcan en negrita y subrayado ($p < 0,05$).

	Bs As	Wichí	Mocoví	Resistencia	MNP
Bs As					
Wichí	<u>0,1411</u>				
Mocoví	<u>0,0687</u>	<u>0,0385</u>			
Resistencia	<u>0,0108</u>	<u>0,1006</u>	0,0296		
MNP	<u>0,0442</u>	<u>0,0899</u>	<u>0,0580</u>	<u>0,0276</u>	

Referencias: Bs As= Buenos Aires, MNP=Misión Nueva Pompeya

Todos los valores p resultaron menores a 0,05, excepto el de la comparación de Mocoví con Resistencia, pero que obtuvo un valor p de 0,0540. Esto indica que hay diferencias significativas entre casi todas las poblaciones, teniendo en cuenta el conjunto de los 10 X-SNPs.

Los valores de FST más altos se obtuvieron en la comparación de Buenos Aires con Wichí, y le sigue Buenos Aires con Mocoví.

Análisis de Varianza Molecular (AMOVA)

Para el AMOVA de los X-SNPs, se crearon tres grupos de poblaciones:

- Wichí con Mocoví, comunidades aborígenes nativas de la provincia de Chaco.
- Buenos Aires.
- Resistencia con Misión Nueva Pompeya, ambas ciudades de la provincia de Chaco.

En el análisis de AMOVA obtenemos tres porcentajes de variación: entre los grupos armados, entre las poblaciones dentro de los grupos y entre los individuos dentro de las poblaciones. Siempre el porcentaje de variación dentro de las poblaciones es mayor porque indica la diferencia genética entre personas individuales. En este caso, como se ve en la Tabla 20, observamos que el porcentaje de variación entre grupos y dentro de ellos es muy similar. Los valores no son significativos, ya

que la probabilidad p tuvo un valor de 0,216. Esto significa que no hay diferencias significativas entre los grupos de poblaciones comparados.

Tabla 20. Valores del AMOVA para el conjunto de SNPs analizados. Valor $p=0,216$. D.f= grados de libertad, este parámetro junto con la suma de cuadrados nos permiten calcular la significancia de los resultados.

Origen de la variación	d.f	Suma de cuadrados	Componente de la varianza	Porcentaje de variación
Entre grupos	2	45,646	0,0698	3,88
Entre poblaciones dentro de los grupos	2	12,601	0,0605	3,36
Entre individuos	567	946,765	1,6698	88,7229
Total	571	1005,012	1.8001	100

Índice de FST para los X-SNPs de forma individual

A fin de verificar qué polimorfismos fueron responsables de la mayor diferenciación entre poblaciones, se realizó también una comparación marcador a marcador. En la Tabla 21 se muestran los valores de FST y AMOVA obtenidos en estos análisis. Los grupos para el AMOVA son los mismos definidos anteriormente.

El análisis de AMOVA sólo mostró diferencias significativas entre los grupos de poblaciones para dos de los marcadores: rs149910 y rs1991961.

La población de Buenos Aires es similar a la de Resistencia (FST no significativo) para la mayoría de los polimorfismos analizados, excepto en dos de ellos (rs1991961 y rs652), y resultó significativamente diferente a todas las poblaciones excepto a la de Resistencia con respecto a los marcadores rs149910, rs933315 y rs318173.

Además, para los marcadores rs4825889 y rs1781116 no se encontraron diferencias significativas entre la población de Buenos Aires y ninguna de las poblaciones restantes, si bien hay algunos datos significativos entre otros pares de poblaciones (datos no mostrados).

Tabla 21. Valores de los Índices de FST (negrita y subrayado indica $p < 0,05$) y AMOVA para los X-SNPs de forma individual.

	rs1299087	rs149910	rs933315	rs4825889	rs1781116	rs1991961	rs318173	rs652	rs1937193	rs1781104
Bs As-Resistencia	0,0085	0,0143	0,0048	0,0140	0,0126	<u>0,0188</u>	0,0084	<u>0,0400</u>	-0,0059	0,0100
Bs As-Wichí	<u>0,2962</u>	<u>0,2594</u>	<u>0,4156</u>	0,0053	0,0057	-0,0093	<u>0,0694</u>	0,0181	0,0187	0,0288
Bs As-Mocoví	0,0205	<u>0,2467</u>	<u>0,1761</u>	-0,0186	0,0040	-0,0198	<u>0,1432</u>	-0,0015	-0,0152	-0,0287
Bs As-MNP	-0,0086	<u>0,0726</u>	<u>0,0777</u>	0,0049	0,0220	0,0008	<u>0,0282</u>	<u>0,1239</u>	<u>0,0546</u>	<u>0,0389</u>
AMOVA										
% de Variación entre grupos	7,80 p=0,152	<u>11,93</u> p=0,043	15,94 p=0,093	-3,20 p=0,758	-5,96 p=0,924	<u>1,76</u> p=0,042	-1,05 p=0,421	2,93 p=0,162	-3,69 p=0,819	0,20 p=0,397
% de Variación entre poblaciones, dentro de los grupos	3,52	0,17	3,63	4,65	7,85	-1,20	7,44	0,52	5,06	1,11

Referencias: Bs As= Buenos Aires, MNP=Misión Nueva Pompeya

Comparación con otras poblaciones del mundo

Para hacer una comparación con otras poblaciones del mundo se utilizaron los porcentajes de frecuencias alélicas para algunos de los marcadores del sitio web 1000 Genomes (<http://www.1000genomes.org>). Se eligieron datos de Italia y España (representando a Europa), México, Perú y Colombia (representando a América), y África (como grupo poblacional más alejado; dicha información incluye datos de diferentes países africanos). Con estas frecuencias se calcularon los índices de FST y el AMOVA. El análisis se hizo marcador a marcador y los datos que se usaron en la comparación, extraídos de la página web, se detallan en el Apéndice 5.

Los índices de FST obtenidos se muestran en la Tabla 22. La población de Buenos Aires obtuvo diferencias significativas con la población de África para todos los marcadores excepto para rs1991961 y rs1781104. Pero estos marcadores no mostraron diferencias significativas con el resto de las poblaciones, excepto rs1781104, que mostró diferencias entre España y otras poblaciones (datos no mostrados).

Para el resto de los marcadores, se obtuvo un mayor número de valores de FST significativos entre Buenos Aires y poblaciones americanas, que entre Buenos Aires y las poblaciones europeas.

Tabla 22. Valores de FST obtenidos para los distintos X-SNPs con otras poblaciones del mundo. Los valores en negrita y subrayado son significativos ($p < 0,05$)

	rs1299087	rs149910	rs933315	rs4825889	rs1991961	rs652	rs1937193	rs1781104
Bs As-África	<u>0,2375</u>	<u>0,2178</u>	<u>0,1326</u>	<u>0,2270</u>	0,0155	<u>0,4899</u>	<u>0,1666</u>	-0,0079
Bs As-Colombia	0,0020	<u>0,0472</u>	<u>0,0341</u>	<u>0,0383</u>	-0,0014	0,0116	-0,0075	-0,0069
Bs As-Perú	<u>0,0583</u>	<u>0,2841</u>	<u>0,2141</u>	0,0046	-0,0073	<u>0,0641</u>	<u>0,0358</u>	-0,0061
Bs As-México	-0,0070	<u>0,1966</u>	<u>0,0341</u>	0,0061	-0,0066	<u>0,0501</u>	-0,0056	0,0137
Bs As-España	-0,0057	0,0020	<u>0,0292</u>	<u>0,1716</u>	-0,0065	-0,0074	<u>0,1173</u>	<u>0,0724</u>
Bs As-Italia	-0,0057	-0,0006	-0,0005	<u>0,0507</u>	-0,0014	0,0068	<u>0,0274</u>	0,0137

Referencias: Bs As=Buenos Aires.

Para el análisis de AMOVA se agruparon las poblaciones en los siguientes cuatro grupos:

- Buenos Aires.
- México, Perú y Colombia (representarían América).
- España e Italia (representarían Europa).
- África.

Se realizó un AMOVA inicial comparando estas poblaciones. Para algunos de los marcadores se realizó un segundo análisis, agrupando Buenos Aires con América o con Europa (quedando 3 grupos de poblaciones), dependiendo de los valores de FST obtenidos (cuando se obtuvieron valores significativos de FST, se mantuvieron los agrupamientos iniciales). En ningún análisis se agrupó Buenos Aires con África porque los índices de FST para la mayoría de los marcadores mostraron diferencias significativas. Los resultados se muestran en la Tabla 23.

El SNP rs149910 mostró diferencias significativas entre grupos cuando se corrió el AMOVA agrupando Buenos Aires con las poblaciones europeas. Lo mismo ocurrió con el marcador rs933315.

En cuanto al SNP rs4825889, el primer AMOVA no reveló diferencias significativas, y en este caso se agrupó Buenos Aires con las poblaciones americanas para correr el segundo AMOVA. El segundo AMOVA no reveló diferencias significativas, y se destaca que los resultados de FST no diferenciaron poblaciones según su continente (datos no mostrados). Algo muy similar se observó para el locus rs1937193, aunque en este caso los resultados del segundo AMOVA sí fueron significativos.

Los resultados del AMOVA en la comparación de los cuatro grupos armados para el marcador rs1781104 no dieron resultados significativos y no se realizó ningún otro agrupamiento ya que todas las poblaciones parecen ser similares (FST no significativo) excepto España (datos no mostrados).

Tabla 23. Valores de AMOVA obtenidos entre grupos. Los valores en negrita y subrayado resultaron significativos ($p < 0,05$). Se realizó un primer AMOVA entre los cuatro grupos armados. Para algunos de los marcadores se realizó un segundo AMOVA, agrupando la población de Buenos Aires con Europa o con América, según los valores de FST obtenidos anteriormente.

	rs1299087	rs149910	rs933315	rs4825889	rs1991961	rs652	rs1937193	rs1781104
% variación (4 grupos)	6,95	10,07	8,55	5,94	0,39	<u>20,37</u>	9,49	1,01
Valor p	0,06	0,115	0,109	0,169	0,263	0,005	0,084	0,281
% variación agrupando Bs As con Europa	<u>8,83</u>	<u>13,96</u>	<u>10,84</u>	--	--	<u>23,58</u>	--	--
Valor p	0,014	0,030	0,043	--	--	0,001	--	--
% variación agrupando Bs As con América	--	--	--	9,69	--	--	<u>13,62</u>	--
Valor p	--	--	--	0,06	--	--	0,021	--

Referencia: Bs As=Buenos Aires

El marcador rs1991961 no mostró diferencias significativas en ninguno de los análisis. Todas las poblaciones parecen tener frecuencias similares para los alelos de este marcador (incluso África). En contraste, el marcador rs652 mostró diferencias significativas en el primer AMOVA con un alto porcentaje de variación entre grupos. Se planteó un segundo AMOVA, agrupando Buenos Aires con Europa, que también reveló diferencias significativas, aumentando en un 3% el porcentaje de variación.

Polimorfismos de Inserción-Delección

Se compararon los resultados obtenidos para los X-Indels con datos de poblaciones nativas Wichí y Mocoví de la provincia de Chaco, Resistencia, capital de Chaco, y Misión Nueva Pompeya, ciudad de la provincia de Chaco (Glesmann, 2015). También se comparó con datos de Viedma, capital de Río Negro (Bridi *et al.*, 2013).

Índice de FST entre poblaciones (para todos los X-Indels)

Se corrieron todos los datos en el Arlequín obteniéndose los valores de FST que se muestran en la Tabla 24. Todos los valores de p resultaron significativos, excepto entre Buenos Aires y Viedma, indicando que todas las poblaciones son genéticamente distintas en estos marcadores, excepto Viedma y Buenos Aires, que son parecidas. Se ve además que hay una menor diferencia entre las poblaciones urbanas (Resistencia vs Buenos Aires; Misión Nueva Pompeya vs Buenos Aires) que entre las poblaciones urbanas y nativas amerindias (Buenos Aires vs Mocoví, Buenos Aires vs Wichí).

Tabla 24. Índices de FST entre las poblaciones para los marcadores X-Indels analizados. Todos los valores en negrita y subrayado resultaron significativos ($p < 0,05$)

	Bs As	Wichí	Mocoví	Resistencia	Viedma	MNP
Bs As						
Wichí	<u>0,1626</u>					
Mocoví	<u>0,1615</u>	<u>0,0513</u>				
Resistencia	<u>0,0300</u>	<u>0,1145</u>	<u>0,0804</u>			
Viedma	-0,0039	<u>0,1354</u>	<u>0,1235</u>	<u>0,0142</u>		
MNP	<u>0,0880</u>	<u>0,0430</u>	<u>0,0232</u>	<u>0,0319</u>	<u>0,0631</u>	

Referencias: Bs As=Buenos Aires, MNP=Misión Nueva Pompeya.

Análisis de Varianza Molecular (AMOVA)

Para la corrida del AMOVA para los X-Indels se crearon tres grupos de poblaciones, y se agregó la población de Viedma, que estaba tipificada para estos marcadores (Bridi *et al.*, 2013):

- Wichí y Mocoví.
- Resistencia y Misión Nueva Pompeya.
- Buenos Aires y Viedma, que se agruparon por ser los dos que no obtuvieron un valor de FST significativo.

Se observan los resultados en la Tabla 25. En este análisis observamos que la variación molecular es mayor entre los grupos que entre las poblaciones dentro de los grupos. Sin embargo, si bien se observa una tendencia con un valor $p=0,083$, las diferencias no son significativas.

Tabla 25. Resultados del AMOVA para los X-Indels entre los tres grupos de poblaciones armados. Valor $p=0,083$.

Origen de la variación	d.f	Suma de cuadrados	Componente de la varianza	Porcentaje de variación
Entre grupos	2	41,396	0,0788	6,61
Entre poblaciones dentro de los grupos	3	9,749	0,0247	2,07
Entre individuos	640	696,815	1,0888	91,32
Total	645	747,960	1,1923	100

Índice de FST para los X-Indels de forma individual

Nuevamente se hicieron corridas con cada uno de los marcadores, para ver cuáles presentan la mayor variación entre poblaciones y son responsables de los valores de FST y AMOVA obtenidos.

Los grupos para el AMOVA son los mismos que se definieron arriba. Los resultados se muestran en la Tabla 26.

El análisis de varianza molecular (AMOVA) resultó significativo para dos de los marcadores: MID 193 y MID 1705, que son los mayores responsables de las variaciones observadas entre los grupos armados.

En cuanto a la comparación de las poblaciones entre sí, el marcador MID 193 dio índices de FST significativamente diferentes entre Buenos Aires y el resto de las poblaciones, excepto Viedma. Además, en el marcador MID 3754 se diferenció la población de Buenos Aires de las poblaciones Wichí, Mocoví, y de Misión Nueva Pompeya, pero no de las poblaciones de Resistencia ni de Viedma.

En el marcador MID 1705 la población de Buenos Aires obtuvo un FST significativo sólo con las poblaciones urbanas de Chaco: Misión Nueva Pompeya y Resistencia.

El marcador MID 3756 arrojó valores de FST no significativos entre todas las poblaciones (Tabla 26 y datos no mostrados), y porcentajes de variación prácticamente

nulos, indicando que las distribuciones alélicas y genotípicas para este marcador son muy similares en todas las poblaciones.

Tabla 26. Valores del índice de FST y porcentajes de variación del análisis de varianza molecular (AMOVA) para cada X-Indel de forma individual. Los valores en **negrita** y **subrayado** son significativos ($p < 0,05$).

	MID 1540	MID 193	MID 3754	MID 1705	MID 3756
Bs As-Resistencia	-0,0015	<u>0,0934</u>	0,0041	<u>0,0690</u>	-0,0015
Bs As-Wichí	<u>0,0868</u>	<u>0,1862</u>	<u>0,4182</u>	-0,0027	-0,0057
Bs As-Mocoví	0,0112	<u>0,2834</u>	<u>0,3711</u>	0,0072	-0,0065
Bs As-MNP	-0,0156	<u>0,1160</u>	<u>0,2370</u>	<u>0,1290</u>	-0,0031
Bs As-Viedma	-0,0099	0,0006	-0,0019	-0,0098	-0,0077
AMOVA					
% Variación entre grupos	-0,17 p=0,406	<u>10,25</u> p=0,010	17,12 p=0,071	<u>5,92</u> p=0,010	-0,59 p=0,720
% Variación entre poblaciones, dentro de los grupos	3,64	-0,22	6,07	-0,62	0,22

Referencias: Bs As=Buenos Aires, MNP=Misión Nueva Pompeya.

Comparación con otras poblaciones del mundo

Para los X-Indels, sólo se encontraron datos de frecuencias alélicas de otros lugares del mundo para los marcadores MID 193 Y MID 1540. Los datos obtenidos de la página 1000 Genomes (<http://www.1000genomes.org>) se muestran en el Apéndice 7.

Los valores de FST y AMOVA obtenidos se muestran en la Tabla 27. Los grupos para el AMOVA fueron los mismos cuatro grupos que se armaron para el AMOVA de los X-SNPs (ver página 61).

Tabla 27. Valores de FST y AMOVA (entre grupos) obtenidos de la comparación de las frecuencias alélicas de Buenos Aires y otras regiones del mundo, para los X-Indels MID 1540 y MID 193. Los valores en negrita y subrayado son significativos ($p < 0,05$).

FST	MID 1540	MID 193
Bs As-África	<u>0,3076</u>	<u>0,1240</u>
Bs As-Colombia	0,0062	-0,0032
Bs As-Perú	0,0002	<u>0,2181</u>
Bs As-México	-0,0076	<u>0,0771</u>
Bs As-España	-0,0047	0,0027
Bs As-Italia	-0,0047	-0,0077
AMOVA		
% de variación (4 grupos)	<u>9,88</u>	7,73
Valor p	0,009	0,179

El marcador MID 1540 mostró diferencias significativas en el AMOVA, pero sólo se detectaron valores de FST significativos entre África y el resto de las poblaciones. Sabiendo esto, se realizó un segundo AMOVA excluyendo el grupo africano, y se obtuvo una variación de 0,57% y un valor p de 0,202, mostrando que la gran variación obtenida en el primer AMOVA era entre el grupo “África” con el resto de los grupos.

En cuanto al MID 193, los índices de FST muestran diferencias entre Buenos Aires con México, Perú, y África. Los valores del AMOVA para los cuatro grupos comparados no fueron significativos. Se realizó un segundo AMOVA, agrupando Buenos Aires con las poblaciones europeas, y se obtuvo una variación entre grupos de 11,97%, y un valor p límite de 0,058. Esto indica que esta última agrupación sería la más correcta, siendo Buenos Aires más parecido a Europa que a América para este marcador.

ANÁLISIS CONJUNTO DE TODOS LOS MARCADORES (INDELS + SNPS)

Haplotipos

Se encontraron 41 haplotipos diferentes teniendo en cuenta todos los marcadores en conjunto, y ninguno se repitió. La frecuencia haplotípica para cada uno fue de 0,0244. Los diferentes haplotipos se muestran en el Apéndice 8.

Desequilibrio de ligamiento

Se realizó una corrida en el Arlequín con todos los datos para determinar si había marcadores del tipo Indel y del tipo SNP ligados. Se corrieron los datos femeninos (N=70), obteniéndose 4 grupos de ligamiento además de los descriptos anteriormente para el análisis de SNPs y el de Indels por separado:

- rs652 y MID 1705
- rs1781104 y MID 3756
- rs1781116 y MID 3756
- rs149910 y MID 1705

Índice de FST

Los resultados de FST para todos los marcadores revelan diferencias significativas entre todas las poblaciones (Tabla 28). Los valores más altos se observan entre las poblaciones de Buenos Aires y Wichí, Buenos Aires y Mocoví, y Resistencia y Wichí.

Análisis de Varianza Molecular (AMOVA)

Al correr el AMOVA, los valores de variación entre grupos para todos los marcadores en conjunto no arrojaron resultados significativos entre los grupos de poblaciones: el valor p fue de 0,195 (valor incluso más alto que en la comparación de los X-Indels solos). Los grupos fueron los mismos que los armados anteriormente en la corrida de los X-SNPs (ver página 58) y los resultados se muestran en la Tabla 29.

Tabla 28. Índices de FST entre las poblaciones para todos los marcadores analizados en el cromosoma X. Todos los valores resultaron significativos ($p < 0,05$) y se resaltan en negrita y subrayado.

	Bs As	Wichí	Mocoví	Resistencia	MNP
Bs As					
Wichí	<u>0,1560</u>				
Mocoví	<u>0,1194</u>	<u>0,0399</u>			
Resistencia	<u>0,0186</u>	<u>0,1109</u>	<u>0,0600</u>		
MNP	<u>0,0617</u>	<u>0,0769</u>	<u>0,0503</u>	<u>0,0295</u>	

Referencias: Bs As=Buenos Aires, MNP=Misión Nueva Pompeya.

Tabla 29. Resultados del AMOVA para todos los marcadores del cromosoma X entre los tres grupos de poblaciones armados. El valor p resultó de 0,195, por lo que la diferencia entre grupos no es significativa.

Origen de la variación	d.f	Suma de cuadrados	Componente de la varianza	Porcentaje de variación
Entre grupos	2	78,228	0,14176	4,74
Entre poblaciones dentro de los grupos	2	18,893	0,08958	2,99
Entre individuos	541	1493,672	2,76095	92,27
Total	545	1590,793	2,99229	100

ANÁLISIS DE LA ESTRUCTURA POBLACIONAL

Se corrieron los datos de las poblaciones de Buenos Aires, Wichí, Mocoví, Misión Nueva Pompeya y Resistencia con el programa Structure 2.3.4 (Pritchard *et al.*, 2000) para ver la estructura poblacional. Se propuso un modelo de ancestría mixta, con tres posibles ancestrías (que podrían ser americana, europea y africana). Los gráficos

obtenidos se observan en las Figuras 24 y 25. En la Figura 24 se muestra el *triangle plot*, un gráfico triangular en el que cada vértice del triángulo representa un agrupamiento. Cada población tiene asignado un color, y los puntos representan individuos de la misma. El programa trata de agrupar a los individuos según los posibles ancestros comunes. El gráfico obtenido no muestra una estructura muy definida ya que se ven puntos de todas las poblaciones en casi todos los vértices. Sin embargo, puede estimarse una tendencia de agrupamiento de las poblaciones de Buenos Aires (fucsia) y Resistencia (rojo) hacia el agrupamiento 2. Se ve entonces una tendencia al agrupamiento de las poblaciones urbanas entre sí. Además se observa que la población Wichí (verde) tiende a estar alejada de ese agrupamiento. Se observa que la población de Misión Nueva Pompeya (amarillo) es la más dispersa.

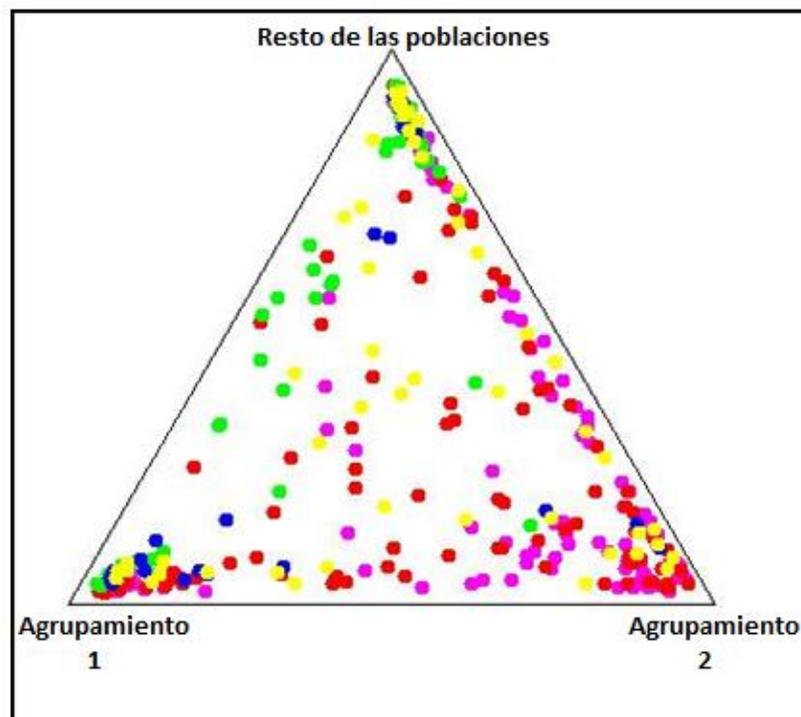


Figura 24. *Triangle plot* obtenido en la corrida del Structure para $K=3$. Los colores representan las diferentes poblaciones: Buenos Aires (fucsia), Resistencia (rojo), Wichí (verde), Mocoví (azul), y Misión Nueva Pompeya (amarillo).

En la Figura 25 se observa el *Bar plot in single line* obtenido. En este gráfico los números en el eje x representan las cinco poblaciones. Cada individuo está representado por una línea vertical particionada en k (3) segmentos coloreados que representan las fracciones de pertenencia asignadas a cada agrupamiento. La

población de Buenos Aires (5) se ve más parecida a la de Resistencia (1), si bien se ve que tiene mayor proporción de la ancestría indicada en color verde y menos de la indicada en color azul. A su vez, en comparación con la población Wichí (2) es notable que esta última población es más cerrada, mientras que la bonaerense es una población más abierta al flujo genético.

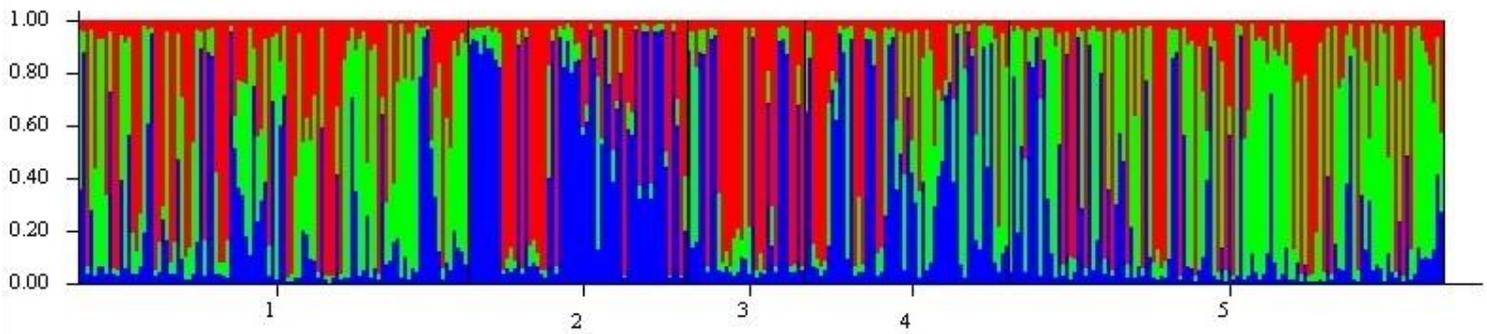


Figura 25. *Bar plot in single line.* Los colores representan los diferentes orígenes ancestrales y las líneas verticales a cada una de las personas de las poblaciones: Resistencia (1), Wichí (2), Mocoví (3), Misión Nueva Pompeya (4) y Buenos Aires (5).

DISCUSIÓN

La diversidad genética humana se moldea tanto por factores demográficos, como biológicos, y tiene implicaciones fundamentales para la comprensión de la base genética de las enfermedades (Li *et al.*, 2008). El estudio de la variabilidad en el cromosoma X nos da información no sólo de las fuerzas de cambio, como la selección natural, sino también sobre las diferencias ocurridas en la historia de hombres y mujeres de una población (Keinan *et al.*, 2009). Gracias a esto, el análisis de la variación genética ligada al cromosoma X puede ayudar a tener un enfoque diferente de cómo influyen la demografía y la selección en el genoma humano (Casto *et al.*, 2010).

En el presente trabajo se buscó comprender cómo la fusión cultural ocurrida en la población bonaerense incidió en la distribución de su variación genética, a través del estudio de dos tipos de polimorfismos genéticos de la región no pseudoautosómica del cromosoma X, SNPs e Indels, generando un patrón particular de variación.

VARIACIÓN DENTRO DE LA POBLACIÓN DE BUENOS AIRES

En el presente trabajo, todos los X-Indels analizados se ajustaron al equilibrio de Hardy-Weinberg, por lo cual en este momento no estaría ocurriendo en forma intensiva ningún proceso de cambio poblacional. Aunque es evidente la existencia de flujo genético en la población bonaerense, ya sea desde el interior del país como desde países limítrofes (INDEC, Censo 2010), éste no estaría ejerciendo una influencia marcada, o bien no se observan cambios a través del análisis de estos marcadores. Lo mismo se observó para los X-SNPs, excepto en el caso del marcador rs149910, el cual presentó un déficit de heterocigotas. Este desajuste al equilibrio de Hardy-Weinberg puede deberse a varias causas: a) el locus puede estar bajo selección; b) pueden existir alelos nulos en la población; c) pueden existir altos niveles de consanguinidad, producto de apareamiento entre individuos emparentados; d) o bien se presenta una cierta subestructuración reproductiva de la población (efecto Wahlund) (Aranguren-Méndez *et al.*, 2005). Por tratarse de marcadores no codificantes, la

selección natural no estaría ejerciendo un efecto directo, por lo cual ésta no sería la causa posible. Se habla de alelos nulos cuando no pueden ser amplificados por PCR, probablemente debido a una mutación en la zona de hibridación del *primer* (Aranguren-Méndez *et al.*, 2005). La presencia de alelos nulos en algunas de las mujeres también podría llevar a obtener una elevada tasa de homocigosis para el SNP en cuestión. Una manera de evidenciar alelos nulos en marcadores de cromosoma X es a través del análisis de muestras masculinas. En el caso del marcador rs149910, se obtuvo amplificación positiva en todos los varones, excepto dos. De este modo habría una probable presencia de mutaciones que impidan la amplificación en un 4% de la población, aunque para confirmar esto se hace necesario el análisis del polimorfismo por secuenciación, utilizando *primers* más externos. Cabe destacar que este SNP también presentó desajuste del equilibrio en otra ciudad argentina, Resistencia (Glesmann, 2015). La posibilidad de consanguinidad se excluye, ya que siendo la población bonaerense tan grande y debido a las costumbres sociales que presenta, es muy poco probable que exista consanguinidad en nuestra población, y aunque así fuera, debería reflejarse en un mayor número de marcadores. Con respecto a la subestructuración, la población bonaerense está conformada por individuos de una gran diversidad de orígenes, y entre ellos suelen constituirse colectividades más o menos abiertas que pueden presentar una subestructura en relación con la población general (efecto Wahlund).

Los altos valores de heterocigosis y las altas diversidades génicas obtenidas para el resto de los marcadores en la población de Buenos Aires reflejan una importante variabilidad genética, considerando que para un marcador bialélico, el mayor valor que se puede obtener de heterocigosis en el equilibrio es 0,5 (Aranguren-Méndez *et al.*, 2005).

Con respecto al ligamiento entre marcadores, en la población de Buenos Aires se encontraron dos pares de X-Indels ligados, uno de los cuales parece corresponderse con la localización citogenética de los marcadores, que se encuentran muy cerca entre sí. Por el contrario, es notable que el par MID 3754-MID 1540 presenta desequilibrio de ligamiento, tratándose de los dos Indels más alejados de

los cinco estudiados. Existen hallazgos similares de otros autores que muestran la existencia de desequilibrio de ligamiento entre polimorfismos de cromosoma X ubicados a más de 10Mb de distancia (Li *et al.*, 2010), lo cual puede deberse al menor grado de recombinación que sufre el cromosoma X a nivel poblacional (ya que en los varones no recombina) (Schaffner, 2004). De todos modos es notable el hallazgo de dos pares de Indels ligados en la población de Buenos Aires, considerando que en las poblaciones urbanas de Misión Nueva Pompeya y Resistencia, tanto como en la población amerindia Wichí, se observó desequilibrio de ligamiento sólo para un par de estos X-Indels (Glesmann, 2015). Esto también podría responder a una subestructuración de la población bonaerense.

Entre los X-SNPs se encontraron ocho pares de marcadores ligados, contrastando con la cantidad obtenida para poblaciones nativas amerindias de Chaco, que fue de quince pares (Glesmann, 2015). Los SNPs ligados parecen estar definiendo un bloque de ligamiento en la zona citogenética Xq25, donde los marcadores se encuentran muy cercanos entre sí. En particular rs1781104-rs4825889-rs1781116-rs1937193 presentan los valores de ligamiento más significativos, por lo cual podría evaluarse la posibilidad de analizarlos en forma de haplotipo, dado que en un trabajo previo también se observó un marcado ligamiento entre estos marcadores (Glesmann, 2015). Asimismo, se obtuvieron desequilibrios de ligamiento entre X-Indels y X-SNPs, lo cual obedece a que éstos se encuentran en la misma zona cromosómica.

La relación nº haplotipos/nº hombres para los marcadores X-Indels (18/45=40% de diversidad haplotípica) fue menor que la obtenida por Glesmann (2015) en poblaciones nativas amerindias (48% y 52% para poblaciones Wichí y Mocoví, respectivamente). Esto si bien resulta inesperado en una población abierta al flujo migratorio desde distintos orígenes, es consistente con los resultados obtenidos del análisis de ligamiento. Se destaca que el haplotipo más frecuente de la población de Buenos Aires (alelo Inserción para todos los marcadores), también apareció con una frecuencia alta en la población de Resistencia, y no se encontró entre las poblaciones nativas amerindias de Chaco (Glesmann, 2015), siendo más probablemente un haplotipo de origen no nativo. Por otra parte, el número de haplotipos para los SNPs

(34/41=83% de diversidad haplotípica) fue elevado, e incluso considerando todos los marcadores en conjunto, ningún haplotipo total se repitió dentro del espacio muestral (41/41). Por lo tanto la región Xq25 presentaría bloques de ligamiento entre los marcadores que allí se encuentran, y esto podría aprovecharse para obtener estimaciones más precisas de mezcla interétnica en poblaciones que provienen de diferentes ancestrías, según lo propuesto por Resque *et al.* (2010), entre otros.

VARIACIÓN ENTRE LA POBLACIÓN DE BUENOS AIRES Y OTRAS POBLACIONES ARGENTINAS

En coincidencia con la variabilidad hallada en los bonaerenses, otras poblaciones urbanas han mostrado alta variabilidad para los marcadores analizados, en particular las de Resistencia, Viedma y Misión Nueva Pompeya (Bridi *et al.*, 2013; Glesmann, 2015). Por otra parte, es notable el contraste con lo hallado en poblaciones con muy poco flujo génico, como las comunidades nativas amerindias Wichí y Mocoví de la provincia de Chaco, en las que se notaron bajas frecuencias de los genotipos heterocigotas y altos porcentajes de homocigosis (Glesmann, 2015). Estos resultados son consistentes con las diferencias significativas reportadas en el cromosoma X entre Buenos Aires y las mismas poblaciones nativas chaqueñas, para marcadores X-STR (Glesmann *et al.*, 2013, Martina, 2006).

Gusmão *et al.* (2008) utilizaron un decaplex de X-STRs y no encontraron diferencias entre los bonaerenses y las poblaciones argentinas de Entre Ríos, Córdoba, Misiones y Río Negro. El análisis realizado en nuestro trabajo tampoco permitió diferenciar las poblaciones de Buenos Aires y Viedma (Río Negro), pero sí obtuvimos FSTs significativos con poblaciones de la provincia de Chaco. Esta diferencia bien podría deberse al alto porcentaje de pobladores nativos originarios que viven en esa provincia, ya que es cuatro veces mayor al que se encuentra en el resto de las provincias antes mencionadas (INDEC, Censo 2010) elevando en consecuencia el componente genético nativo americano de esas poblaciones. Además, el 58% de los pobladores que se reconocen como nativos o descendientes de nativos en esta provincia, viven en áreas urbanas (INDEC, Censo 2010), generando entre los chaqueños

una composición genética más miscegenada con poblaciones nativas. Esto último se condice tanto con datos históricos como con un estudio genético realizado por Avena y colaboradores (2012), en el que se encontró un componente nativo mayor tanto en el Noroeste argentino como en el Noreste argentino, en comparación con lo encontrado en la población bonaerense. Di Fabio Rocca y colaboradores (2015) aportaron datos adicionales acerca de la composición genética de la población de Resistencia, mostrando un 67,9% de componente mitocondrial nativo americano.

Los resultados del análisis de varianza molecular para el conjunto de X-Indels no fueron concluyentes, pero se nota una tendencia a la diferenciación entre las poblaciones de capitales no chaqueñas (entre las que se encuentra Buenos Aires), y las poblaciones urbanas chaqueñas. Esta tendencia podría eventualmente confirmarse utilizando un mayor número de marcadores en la comparación. Los resultados del AMOVA para los SNPs y para todos los marcadores en conjunto no parecen revelar diferencias entre los grupos armados, muy probablemente debido a la gran varianza que existe entre las poblaciones que componen cada grupo.

Los valores de F_{ST} fueron más altos al comparar a los bonaerenses con poblaciones nativas chaqueñas que con poblaciones urbanas de Argentina, para la mayor parte de los marcadores. Esta mayor diferenciación se debería a la ausencia de flujo génico de individuos de poblaciones urbanas con nativos y también por el reducido tamaño poblacional de estos últimos, en comparación con ciudades que siguen creciendo y recibiendo el aporte genético de otras poblaciones (Bridi *et al.*, 2013; Glesmann *et al.*, 2013).

Con respecto al análisis de la estructura poblacional, de ambos gráficos obtenidos en el Structure se entiende que la población de Buenos Aires es genéticamente más parecida a la de Resistencia que a las poblaciones nativas chaqueñas en cuanto al cromosoma X. Sin embargo también presenta patrones de variación particulares con una distribución dispersa, que la hacen diferente del resto de las poblaciones, incluso de Resistencia. Esto concuerda con los valores de F_{ST} obtenidos anteriormente, en el que se encontraron un menor número de marcadores

que pudieran diferenciar a las poblaciones de Buenos Aires y Resistencia, respecto de otras poblaciones con las que se la comparó.

VARIACIÓN ENTRE LA POBLACIÓN DE BUENOS AIRES Y OTRAS POBLACIONES DEL MUNDO

Al comparar los resultados obtenidos para la población de la provincia de Buenos Aires, con la información extraída de 1000 Genomes (<http://www.1000genomes.org>) obtuvimos diferencias en un mayor número de marcadores con la población de Perú, y, en menor medida, con la población Colombiana. Si bien éstas también son poblaciones americanas, esta diferenciación podría deberse a que esos países poseen tanto un componente autóctono como africano mayor al de los bonaerenses, según un estudio realizado por Avena y colaboradores (2006). Numéricamente, se ha detectado un componente autóctono en Lima (Perú) del 56%, mientras que en la ciudad de Buenos Aires se ha descrito un porcentaje de contribución indígena del 15,8 % (Avena *et al.*, 2006). La contribución africana es muy importante en ciudades en las que el tráfico de esclavos fue mucho más intenso, como en Barranquilla (Colombia), donde se obtuvo un porcentaje de 84%, mientras que en la ciudad de Buenos Aires se detectó sólo un 4,3% de componente africano (Avena *et al.*, 2006).

Bobillo y colaboradores (2009) llevaron a cabo un análisis de diferenciación entre Argentina, Portugal y España, utilizando diez marcadores del tipo STR del cromosoma X. Sus resultados muestran más similitudes que diferencias con estos países europeos, no habiéndose encontrado diferencias únicamente en dos y cuatro X-STR con Portugal y España, respectivamente. En nuestro trabajo, se obtuvieron diferencias con España también en cuatro (SNPs) de los diez marcadores comparados, y en dos (SNPs) de los diez con Italia. De esta manera, si bien los STR presentan una tasa mutacional notablemente mayor (Kaessmann *et al.*, 2002; Katoh *et al.*, 2002; Xu & Fu, 2004) ésta no estaría afectando los resultados, ya que se obtuvo lo mismo comparando con los marcadores de secuencia única analizados en el presente trabajo, de tasa mutacional más baja.

VARIABILIDAD DE INDELS Y SNPS

Dentro del conjunto de X-Indels analizados, podemos destacar a los marcadores MID 193, MID 3754 y MID 1705 como los más informativos desde el punto de vista poblacional, mostrando diferencias entre poblaciones urbanas y amerindias. En contraste, los marcadores MID 3756 y MID 1540 no parecen presentar frecuencias diferentes entre las poblaciones comparadas, ya que mostraron similitud, con excepción del MID 1540 al comparar poblaciones africanas con poblaciones no africanas. De esta manera, estos últimos dos Indels tendrían menor utilidad para estudios poblacionales.

De acuerdo con datos bibliográficos relevados en otras partes del mundo (<http://www.1000genomes.org>) el Indel que sugiere con más claridad una ancestría europea en la población de Buenos Aires es el MID 193, por lo tanto podría ser de utilidad en estudios poblacionales para diferenciar grupos de individuos de distinta procedencia.

Por otro lado, los X-SNPs que mostraron una marcada similitud con poblaciones europeas fueron rs149910 y rs933315 y una diferencia significativa con poblaciones latinoamericanas (México, Columbia y Perú). El primero no debería ser utilizado en estudios poblacionales sin antes evaluar en profundidad las causas reales del desajuste al equilibrio Hardy-Weinberg en ese locus. Glesmann (2015) propuso el alelo T de este marcador como uno de los alelos que conformaría un haplotipo más probablemente perteneciente a poblaciones nativas originarias chaqueñas y se observó una gran prevalencia del alelo G en poblaciones europeas, en contraste con las poblaciones americanas (<http://www.1000genomes.org>). Sin embargo en este trabajo no se halló dicha tendencia.

Finalmente, los marcadores rs1991961 y rs1781104 aportan menos información para la diferenciación de los bonaerenses con otras poblaciones, ya sea a nivel nacional o mundial. Estos marcadores, junto a los MID 3756 y MID 1540, son buenos candidatos a ser usados en identificación humana (ciencias forenses y evaluación de parentescos) ya que presentan frecuencias muy similares en las poblaciones. Por lo tanto en el futuro será muy útil determinar los parámetros necesarios para evaluar

su posible aplicación en esta disciplina (Poder de Exclusión y Poder de Discriminación) (Fung & Hu, 2008). El Poder de Exclusión de un marcador utilizado para casos de paternidad, es una forma de caracterizar su habilidad para excluir a un falso padre, mientras que el Poder de Discriminación es una medida de la capacidad que tiene ese locus para distinguir a un individuo de otro. Ambos parámetros están sumamente relacionados.

Algunos trabajos que abordan el estudio del cromosoma X han utilizado un número mayor de marcadores (Li *et al.*, 2008; Tomas *et al.*, 2008; Edelman *et al.*, 2009; Casto *et al.*, 2010; Freitas *et al.*, 2010; Resque *et al.*, 2010; Pereira *et al.*, 2011; Pereira *et al.*, 2012; Ibarra *et al.*, 2014;). En la mayoría de ellos se utilizaron métodos computacionales para evaluar información obtenida en bases de datos, recurriendo a técnicas de secuenciación o reacciones de múltiplex de alto rendimiento que les permitieron genotipificar un gran número de polimorfismos en una sola reacción. Si bien en este trabajo se utilizó un número menor de marcadores, la información obtenida es de utilidad como una aproximación inicial en el área, mostrando resultados consistentes con la historia demográfica bonaerense.

Queda pendiente en el futuro, estudiar la diversidad del cromosoma X de los bonaerenses utilizando un mayor número de marcadores para optimizar este análisis.

CONCLUSIONES

- La población bonaerense es muy diversa genéticamente en la zona del cromosoma X estudiada en este trabajo, dados los altos valores de heterocigosis, el alto número de haplotipos, y otros parámetros poblacionales.

- No se evidenciaron fuerzas de cambio poblacional marcadas que estén afectando a la variación de estos marcadores del cromosoma X.

- Los resultados arrojan similitudes genéticas importantes en el cromosoma X con las poblaciones europeas, y diferencias genéticas notables con poblaciones nativas originarias de Argentina, lo cual concuerda con los antecedentes históricos poblacionales de la población bonaerense.

- Se propone analizar en un futuro la región que incluye los SNPs rs1781104, rs4825889, rs1781116 y rs1937193 para un análisis de haplotipos en tríos madre-padre-hija o dúos madre-hijo, a fin de evaluar en mayor profundidad el grado de ligamiento entre estos marcadores.

- Los marcadores que son de mayor aplicación para estudios poblacionales son MID 3754, MID 193, rs149910 y rs933315, mientras que los marcadores rs1991961, rs1781104, MID 3756 y MID 1540 podrían tener aplicación en ciencias de identificación, y queda pendiente su evaluación.

BIBLIOGRAFÍA

- Ajmone Marsan P. & Oldenbroek K. 2010. Marcadores moleculares: una herramienta para explorar la diversidad genética. En: La situación de los recursos zoogenéticos mundiales para la alimentación y la agricultura. Comisión de Recursos Genéticos para la Alimentación y la Agricultura. Organización de las Naciones Unidas para la Agricultura y la Alimentación (FAO). Roma, Italia. 555 pp. Disponible en: <http://www.fao.org/docrep/012/a1250s/a1250s00.htm>. Último acceso: febrero de 2016.
- Aranguren-Méndez J. A., Román-Bravo R., Isea W., Villasmil Y. & Jordana J. 2005. Los microsatélites (STRs), marcadores moleculares de ADN por excelencia para programas de conservación: una revisión. Asociación Latinoamericana de Producción Animal. 13(1):30-42.
- Avena S., Goicoechea A., Rey J., Dugoujon J., Dejean C. & Carnese F. 2006. Gene mixture in a population sample from Buenos Aires City. *Medicina* (Buenos Aires) 66(2): 113-118.
- 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., González Burchard E., Parolin M. L., Goicoechea A., Acreche N., Boquet M., Ríos Part M. de 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(4): e34695. DOI:10.1371/journal.pone.0034695.
- Bobillo C., Sala A., Gusmão L. & Corach D. 2009. Genetic analysis of 10 X-STRs in Argentinian population. *Forensic Science International: Genetics*. 5(1):14-16.
- Bridi J. R., Di Santo Meztler G. P., Raggio M. C. & Catanesi C. I. 2013. Variación de 5 InDels de cromosoma X en la población de Viedma. XI Jornadas Nacionales de Antropología Biológica (póster). Ciudad Autónoma de Buenos Aires.
- Brinkmann B., Klitsch M., Neuhuber F., Hühne J. & Rolf B. 1998. Mutation Rate in Human Microsatellites: Influence of the Structure and Length of the Tandem Repeat. *American Journal of Human Genetics*. 62:1408-1415.
- Casto A. M., Li J. Z., Absher D., Myers R., Ramachandran S. & Feldman M. W. 2010. Characterization of X-Linked SNP genotypic variation in globally distributed human populations. *Genome Biology*. 11:R10. DOI: 10.1186/gb-2010-11-1-r10.

- Checa Caratachea M. A.** 2007. Polimorfismos genéticos: importancia y aplicaciones. *Revista del Instituto Nacional de Enfermedades Respiratorias*. México. 20(3):213-221.
- Di Fabio Rocca F., Glesmann L., Patiño J., Dejean C., Carnese F. R., Catanesi C. I., Avena S.** 2015. Determinación de linajes uniparentales en una muestra poblacional de la ciudad de Resistencia, Chaco. XII Jornadas Nacionales de Antropología Biológica (póster). Corrientes, Argentina.
- Edelmann J., Hering S., Augustin C. & Szibor R.** 2009. Indel polymorphisms—An additional set of markers on the X-chromosome. *Forensic Science International: Genetics Supplement Series* 2:510–512.
- Excoffier L. & Lischer H. E. L.** 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*. 10: 564-567.
- Floria C. A. & García Belsunce C. A.** 1992 *Historia de los Argentinos*. Ed. Larousse. Buenos Aires. 1001 pp.
- Freitas N. S. C., Resque R. L., Ribeiro-Rodrigues E. M., Guerreiro J. F., & Ney P. C. Santos., Ribeiro-dos-Santos A. & Santos S.** 2010. X-linked insertion/deletion polymorphisms: forensic applications of a 33-markers panel. *International Journal of Legal Medicine*. 124:589–593.
- Fung W. K. & Hu Y-Q.** 2008. *Statistical DNA forensics. Theory, methods and computation*. Ed. John Wiley & sons. San Francisco, USA. 239 pp.
- Garbus I., Gómez M. & Echenique V.** 2010 *Herramientas básicas de ingeniería genética*. En: *Biotecnología y Mejoramiento Vegetal II*. Levitus G., Echenique V., Rubinstein C., Hopp E., Mroginski L., (Eds.). Ediciones Instituto Nacional de Tecnología Agropecuaria (INTA). 647 pp.
- Glesmann L.** 2015. *Análisis de polimorfismos genéticos de cromosoma X y búsqueda de marcadores étnicos chaqueños (Tesis doctoral)*. Facultad de Ciencias Naturales y Museo, Universidad Nacional de La Plata, La Plata. 138pp.
- Glesmann L. A., Martina P. F. & Catanesi C. I.** 2013. Genetic variation of X-STRs in the Wichí population from Chaco province, Argentina. *Human Biology* .85(5):687-98
- Griffiths A. J. F., Miller J. H., Suzuki D. T., Lewontin R. C. & Gelbart W. M.** 2002. *Genética*. Ed McGraw-Hill Interamericana. España. 845pp.

- Gusmão L., Sánchez-Diz P., Alves C., Gomes I., Zarrabeitia M. T., Abovich M., Atmetlla I., Bobillo C., Bravo L., Builes J., Cainé L., Calvo R., Carvalho E., Carvalho M., Cicarelli R., Catelli L., Corach D., Espinoza M., García O., Malaghini M., Martins J., Pinheiro F., Porto M. J., Raimondi E., Riancho J. A., Rodríguez A., Rodríguez A., Rodríguez Cardozo B., Schneider V., Silva S., Tavares C., Toscanini U., Vullo C., Whittle M., Yurrebaso I., Carracedo A. & Amorim A. 2008. A GEP-ISFG collaborative study on the optimization of an X-STR decaplex: data on 15 Iberian and Latin American populations. *International Journal of Legal Medicine*. 123(3):227-234.
- Hedrick P. 2000. *Genetics of populations*. Jones and Bartlett Publishers International. USA. 553pp.
- Ibarra A., Restrepo T., Rojas W., Castillo A., Amorim A., Martínez B., Burgos G., Ostos H., Álvarez K., Camacho M., Suarez Z., Pereira R. & Gusmão L. 2014. Evaluating the X Chromosome-Specific Diversity of Colombian Populations Using Insertion/Deletion Polymorphisms. *PloS One*. 9(1):e87202 DOI: 10.1371/journal.pone.0087202
- International Human Genome Sequencing Consortium. 2004. Finishing the euchromatic sequence of the human genome. *Nature*. 431:931-945.
- Kaessmann H., Zöllner S., Gustafsson A. C., Wiebe V., Laan M., Lundeberg J., Uhlén M. & Pääbo S. 2002. Extensive Linkage Disequilibrium in Small Human Populations in Eurasia. *The American Journal of Human Genetics*. 70:673-685.
- Katoh T., Mano S., Ikuta T., Munkhbat B., Tounai K., Ando H., Munkhtuvshin N., Imanishi T., Inoko H., Tamiya G. 2002. Genetic Isolates in East Asia: A Study of Linkage Disequilibrium in the X Chromosome. *American Journal of Human Genetics*. 71:395-400.
- Keinan A., Mullikin J. C., Patterson N. & Reich D. 2009. Accelerated genetic drift on chromosome X during the human dispersal out of Africa. *Nature Genetics*. 41(1): 66-70.
- Lewin B. (Ed.) 2008. *Genes IX*. Editorial McGraw-Hill. México. 844pp.
- Li J. Z., Absher D. M., Tang H., Southwick A. M., Casto A. M., Ramachandran S., Cann H. M., Barsh G. S., Feldman M., Cavalli-Sforza L. L. & Myers R. M. 2008. Worldwide Human Relationships Inferred from Genome-Wide Patterns of Variation. *Science*. 319:1100-1104

- Li L., Li C., Zhang S., Zhao S., Liu Y. & Lin Y. 2010. Analysis of 14 highly informative SNP markers on X chromosome by TaqMan SNP genotyping assay. *Forensic Science International: Genetics*. 4:145-148.
- Lyon M. F. 1961. Gene action in the X-chromosome of the mouse. *Nature*. 190:372-373.
- Martina P.F. 2006. Estudio de marcadores microsatélites del cromosoma X en poblaciones humanas (Tesis de grado). Facultad de Ciencias Exactas, Químicas y Naturales, Universidad Nacional de Misiones. Posadas, Misiones. 74pp.
- Martínez M. C., Helguera M. & Carrera A. 2010. Marcadores moleculares. En: *Biotecnología y Mejoramiento Vegetal II*. Levitus G., Echenique V., Rubinstein C., Hopp E., Mroginski L., (Eds.). Ediciones Instituto Nacional de Tecnología Agropecuaria (INTA). 647pp.
- Mills R. E., Luttig C., Larkins C. T., Larkins C. E., Beauchamp A., Tsui C., Pittard S. W. & Devine S. E. 2006. An initial map of insertion and deletion (INDEL) variation in the human genome. *Cold Spring Harbor Laboratory Press*. 16:1182-1190.
- Pereira R., Pereira V., Gomes I., Tomas C., Morling N., Amorim A., Prata M. J., Carracedo A. & Gusmão L. 2012. A method for the analysis of 32 X chromosome insertion deletion polymorphisms in a single PCR. *International Journal of Legal Medicine*. 126:97-105
- Pereira V., Tomas C., Amorim A., Morling N., Gusmão L. & Prata M. J. 2011. Study of 25 X-chromosome SNPs in the Portuguese. *Forensic Science International: Genetics*. 5:336-338.
- Pierce B. A. 2010. *Genética: un enfoque conceptual*. Editorial médica panamericana. España. 832pp.
- Pritchard J. K., Stephens M. & Donnelly P. 2000. Inference of Population Structure Using Multilocus Genotype Data. *Genetics*. 155: 945-959.
- Reich D. E., Cargill M., Bolk S., Ireland J., Sabeti P. C., Richter D. J., Lavery T., Kouyoumjian R., Farhadian S. F., Ward R. & Lander E. S. 2001. Linkage disequilibrium in the human genome. *Nature*. 411:199-204.
- Resque R. L., Freitas N. S. C., Ribeiro-Rodrigues E. M., Guerreiro J. F., Santos N. P. C., Ribeiro-dos-Santos A., Zago M. A. & Santos S. 2010. Estimates of Interethnic Admixture in the Brazilian Population Using a Panel of 24 X-Linked Insertion/Deletion Markers. *American Journal of Human Biology*. 22(6):849-852.

- Schaffner S. F.** 2004. The X chromosome in population genetics. *Nature Reviews Genetics*. 5:43-51.
- Solá M. D.** 2009. *Aborígenes Argentinos*. Editorial Gradifco. Buenos Aires, Argentina. 272pp.
- Szibor R.** 2007. X-chromosomal markers: Past, present and future. *Forensic Science International Genetics*. 1: 93-99.
- Szibor R., Krawczak M., Hering S., Edelman J., Kuhlisch E. & Krause D.** 2003. Use of X-linked markers for forensic purposes. *International Journal of Legal Medicine*. 117:67-74.
- Tomas C., Sanchez J. J., Barbaro A., Brandt-Casadevall C., Hernandez A., Dhiab M. B., Ramon M. & Morling N.** 2008. X-chromosome SNP analyses in 11 human Mediterranean populations show a high overall genetic homogeneity except in North-west Africans (Moroccans). *BMC Evolutionary Biology*. 8:75.
- Weber J. L. & Wong C.** 1993. Mutation of human short tandem repeats. *Human Molecular Genetics*. 2(8):1123-1128.
- Weber J. L., David D., Heil J., Fan Y., Zhao C. & Marth G.** 2002. Human Diallelic Insertion/Deletion Polymorphisms. *The American Journal of Human Genetics*. 71:854-862.
- Xu H. & Fu Y-X.** 2004 Estimating effective population size of mutation rate with microsatellites. *Genetics*. 166: 555-563.

REFERENCIAS ELECTRÓNICAS

- Encuesta Complementaria de Pueblos Indígenas (ECPI) 2004-2005, Complementaria del Censo Nacional de Población, Hogares y Viviendas 2001. 2004-2005. http://www.indec.gov.ar/micro_sitios/webcenso/ECPI/pueblos/ampliada_index_nacionales.asp?mode=00. Último acceso: febrero de 2016.
- GraphPad Prism 6.07. <http://www.graphpad.com/>. Último acceso: enero de 2016.
- INDEC: Instituto Nacional de Estadísticas y Censos. <http://www.indec.gov.ar>. Último acceso: enero de 2016.
- Sitio Web 1000 Genomes. <http://www.1000genomes.org>. Último acceso: febrero de 2016.

APÉNDICE

APÉNDICE 1. Carta de información y formulario de consentimiento para el proyecto.

CARTA DE INFORMACIÓN PARA LOS PARTICIPANTES DE “ANÁLISIS DE POLIMORFISMOS GENÉTICOS DE CROMOSOMA X”

Lo invitamos a participar en una investigación que tiene por objetivo estudiar ciertas regiones de su ADN que son variables entre los individuos de una población.

En las poblaciones humanas los rasgos genéticos se van transmitiendo de una generación a otra. Algunas tienen un efecto visible, como en el color de ojos, mientras que otras tienen un efecto que no es visible directamente, y para analizarlas se requieren técnicas de laboratorio.

En el laboratorio de Diversidad Genética del IMBICE (CICPBA-CONICET-UNLP) utilizamos técnicas de laboratorio molecular para estudiar la variación de distintos rasgos heredables que no tienen ningún efecto en las personas que los presentan.

Para ello, le solicitamos su consentimiento para donar una pequeña muestra de saliva.

Su participación no implica riesgos físicos ni morales para su persona, y su muestra será tratada de forma completamente anónima y confidencial. No existen beneficios para el donante, sino que el beneficio directo de este estudio es para la comunidad local, por el aporte al conocimiento de la variación genética de su población.

Su participación en esta investigación es absolutamente voluntaria, y Ud. puede abandonar el estudio en cualquier momento si así lo desea. Si Ud. acepta participar, tenga a bien completar y firmar el consentimiento adjunto. La persona responsable de tomar la muestra, le informará sobre el proyecto y le aclarará todas sus dudas, respondiendo claramente a sus preguntas.

Teléfono de contacto: 0221-421-0112

Correo electrónico de contacto: ccatanesi@imbice.gov.ar

FORMULARIO DE CONSENTIMIENTO INFORMADO “ANÁLISIS DE POLIMORFISMOS GENÉTICOS DE CROMOSOMA X”

Yo,..... (Nombres y apellidos completos), de..... Años (edad), en calidad de donante de muestra biológica, he leído la carta de información “ANÁLISIS DE POLIMORFISMOS GENÉTICOS Y BÚSQUEDA DE MARCADORES ÉTNICOS DE CROMOSOMA X” y he podido hacer preguntas sobre la investigación a..... (Responsable de la toma de muestra). Declaro haber entendido la carta, comprendiendo que mi participación es voluntaria, gratuita, anónima y confidencial. Comprendo también que puedo retirarme de la investigación cuando quiera sin tener que dar explicaciones, y que puedo contar con los datos obtenidos si lo requiero. Por ello, presto libre conformidad para que se tome una muestra de saliva de mi persona, para estudiar rasgos genéticos con fines poblacionales.

Marque lo que corresponde:

- Al término del proyecto, SÍ - NO autorizo a que se almacene la muestra de ADN para estudios futuros que cuenten con el aval de un Comité de Ética.

Lugar y fecha:

Firma y aclaración del donante:

Firma y aclaración del responsable de la toma de muestra:
.....

Firma y aclaración del investigador responsable del proyecto:
.....

Firma y aclaración de un testigo:

APÉNDICE 2. Secuencias de los *primers* utilizados para las reacciones de amplificación de los X-SNPs. En negrita se muestran las bases que difieren en el par de *primers* alelo específicos.

Marcador	Secuencia <i>primer Forward</i> (5- 3)	Secuencia <i>primer Reverse</i> (5- 3)
rs318173	GGTTTCTGTCTATGACAATTC	<ul style="list-style-type: none"> CTCTGGGTATATTCAGAAGCA CTCTGGGTATATTCAGAAGCC
rs4825889	TCCCTCTTTGGCTCCCCTAAA	<ul style="list-style-type: none"> ATGAACCGAGTTTTGATTTGGAAT ATGAACCGAGTTTTGATTTGAAAC
rs933315	<ul style="list-style-type: none"> GCTGGGAAGACCAAAAAACAAGCAA GCTGGGAAGACCAAAAAACAATCAG 	CAATGCCCCACAGAGATTTTAGTG
rs1991961	ACGGAGAACCATGTGCTTAGCA	<ul style="list-style-type: none"> TCTGACTAAGTGGAGACAACCG TCTGACTAAGTGGAGACACCCT
rs149910	GGGCTCATTGCAGATGTTGAC	<ul style="list-style-type: none"> ATCCATTGTTTCATTTCAAGATATC ATCCATTGTTTCATTTCAAGCTATA
rs1937193	GGGACTTTCTAGCATGACTGT	<ul style="list-style-type: none"> GCCTCAAACCTCTTTCTTCTCAT GCCTCAAACCTCTTTCTTCACAA
rs1781116	<ul style="list-style-type: none"> GCAAATATGTTGGCAAGACATAC GCAAATATGTTGGCAAGACGTAT 	CCCTTCCCCACCTTGGATAG
rs652	<ul style="list-style-type: none"> CTGTCCCACACCTACAGTGCG CTGTCCCACACCTACATTGCT 	ACTGGTACCTGTGCATGCATTT
rs1781104	<ul style="list-style-type: none"> GTCTGTCTTGTCCAATGCCGAGA GTCTGTCTTGTCCAATGCCTAGC 	GGCAAATTGACTGGAAGCCA
rs1299087	TTTTAGAGGGATTCAGACACAAA	CGCTTTGCAATTGGTTGTAT

APÉNDICE 3. Secuencias de los *primers* utilizados para la amplificación de los X-indels, números de identificación y alelos.

Marcador	Nº rs	Alelos	Secuencia <i>primer Forward</i> (5'-3')	Secuencia <i>primer Reverse</i> (5'-3')
MID 1540	rs2307736	-/TTCCAACCAGCAGCAGGGTT	CTGTCAAGTCCAAATCAAGCAG	TTATACTTCCAAACAGCCCACA
MID 1705	rs2307901	-/TAGA	TCATCCATGTTTCTGGGTAT	GAAGCAAGAATGTGTTGTCAT
MID 3754	rs57608175	-/GGTCATCACGAG	GATCCTTTGGTTTGCACAATG	CCCTGCTCTAATTCTCCCACT
MID 3756	rs72270198	-/ TGGCAGAGCCATT	CAGCGTTCATATCAGCACTCA	TCCGCATCACCAGTTAAAGAC
MID 193	rs16632	-/AATA	TTTGAAAACACAGAAATGCA	CTGAAGTGACCATTGACCTC

APÉNDICE 4. Haplotipos encontrados entre los 41 hombres muestrales para los X-SNPs. Total: 34 haplotipos.

rs652	rs318173	rs1299087	rs933315	rs1991961	rs149910	rs1937193	rs4825889	rs1781116	rs1781104	Nº individuos	Frec.
G	G	A	A	T	G	A	G	T	C	3	0.0732
T	G	G	A	T	T	A	G	T	C	1	0.0244
T	T	A	A	T	G	A	G	T	C	1	0.0244
G	T	G	A	T	G	A	G	T	C	3	0.0732
G	G	G	A	T	G	T	A	C	C	1	0.0244
T	T	G	G	G	G	A	G	T	C	1	0.0244
T	T	G	A	G	T	A	G	T	C	2	0.0488
T	T	G	G	T	T	A	G	T	C	1	0.0244
G	T	A	A	T	G	T	A	C	C	1	0.0244
T	T	A	A	G	T	A	G	T	C	1	0.0244
T	T	A	A	G	G	A	G	T	C	1	0.0244
G	T	A	A	T	G	A	G	T	C	2	0.0488
G	G	A	A	G	G	T	A	T	C	1	0.0244
G	T	A	G	G	G	A	G	T	C	1	0.0244
G	G	G	A	G	G	T	A	C	A	1	0.0244
T	T	A	G	T	G	T	A	C	A	1	0.0244
G	T	G	G	T	T	A	G	T	C	1	0.0244

Referencias: Frec= frecuencia del haplotipo.

APÉNDICE 4. Haplotipos encontrados entre los 41 hombres muestrales para los X-SNPs. Total: 34 haplotipos. (Continuación)

rs652	rs318173	rs1299087	rs933315	rs1991961	rs149910	rs1937193	rs4825889	rs1781116	rs1781104	Nº individuos	Frec.
G	T	A	G	T	G	A	G	T	C	2	0.0488
G	G	G	G	G	G	T	A	C	A	1	0.0244
T	T	G	G	T	G	A	G	T	C	1	0.0244
G	G	A	A	G	G	A	G	T	C	1	0.0244
G	T	G	A	G	G	A	G	T	C	1	0.0244
G	T	G	G	G	G	A	G	T	C	1	0.0244
G	T	G	G	G	G	T	A	C	A	1	0.0244
G	T	A	G	G	T	T	A	C	A	1	0.0244
T	G	A	A	G	T	A	G	T	C	1	0.0244
G	T	G	A	G	G	T	A	C	A	1	0.0244
G	T	A	A	G	G	T	A	C	C	1	0.0244
G	T	G	G	G	T	T	A	C	A	1	0.0244
G	T	A	G	T	T	T	A	C	A	1	0.0244
G	G	A	A	G	G	T	A	C	A	1	0.0244
T	T	A	A	T	T	T	A	C	A	1	0.0244
G	T	G	A	G	T	T	A	C	A	1	0.0244
G	T	G	G	T	G	A	G	T	C	1	0.0244

Referencias: Frec= frecuencia del haplotipo.

APÉNDICE 5. Haplotipos de Indels obtenidos a partir de los 45 hombres. Total: 18 haplotipos.

MID 1540	MID 1705	MID 3754	MID 3756	MID 193	Nº individuos	Frecuencia Haplotípica
1	1	1	0	0	3	0.0667
1	1	0	0	0	3	0.0667
1	0	1	1	0	2	0.0444
1	1	0	1	1	3	0.0667
1	1	0	0	1	2	0.0444
0	1	0	1	1	4	0.0889
1	0	1	0	0	2	0.0444
1	1	1	1	1	6	0.1333
0	1	0	1	0	2	0.0444
0	1	1	0	0	4	0.0889
1	1	1	0	1	1	0.0222
0	1	0	0	1	2	0.0444
0	0	0	1	0	1	0.0222
0	1	1	1	1	5	0.1111
0	0	0	1	1	1	0.0222
0	1	0	0	0	2	0.0444
1	1	1	1	0	1	0.0222
1	0	0	1	1	1	0.0222

APÉNDICE 6. Datos obtenidos de la página Web 1000 Genomes, de porcentajes de frecuencias alélicas de los diferentes X-SNPs. Los datos “México” pertenecen a la ciudad de Los Ángeles, California. Los datos de España pertenecen a la población ibérica de ese país. Los datos de Italia pertenecen a la región de Toscana.

Marcador	Alelos	África	Colombia	Perú	México	España	Italia
rs1299087	A	9	37	26	42	47	47
	G	91	63	74	58	53	53
rs149910	G	42	61	36	44	82	71
	T	58	39	64	56	18	29
rs933315	A	37	50	29	50	77	70
	G	63	50	71	50	23	30
rs4825889	A	79	58	35	51	74	60
	G	21	42	65	49	26	40
rs1991961	G	32	37	41	45	40	37
	T	68	63	59	55	60	63
rs652	G	17	66	57	59	76	82
	T	83	34	43	41	24	18
rs1937193	A	83	52	68	50	28	40
	T	17	48	32	50	73	60
rs1781104	A	35	37	32	45	55	45
	C	65	63	68	55	45	55

APÉNDICE 7. Datos obtenidos de la página Web 1000 Genomes, de porcentajes de frecuencias alélicas de los X-Indels MID 1540 Y MID 193. Los datos “México” pertenecen a la ciudad de Los Ángeles, California. Los datos de España pertenecen a la población ibérica de ese país. Los datos de Italia pertenecen a la región de Toscana.

Marcador	Alelos	África	Colombia	Perú	México	España	Italia
MID 1540	I	97	66	64	57	54	54
	D	3	34	36	43	46	46
MID 193	I	91	64	33	48	75	69
	D	9	36	67	52	25	31

APÉNDICE 8. Haplotipos totales obtenidos a partir de la muestra masculina de Buenos Aires (n=41). Cada hombre presentó un haplotipo diferente, y la frecuencia haplotípica resultó de 0.0244. Para los Indels, el 1 representa la inserción y el 0 representa la delección.

rs652	rs318173	rs1299087	rs933315	rs1991961	rs149910	rs1937193	rs4825889	rs1781116	rs1781104	MID 1540	MID 1705	MID 3754	MID 3756	MID 193
G	G	A	A	T	G	A	G	T	C	1	1	1	0	0
T	G	G	A	T	T	A	G	T	C	1	1	0	0	0
T	T	A	A	T	G	A	G	T	C	1	0	1	1	0
G	T	G	A	T	G	A	G	T	C	1	1	0	1	1
G	G	G	A	T	G	T	A	C	C	1	1	0	0	1
T	T	G	G	G	G	A	G	T	C	0	1	0	1	1
T	T	G	A	G	T	A	G	T	C	1	0	1	1	0
T	T	G	G	T	T	A	G	T	C	1	0	1	0	0
G	T	A	A	T	G	T	A	C	C	1	1	1	1	1
T	T	A	A	G	T	A	G	T	C	0	1	0	1	0
T	T	A	A	G	G	A	G	T	C	0	1	0	1	1
G	T	A	A	T	G	A	G	T	C	0	1	1	0	0
G	G	A	A	G	G	T	A	T	C	1	1	1	0	1
G	T	A	G	G	G	A	G	T	C	0	1	1	0	0
G	G	G	A	G	G	T	A	C	A	1	1	1	1	1
T	T	A	G	T	G	T	A	C	A	0	0	0	1	0
G	T	G	G	T	T	A	G	T	C	0	1	0	1	0
G	T	A	G	T	G	A	G	T	C	1	1	1	1	1
G	G	G	G	G	G	T	A	C	A	0	1	1	1	1
T	T	G	G	T	G	A	G	T	C	0	0	0	1	1
G	G	A	A	G	G	A	G	T	C	0	1	1	1	1

APÉNDICE 8. Haplotipos totales obtenidos a partir de la muestra masculina de Buenos Aires (n=41). Ningún haplotipo se repitió, y la frecuencia haplotípica resultó de 0.0244. Para los Indels, el 1 representa la inserción y el 0 representa la delección. (Continuación)

rs652	rs318173	rs1299087	rs933315	rs1991961	rs149910	rs1937193	rs4825889	rs1781116	rs1781104	MID 1540	MID 1705	MID 3754	MID 3756	MID 193
G	T	G	A	G	G	A	G	T	C	0	1	1	1	1
G	T	G	G	G	G	A	G	T	C	0	1	0	0	1
G	T	G	G	G	G	T	A	C	A	0	1	1	1	1
G	T	A	G	G	T	T	A	C	A	0	1	0	0	0
G	T	G	A	T	G	A	G	T	C	1	1	1	0	0
G	T	G	A	T	G	A	G	T	C	1	1	0	0	0
T	G	A	A	G	T	A	G	T	C	1	0	1	0	0
G	T	G	A	G	G	T	A	C	A	1	1	1	0	0
G	T	A	A	T	G	A	G	T	C	1	1	0	1	1
G	T	A	A	G	G	T	A	C	C	1	1	1	1	1
G	G	A	A	T	G	A	G	T	C	1	1	1	1	1
G	T	G	G	G	T	T	A	C	A	0	1	1	0	0
G	T	A	G	T	G	A	G	T	C	0	1	0	0	0
G	G	A	A	T	G	A	G	T	C	0	1	0	1	1
G	T	A	G	T	T	T	A	C	A	0	1	1	0	0
G	G	A	A	G	G	T	A	C	A	1	1	0	0	1
T	T	A	A	T	T	T	A	C	A	1	1	0	1	1
G	T	G	A	G	T	T	A	C	A	1	1	1	1	1
G	T	G	G	T	G	A	G	T	C	1	1	0	0	0
T	T	G	A	G	T	A	G	T	C	1	0	0	1	1

